Project Plan for Digitization

A Structural Manual for Policy Development at the Greater West Bloomfield Historical Society

Jeff Nash, Sara Sterkenburg, Paul Wentzell

December 13, 2011

Written as a partial requirement of SI 644: Advanced Preservation
University of Michigan, School of Information

Table of Contents

EXECUTIVE SUMMARY **2**

SECTION 1: CONTEXT FOR PROJECT **2**

Section 1.1: Preservation Needs Assessment Findings 2

Section 1.2: Collection Importance and Uses 4

Section 1.3: Risks and Cost-Benefit Analysis of Digitization 4

SECTION 2: PROJECT DESCRIPTION **6**

Section 2.1: Project Goals 7

Section 2.2: Selection 9

SECTION 3: PREPARATION FOR DIGITIZATION **12**

Section 3.1: Labor 12

Section 3.2: Equipment 16

Section 3.3: Training 16

Section 3.4: Location and Facility 16

Section 3.5: Budget 16

Section 3.6: Timeline 17

Section 3.7: Reflecting on the Value Added by Digitization 18

SECTION 4: TECHNICAL STANDARDS AND QUALITY ASSURANCE **18**

Section 4.1: Digital Object Control and Standardization 18

Section 4.1.1: Hardware *19*

Section 4.1.2: Software *20*

Section 4.1.3: Image Capture Parameters *20*

Section 4.1.4: Quality Control *21*

Section 4.1.5: Metadata and Digital Object Description *22*

Section 4.2: Digital Object Management 23

Section 4.2.1: Data Management Personnel *24*

Section 4.2.2: Backing-Up Data *24*

Section 4.2.3: Digital Storage *25*

Section 4.2.4: Quality Assurance *25*

Section 4.2.5: Metadata Migration *26*

Section 4.2.6: Image Management *27*

Section 4.2.7: Managing Collection Scope *27*

SECTION 5: POLICY REEVALUATION **28**

CONCLUSION **30**

APPENDICES **32**

BIBLIOGRAPHY **37**

Executive Summary

After conducting an initial *Needs Assessment Survey* with the Greater West Bloomfield Historical Society (GWBHS) earlier this year, we chose to follow up on our recommendations and develop a secondary document: a detailed project plan that could aid the society in a collection digitization project. The plan that follows will aim to do two things: 1.) establish the necessity of in-depth policy development prior to actual digitization, and 2.) provide recommendations on best-practice metrics that should be followed to successfully allocate resources to the project (including employees, volunteers and interns, finances, facilities, possible vendors, and equipment) and create preservation-quality images that will both provide increased access to the GWBHS collections, and maintain the intellectual value of the museum's cultural artifacts for years to come.

In the sections that follow, we will establish reasons for electing to draft a plan for digitization, discuss goals of the effort (and their alignment with major missions of the GWBHS at large), and underline the importance of selecting appropriate and in-demand content for various phases of digitization. Later our discussion will move toward preparation requirements, technological equipment, metrics necessary for digitizing consistent and standard images, and how to manage digitized data for the long-term. Finally, we will detail the importance of frequent reevaluation of digitization policies. Our hope is that this plan will provide a starting point from which the GWBHS will be able to begin an effort with the ability to reach new audiences, gain control over its collections, and provide culturally relevant material to its community for the long-term.

Section 1: Context for Project

Section 1.1: Preservation Needs Assessment Findings

Our initial preservation assessment survey for the Greater West Bloomfield Historical Society (GWBHS) revealed a fragile environment for the long-term preservation of local historic artifacts and documents. In our initial assessment, we evaluated the condition of the building environment for collection storage, the disaster preparedness of the organization, preservation selection and other policies, and the physical handling of the collections. Our goal in the assessment was to provide feedback on the overall preservation environment, and offer suggestions to improve the overall storage and exhibit environments. As a direct result of this needs assessment survey, and in listening to the museum in its desire to

expand digital access, we have decided to draft a project plan for digitization to help the museum preserve its collections in the face of a fragile preservation environment.

According to our initial assessment, the building environment suffers from many preservation risks that threaten the longevity of the society's collections. The main basement storage area is the most notable risk area, where artifacts including historic tax records are exposed to insects, moisture and mold, and exposed plumbing. The underground storage area also raises concern for possible flooding scenarios, especially considering the GWBHS building sits very near a lake. In the main exhibit area, photographs, documents and other possible photosensitive materials are subject to prolonged exposure to sunlight through un-shuttered windows. Temperature levels in both the basement storage and exhibit rooms vary drastically, and goes unregulated for the most part throughout the year. Additionally, many artifacts are not cordoned off in the main exhibit areas to prevent accidental damage from patron handling.

In terms of disaster preparedness, our initial assessment revealed the GWBHS has many areas to improve. We noticed that GWBHS currently lacks fire alarm systems and extinguishers throughout the building, creating a risk environment for fire damage. Also, the GWBHS does not use pallets in the basement storage area consistently, and some artifacts are stored on the ground, putting them at risk for flood damage. Because the building is left unattended for long periods of time, many disaster situations may go unnoticed for a period of time in which it may be too late to rescue important artifacts.

The GWBHS currently has no formal policies for the preservation of its artifacts. Most documents or artifacts that are deemed important are usually kept under glass (with no regular maintenance procedure), or are sheathed in some sort of protective covering such as plastic. Other artifacts are left out in the open, or are stored in the basement with little to no protection. In fact, the organization regularly employs some of its artifacts—like antique farming tools—as hands-on props for local history events held at the museum, further reducing their lifespan. We also discovered the GWBHS has no formal selection procedure for identifying objects that should be better preserved, which has led to the neglect of higher-risk artifacts at the museum.

The GWBHS has a philosophy that encourages patron interactivity and education, sometimes at the expense of preservation of materials. Some items, including photographs and important documents, are kept away from the reach of the public in glass cases or other enclosures. Other objects are free to touch. As mentioned before, the museum routinely uses artifacts as hands-on props during community events. There are no rules for handling

at the museum, and no barriers in place to prevent patrons from handling potentially sensitive artifacts in the exhibit area.

Section 1.2: Collection Importance and Uses

The GWBHS' mission, as indicated on its web site, is "to celebrate and preserve the local heritage of the area, through its products, its technology and its people." Its collections exhibit much of the Irish immigrant farm culture that spread through the region following the purchase of the land from a local Native American tribe. The museum still holds the original land purchase agreement document between the Native Americans and the United States Government. The document on display in the exhibit room is hand-signed by former President John Quincy Adams and the leaders of the native tribe, and is a priceless piece of regional history. Apples and sheep farms thrived in the Greater West Bloomfield area in the 19th century, and settlers copied this apple growing tradition from the Native Americans who lived in the area. This long-time tradition of apple farming is preserved in some forms by the museum, which has many farming-related tools and even an apple cider press from the era. A mural depicting a sheep and a schoolgirl by local artist Roy Gamble adorns one wall of the museum, a product of the region's sheep farming roots.

Even the GWBHS building itself is an artifact of Orchard Lake Village, having served as Orchard Lake Hotel in the mid- to late-1900's, and then as the Orchard Lake City Hall before transforming into an area museum. The collections the building holds are varied, ranging in physical artifact types - farming tools, era clothing, film reels, documents, paintings, newspaper clippings, maps flags and more.

The museum hosts much of its collection for display in an exhibition room and annex, yet a considerable amount of the collections are not on public display (including books, documents, tools, film reels, pictures, tax records, and more), and are stored in a basement storage area or in the GWBHS office. The museum is concern with the fact that a large portion of local residents and researchers cannot see its holdings. Thus, the need to digitize the collection for both preservation and access is crucial to the museum's patronage and continued importance. The museum's express desire to preserve the local heritage of the Greater West Bloomfield heritage in its artifacts and promote access online led us to draft this project plan for digitization.

Section 1.3: Risks and Cost-Benefit Analysis of Digitization

The purpose of this digitization project is to migrate the physical collections of GWBHS to digital format for long-term preservation and access. Taking into account the risk-laden environment at GWBHS as exhibited in the needs assessment survey, there are many factors that can hinder the progress of our proposed digitization project. Placing priority on a digitization plan necessarily prohibits the implementation of other—arguably more pressing—plans like a disaster and danger preparedness plan, or an exhibition environment maintenance plan. For example, if a flood, fire or theft occurred during the course of the digitization project and compromised valuable artifacts, this could have drastic repercussions for the museum that had no disaster preparedness. The digitization would be interrupted, and both digital and physical artifacts could be compromised or completely unusable.[1]

Digitization does not guarantee the continued lifespan of the physical artifacts housed in the museum building, nor do the digital object copies convey the exact information of the physical artifact they document. In fact, digitization may even further diminish the quality of the collections, as the museum runs the risk of damaging already fragile documents during handling. These represent just a sample of the considerations that played a role in the course of developing this digitization plan.

On a different note, in choosing to write a project plan for digitization we had to weigh the importance of digital technologies and the need to provide access to patrons and disseminate the cultural history of the region to residents and researchers alike. Currently, the museum is seldom open to the public, and only a handful of patrons visit the museum each month. Therefore, the artifacts in the museum are susceptible not only to wear-and-tear from the storage environment, but just as important, their existence is largely unknown to the very community they represent. The digitization plan would allow GWBHS to start crafting a bigger Web presence for their collections using the digital files created from the project for access uses. While this report will suggest specific file formats for digital artifacts and access copies, it will not offer suggestions for the hosting of these digital files on the Web or elsewhere for access. It is expected that the GWBHS will choose its own platform for hosting digitized files over the Internet, and will be responsible for determining the rendering of images online, the usability of the site, and any additional considerations that user access requires.

Another benefit in writing a project plan for digitization is it offers a convenient and organized way to initiate the process of cataloging the collections at GWBHS. Currently, the

[1] Although we will not draft recommendations on those types of policies within this document, the GWBHS should keep other storage-environment concerns in mind when moving forward with plans for the museum.

GWBHS does not have an inventory of its holdings, and there is no consistent catalog method used to mark artifacts for ownership. Over the course of a complete digitization outcome - as every object will necessarily be handled and scrutinized - the GWBHS can implement a system of its preference to label and catalog each of its artifacts. This report will not highlight a specific method for the labeling or cataloging of *physical* objects, but will offer suggestions in terms of the cataloging of *digital* objects.

Section 2: Project Description

With this project plan, our aim is to clearly articulate the process that a quality, in-depth, and well thought out digitization effort should follow. We aim to make recommendations on the types of benchmarks that major policy-development decisions should implement, the amount of labor and time requirement that a project of this scale will require, and resources, planning, equipment, and external resources (including potential vendors) that will need to be solicited. We believe that a fully digitized collection can offer GWBHS patrons unprecedented access to the museum's collections while dually fulfilling its mission to increase access to and knowledge of local community history. Preservation is a necessary reason for this project to be undertaken. However, it cannot solve or eradicate problems presented by the conditions of the analog collection: "digitization is not preservation – it is simply a means of copying original materials. In creating a digital copy, the institution creates a new resource that will itself require preservation."[2] In this plan, we will outline some of the major components of a digitization project that require implementation within a larger digitization policy, and that will also necessitate funding, additional labor, and several years of attention. Where external information or resources are valuable, we provide footnotes and appendices for further information.

For the purposes of clarity, we will refer to four specific types of information individually (according to the National Information Standards Organization (NISO)[3]), and understanding the difference of each of these will be important as this plan continues.

- Collections (organized groups of objects)
- Objects (digital materials)
- Metadata (information about objects and collections)
- Initiatives (programs or projects to create and manage collections)

[2]Preservation and Selection for Digitization leaflet (from the Northeast Document Conservation Center) http://www.nedcc.org/resources/leaflets/6Reformatting/06PreservationAndSelection.php

[3]National Information Standards Organization, specified information types: http://framework.niso.org/node/7

In this section we will lay out a general description of how the GWBHS can carry out a digitization project that follows best-practice methodology. In the two subsections that follow, we will discuss the goals of the digitization effort and how the goals address major missions of the GWBHS museum, and then we will discuss procedures surrounding selecting the most appropriate information for digitization in a series of phases.

Section 2.1: Project Goals

In building from our Preservation Needs Assessment to focus intently on a single project plan for the GWBHS, we recognized their clearly articulated desire for a more comprehensive digital collection, and parallel concern for increasing the access and awareness of the museum's historical collections.

Distinct reasoning to support the choice of a digital collection for improving collections accessibility is evident through the GWBHS' entirely volunteer-based culture. With a small staff and extremely limited visiting hours, very little accessibility has been available to the user community to date. Further, the GWBHS staff seemed particularly interested in and motivated by the possibility of online accessibility to collections, and we believe this would promote the improvement of major museum weaknesses that indirectly relate to accessibility. Resolving some of these weaknesses are direct goals of the project, and they include:

- *Gaining control over the collection*
  As was mentioned in the previous section on *Context for the Project*, the GWBHS collection has not been cataloged. This means that there is no single directory where information about a particular object can be gleaned *or* recorded. This puts items at high-risk for value-loss: volunteer staff that knows the items best could move, or cease their involvement with the GWBHS before that information is preserved. Further, without full collection documentation, there can be no complete understanding of the relationship between particular items or sub-collections, preventing the potential for new context to be seen or utilized by users. We believe that the digitization effort will force cataloging to be implemented, which we hope will subsequently promote follow-through for the long-term care of digital objects.

- *Tracking the condition of physical museum artifacts*

Another problem emanating from the absence of a full collection catalog is that no benchmarks are currently in place to compare past condition of artifacts to their condition today. Thus, the deterioration or changing state of items cannot be determined. During the cataloging process, a precursor to digitization, a detailed description of the physical condition of each item should be included within a database. As a result, the cataloging will benefit increased access through digitization, *and* will provide for improved analog collections.

- *Expand user audience through improved access to collections*
  The GWBHS collections are underutilized in their current environment. With the museum facility open only four hours per month, even the most avid of users are limited in their access to the collections. Online storage of materials will promote use, circulation, and will offer users outside of the West Bloomfield area a chance to view and utilize the collections remotely. The online environment can promote the amount of feedback the museum receives from users by introducing feedback tools or surveys within the interface, so that the GWBHS can better understand and serve the needs and demands of their user community.

- *Increase access without added burden to volunteer staff*
  With the expansion of the GWBHS website and museum content, more access could be delivered without the requirement of additional open museum hours or further time commitments from the thinly-spread volunteer staff.

- *Utilize the help of area-universities for growth and expertise*
  Interns and summer student hires can afford the GWBHS appropriate help at little to no cost. Further, the external word of mouth and news that such opportunities can bring to the museum have the potential to increase awareness (and possibly funding) for the GWBHS and its future efforts.

- *Recording metrics on how long the digitization of particular items should take*
  The *Digitization Manager* or *Assistant Manager* (positions described in *Section 3: labor*) should make sure that the time required for each aspect of the digitization process (including training, cataloging, arrangement of artifacts, actual digitization, description input within software, etc.) is recorded, so that an average or expected amount of time for a given phase can be worked into the expected timeline of the project, and increasingly curated as the process continues. This type of information can help the organization understand the timeline it requires, but also help other similar societies identify the requirements of digitization, keep interns and volunteers

on track by expressing the average amount of time that a given task should require, and better visualize the amount of work that can be accomplished (within one hour, one week, and one month, respectively). In the case that a time crunch requires more end product to be processed faster, the GWBHS will have a clear concept of how much can be accomplished with the hiring of even one additional individual.

- *Carefully digitizing, describing, and arranging digital content for the long-term*
  Our initial Needs Assessment Survey outlined a lot of the problems with the GWBHS' physical storage space, building size and capabilities, and the obstacles it holds for the long-term preservation of physical artifacts. Creating digital representations of these items will not preserve them in the long-term, but it will allow one form of the data, at a minimum, to survive. Additionally, it could raise awareness that later could bring in funding with the potential to start an added initiative for preserving physical collections. The long-term accessibility and preservation of the digital objects, however, will provide unhindered remote access to a wider base of users, fulfilling one of the main missions of the GWBHS.

Goals are essential to the continued motivation for this project, but are also necessary to ensure that the effort *solves* a clear problem. In order to stay organized and be sure that goals are articulated throughout the entire digitization process, the GWBHS will need to develop structured policies for several aspects of the project. A policy for digitization should include benchmarks for measuring and preparing for quality assurance and technical standards, evolving audience needs, increased web traffic, and an appropriate timeline for accomplishing the project. Additionally, making these goals and data available to users will benefit the GWBHS. Highly user-oriented organizations work to maintain transparency[4] wherever possible, and the GWBHS should strive for similar visibility. If goals, needs, and obstacles are articulated clearly throughout the process, users will be aware of progress, and potential volunteers and donors will understand the specific needs in the event that they have the means to help.

Section 2.2: Selection

[4] The *Trusted Repositories Audit Certification* (TRAC) process explains the definition of transparency in the digital environment as: "a repository that exposes its design, specifications, practices, policies, and procedures for risk analysis" The document further asserts that..."digital repositories must be transparent in all practices as they relate to preservation capabilities or assertions made about trusted, long-term management of digital materials." http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

The GWBHS has hundreds, perhaps thousands of items and artifacts between its exhibition space and storage collections. Knowing that these items are not cataloged or processed with unique identifiers is problematic when trying to assess both 1.) the volume of data necessitating digitization and 2.) the form (manuscript, book, farming equipment, clothing, etc.) of each object, because different 3D materials require varying digitization equipment. Thus, not every museum artifact can be digitized in one fell swoop, or over the same period of time - the work must be done in phases. We recommend that the GWBHS prioritize different types of items for digitization, grouping like-items together. Because we do not know the demands of users, or the community's most pressing needs, we find it essential that the volunteer staff that interacts with the content and users most frequently form a committee to discuss selection priorities.

The first phase requires choosing a specific type of museum resource for cataloging, digitization, and addition to the online repository. To get started, our suggestion would be to begin Phase I with 2D materials (i.e. photographs, papers, or manuscripts) for two reasons: 1.) 2D digitization equipment is more commonplace, cheaper to use and rent; 2.) metadata description can be based on items already in the online repository, requiring less research and process upfront to begin the initiative. After 2D documentation, the GWBHS will need to determine the next series of items to be digitized. We hope that the GWBHS uses the potential of the online atmosphere to solicit feedback and requests from patrons so that some metrics can be utilized to measure the demands of users and the uses that the first phase materials have allowed. Such statistics (in addition to budget allotments) will help to facilitate a conversation and decision for successive phases of artifact digitization.

Selection is of utmost importance, but it is also imperative to state that a policy should always back every digitization phase: a clearly selected type of content should be chosen, and the reasoning for its selection should fully support the goals articulated by the project. Should need, goals, funding, or external circumstances shift, the policy should be reevaluated before moving forward. Each phase should be carried out completely (full digitization) before beginning with new content for a successive digitization phase. This way, at least a portion of the collection can be made highly accessible via the intranet, rather than multiple phases being partially initiated; this is the easiest way to streamline the process and deliver content to users at a faster rate.

The aforementioned components should play key roles in the selection process, but other discussions warrant attention as well, depending on whether they are applicable to the digitization of a particular phase:

- *Intellectual property rights and copyright concerns*
  In our interviews with the GWBHS, copyright seemed to be of very little, if any concern to the repository. Most items within the museum have been donated. Beyond that, most pre-date the current copyright benchmark. Those certainties are wonderful for the GWBHS, but we would like to stress the importance of making sure that documentation of donation, clear ownership, and absolute copyright possession lie with the GWBHS before content is posted to the intranet. Copyright infringement is a serious issue, and could create a lawsuit that puts the entire society at risk. If any doubt surrounds an object's copyright standing, it is better to err on the side of conservatism, and seek an external opinion, ask for licensed approval from the copyright holder to make the data available online, or elect not to digitally publish the item for the time being.

  Information should also be provided to users about the copyright status (distinguishing between items whose copyright is owned by GWBHS, which items have fallen out of copyright and into the public domain, and which are being licensed). It is important to note the extent to which users can access, download, retrieve, or pay for digital copies of information. Preferred citations should be made available for patron use in the online environment.

- *Item condition and diversity*
  Many of the artifacts, as listed above, greatly vary. Items can come in different sizes, ages, materials, qualities, and value to the community. Such variety is an asset to the collection, but should also be kept in mind throughout the process. Different equipment, management, and long-term maintenance apply to each unique form. It is important to be responsible when making selections, and to choose items whose full digitization can be funded. Further, copyright and life span can vary as well. Part of the attention given to selecting the next phase of items for digitization should surround which items necessitate digitization. Need can be driven by users, but it can also be driven by deteriorating item condition. In the case that the physical environment cannot sustain preservation, a digital representation of an item is better than total loss. High consideration should be placed on items with reasonable risk for loss, so as to save what is left of their existence before total historical dissolution.

- *Conservation*
  Because the digitization effort will take place simultaneously with cataloging, it is possible that the condition of some items will necessitate conservation work, or that items are determined to be too fragile for digitization. In these cases, measures

should be taken to care for the needs of those items in addition to continuing the digitization process, or to donate those items to repositories and institutions that can conserve or care for them in the long-term. In the event that massively deteriorated items are discovered during the reorganizing and cataloging process, the needs of the collection could necessitate halting the digitization process. It is important to develop these types of scenarios within the digitization policy document, and is valuable to consider their possibility at early stages of policy development.

Section 3: Preparation for Digitization

Before digitization can begin, a working policy needs to be developed by a specified team of GWBHS volunteers, and drafted by a specified deadline. After a draft has been created, the policy should be presented to the GWBHS board, and edited/amended until it has been accepted and approved by the board at large.

In outlining a comprehensive policy to include specific measurement metrics and requirements of the Digitization effort, several work-related components need to be included and drawn out in as much detail as possible. For example policies off of which the GWBHS should draw to develop its own working document, we recommend the *National Library of Australia's digitization policy*,[5] the *NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Documents*,[6] *Technical Guidelines for Digitizing Cultural Heritage Materials*,[7] and *East Carolina University's Joyner Library Policy Manual*.[8] To help develop some of the areas where cost, time, attention, and policy should be allocated; we have included a series of sub-sections below. Each will include a description and any findings or recommendations on the types of considerations that should be made by the GWBHS in policy creation. The subsections will be ordered as follows: labor, equipment, training, location and facility, budget, and timeline.

Section 3.1: Labor

One of the largest allocations for this project will be labor. The GWBHS needs to determine who will be responsible for leading this project, who will help as support staff, whether or not outsourcing work to vendors is a good option, what the costs for each of those avenues

[5] National Library of Australia: Digitization Policy document: http://www.nla.gov.au/policy/digitisation.html
[6] NINCH Guide to Good Practice: http://www.nyu.edu/its/humanities/ninchguide/
[7] Technical Guidelines for Digitizing Cultural Heritage Materials: http://www.digitizationguidelines.gov/guidelines/digitize-technical.html
[8] East Carolina University's Joyner Library Policy Manual: http://digital.lib.ecu.edu/info/DigitalCDP.pdf

will be to the organization, and what prioritizing digitization will mean for the development of other aspects of the GWBHS museum.

Included in labor duties will be the cataloging and reorganization of artifacts, assessing their condition and value to the user community, and prioritizing their digitization within a particular phase. After this will come cleaning the artifacts (as necessary) and digitizing of each. Evaluating external vendors to digitize three-dimensional artifacts, and developing training session on any purchased equipment to use in-house will also be necessary. As a further concern, the GWBHS will need to consider the costs of persons in comparison to the budget allocated to this project. If vendors are hired, for how long will the museum (and how long *can* it) expect to afford the added help? Maximums should be set for each of these factors within the budget policy, and those should be adhered to strictly. Where vendors and experienced corporations cannot be afforded, the museum should think about hiring graduate-student interns from programs that offer training in-line with the needs of the GWBHS. For a closer description of recommended personnel for the digitization project, see the following outline:

- *Digitization Manager* - It is extremely important that someone with a developed understanding of the GWBHS collections, and an investment in their preservation be designated as a manager for this project. Based on our understanding of the personnel at GWBHS, we believe it is also important that someone with a technological background would add helpful experience to the process. For these reasons, we see Buzz Brown or Bill McIsaac as natural candidates. We expect this position to require about 20% full-time-equivalency (FTE) hours (or around 5-8 hours per week). This person will be responsible for overseeing all other personnel and workers on the digitization effort. Some portions of the process may require more intensive work hours, while others time frames may be lighter. As this person is directly responsible for all digitization employees, they should also have a direct hand in hiring the team, interns, and volunteers selected for the project. As a last recommended requirement, the *Digitization Manager* should schedule regular meetings (likely monthly) with all workers associated with the project, so that open communication, goals, and progress with the project are facilitated and known by all associated.

- *Assistant Digitization Manager* - Because our timeline (see subsection on *Timeline* for more detail) expects this process to require several years before total completion, we foresee some likelihood that the designated *Digitization Manager* will not be able to oversee the entire process. If something happens to the project manager, (i.e. injury,

family-related emergencies, or moving from the area), a secondary individual with appropriate institutional knowledge of the project should be in place for continuing progression with digitization. In the situation that a back-up manager takes over full management responsibilities, a new *Assistant Manager* should be appointed. This person should work closely with the *Digitization Manager* and carry extensive responsibilities over the project. They should be familiar with all supervisory staff and able to provide necessary training sessions as needed. This position will also require approximately 20% FTE (5-8 hours per week).

- *Interns* - Experienced students in appropriate fields of study are great resources for inexpensive, motivated help, and will be vital to the digitization project. Given that much of the workforce at the GWBHS is volunteer-based, having interns can drive progress and maintain a focus on the long-term goals of the society. We have suggested that the GWBHS offer full-time, summer semester internships that reflect the study schedule of local university programs. Within the digitization policy, the GWBHS should spell out specifically desired skills and talents in a job description for interns. We have provided suggested position descriptions for internships within **Appendix A**. Interns hired for summer months should be expected to work 50-100% FTE (20-40 hours per week). Possible fall or winter semester interns could work anywhere from 25-50% FTE (10-20 hours per week). Some of this time should be under the supervision of the *Digitization Manager* or *Assistant Manager*, but after interns receive training and an introduction into the GWBHS, they should be able to work at an adequate pace with museum materials on their own.

- *Volunteers* - Low cost help can be extremely helpful, as was discussed in the bullet-point on interns (above), but because volunteers are not guaranteed to have the same valuable skills as students learning to become professionals in the preservation field, some level of GWBHS-based motivation and task delegation must be offered to assure that volunteers are taking on appropriate and helpful tasks for the project, and are consistently working on them. Expectations for volunteers, and any requirements for their training and skill-level should be expressed within the digitization policy. Volunteers should not be required to work a particular amount of hours per week, but a minimum of 10 hours per month should be required, or else a volunteer could lose scope and familiarity with the digitization project, and no longer be able to carry out valuable tasks for the GWBHS. Volunteers should work during hours when either managers or interns are also scheduled, and fill out a time sheet (approved by a manager) upon each visit, so that the GWBHS can keep a record of the amount of hours required by the digitization process. Volunteers should register with the

GWBHS in an official document (that should be provided within the digitization policy) to comply with insurance and non-profit requirements.

- *Vendors* - We believe that external vendors will become a realistic option to be hired by the GWBHS at some level of the digitization process - likely when moving to the digitization of 3D materials. Vendors will require a list of expectations and requirements for technical specifications (see sections 4.1-4.2 for more information), a timeline for completion, and the payment they will receive for their services. The GWBHS should look at potential vendors well ahead of the time when vendors are necessary to the process.

  Depending on the type of artifacts being digitized in the next phase, specific equipment, software, or procedures might be required (i.e. large or oversize artifacts have different needs than silver-gelatin prints), and some vendors specialize in digitizing particular content (which could be beneficial for dealing with various over-size objects within the GWBHS collection that would be otherwise very difficult to digitize in-house). The GWBHS should be aware of which types of vendors will be suitable based on the needs of upcoming digitization phases. Because vendors are proprietary, and interested in the business end of partnerships with institutions in need of digitization, it will be more expensive to have them digitize content than to the digitization internally. A vendor will, however, be able to deliver digitized content much more quickly than the volunteer-staff of the GWBHS would be capable of.

  If vendors are elected for digitization, the GWBHS should take care to consider its needs and relay those to the vendor in a timely manner. Otherwise, the vendor might not understand those needs, or fail to keep them in mind during their work. Any requirements, budget allowances, or expectations for vendors (as well as the timeline they can be afforded) should be explicitly detailed within the digitization policy.

  It can be difficult to find digitization vendors due to the interdisciplinary nature of their work: they can be involved primarily in publishing, copying, records management, or even be a cultural institution themselves. For example, Michigan State University Library's Digital and Multimedia Center serves the student body and digitizes multiple types of objects for inter-university divisions, but also operates an external business.

  Getting involved in cultural institutions and organizations for museums or archives will open the GWBHS up to many vendors that often try to advertise at conferences.

Jill Hurst-Wahl, a field-professional, wrote a relevant blog post about various digitization vendors who were present at the Special Library Association's annual meeting.[9]

Section 3.2: Equipment

Software programs (for scanning/capture, reformatting, and metadata tagging), and hardware (a scanner; computer capable of handling the software, digital files, back-up hard drives) need to be purchased and closely maintained over time. Because of the quick procession of technological developments, any purchases are likely to become obsolete very quickly, so keeping an eye out for evolving systems and software will be important. Scanners, gloves for handling items, archival boxes for storing and housing items after their digitization, and other items unaccounted for in this project plan should be made available as needed.

Section 3.3: Training

As further criterion for consideration, the GWBHS should assess how and at what cost training should be provided for employees on using the software and hardware systems, cataloging and describing the digital objects, choosing and learning metadata schemas for the description, and assessing the condition of each item so that fragile items can be flagged for conservation or higher-level attention.

Labor concerns are a major aspect of this project, and so, requirements and skills should be as well. Student hires or volunteers can (and should) be solicited according to their fit with the organization, but volunteers should be trained, and a project manager should be selected to oversee work flow, make decisions when unexpected scenarios arise, and take care of future training: either administering training on his/her own, or selecting (or hiring) appropriate training facilitators.

Section 3.4: Location and Facility

It is important to also think about a location for digitization and project work. The GWBHS does not have extra room in-house for a digitization project of this scale. Renting another facility may be an option, but vendors hired to digitize could also be attractive if outside working spaces are difficult to locate for the time-span required.

Section 3.5: Budget

[9] URL link to Jill Hurst-Wahl's blog post: http://hurstassociates.blogspot.com/2007/06/digitization-vendors-at-sla-conference_10.html

Given that much of the services are non-paid or volunteer-based, most of the monetary expenditures will be allocated toward obtaining hardware and software (discussed in *Section 4*). A lot of the budget will require measurement in terms of time (and FTE) rather than strict finances. An itemized budget (Microsoft Excel spreadsheet) is provided with expected costs within **Appendix B** of this document.

Section 3.6: Timeline

Although the preparation, policy drafting, selection process, and goal setting aspects of the project are a lot of work, they are requisite for a successful initiative. In moving forward with digitization, it is important to be aware of what is realistic in setting up a timeline. The small staff and very few open hours of the GWBHS museum means that first digitization phase of the project may take longer than desired. Still, an appropriate goal should be set for overall phase I completion, and smaller benchmarks for more finite components should be fit within that timeline too. We have loose constructions for setting benchmarks for the short, medium, and long-term, and those should be amended as fits the GWBHS culture and abilities:

- *Short-term* (1-2 years out): Includes goal setting, policy drafting, equipment allocation, vendor analysis, selection, and Phase I of digitization.

- *Medium-term* (2-3 years out): Once Phase I of digitization has been completed, the "medium-term" will begin, and should include a reasonable amount of time to reflect on the process, consider the value-added by digitization, analyze any data or feedback garnered from users, and re-evaluate the appropriateness and fit of previously constructed policies. Once appropriate changes have been made, and the next round of items are chosen for selection, Phase II of digitization should be carried out to completion.

- *Long-term* (beyond 3-4 years): Additional evaluation periods and successive digitization phases, to be completed as budget, resources, labor, and time can be spared.

A visual timeline is provided in **Appendix C** with a recommended progression of events for the digitization effort. We carefully considered available staff at the GWBHS during its development, and expect the policy drafting and approval alone to take around six months. We also tried to align the effort with summer months for the purposes of hired interns. In

crafting a digitization policy, we recommend that the GWBHS consider our suggestions, but ultimately amend any aspects of this project plan to fit realistic expectations for its own needs.

Section 3.7: Reflecting on the Value Added by Digitization

When the GWBHS has fully drafted a full-scale digitization policy, had it approved by the board, and begun and completed the digitization of Phase I, the society should reflect back upon the processes undergone. The strengths, weaknesses, problems, unforeseen circumstances, and any feedback gleaned during Phase I should be discussed, and added as metrics for going forward with additional digitization phases. Further, before moving along with additional phases, all policy documentation should be reevaluated for its relevancy. Any changes or divergences from the original policy should be added and recorded, and the policy should be amended to reflect the needs required by digitization that have been learned since the original policy development. The GWBHS board should approve this—and any successive—revised policies.

We also recommend that a survey be provided for users to fill out if they utilize the digital holdings. Use analytics should be recorded once digital data is made available to users, and the amount of users should be tracked as they are visiting the site. What will be of particular interest (and value) are the types of items that users are most frequently visiting. This type of data may provide the key to which materials are in the highest demand by the user community, and similar non-digitized materials should be flagged with a higher value for digitization for future phases.

Section 4: Technical Standards and Quality Assurance

Section 4.1: Digital Object Control and Standardization

In constructing the parameters for digitization, it is important to consider that the digital object copy is a separate artifact from the physical object it describes. Therefore, no digital copy will entirely capture the information contained in the original. It remains important, however, to preserve as much information about the original object in its digital record, through the appending of precise metadata for later reference. Additionally, adhering to best practices in file reformatting, color management and reproduction, workflow, and metadata tagging is crucial to describing the integrity and provenance of the original artifact object. For the purposes of our project, we will recommend preserving two copies of the digital object - one as a master copy and another as an access copy for patron use.

In our project plan, digitization will occur in stages, where the first series of stages will include the scanning of documents, images (prints and negatives), or other two-dimensional objects, and the second series of phases will include the digitization of larger two-dimensional objects and even three-dimensional objects through high-resolution photography. For the second series, GWBHS will likely need to seek a vendor to digitize larger objects. Any vendor contracted for digitization should maintain the same standards outlined throughout this report (specifically *Section 4*). In recommending software and hardware for purposes of digitization in the first stage, GWBHS should require a strict adherence to standards set by the U.S. National Archives and Records Administration in terms of spatial resolution (dpi, pixel width and height), tonal resolution/color and bit depth, and file formats.[10]

Section 4.1.1: Hardware

For purposes of processing two-dimensional objects, the GWBHS will need to obtain a computer and monitor capable of efficiently executing Adobe Photoshop software, Epson scanning software, and monitor calibration software (for accurate color reproductions), as these are the three main software we will recommend for basic handling of image/document capture, quality control, reformatting, and metadata tagging. The most demanding of these software programs is Adobe Photoshop CS5 (see **Appendix D** for system requirements). The display monitor should have manual color adjustment controls capable of accurately reproducing color temperature, saturation, and hue of the digital capture. The monitor should meet Adobe Photoshop's minimum recommendations for display resolution at 1280x800 pixels. Based on current pricing trends, our estimate for a minimum specification desktop computer and monitor package is estimated to cost $700.

For scanning, we recommend that the GWBHS purchase an Epson V750-M scanner[11] (approximately $850), which - according to its specifications - is capable of reproducing the necessary tonal resolution, spatial resolution, and color space and depth required by the U.S. National Archives and Records Administration. We have chosen this scanner based on its ubiquitous use for professional applications, high product reviews, and high-quality construction that will make creating professional scans with minimal need for software correction. Keeping in mind that the scans will create very large TIFF output files, it is pertinent that the GWBHS make sure their storage and back-up hard drives (at least two)

[10]Peterson, Kit A. "Digital Master Images: Sample Technical Specifications for Photograph Collections" http://www.loc.gov/rr/print/tp/DgtlMastersSamplSpecsSelctdRcmndFinal7_2004.pdf
[11]Epson Perfection V700 Photo/V 750-M Pro: http://www.epson.com/cmc_upload/0/000/142/238/V700_V750_InfoSheetR1.pdf

are sufficiently large enough (at least 1 terabyte recommended) to hold as many master images and access copies as necessary.

Section 4.1.2: Software

The three types of software we will recommend for digitization include Adobe Photoshop (latest version), Epson scanning software, and the Datacolor Spyder 3 Pro software (or similar monitor calibration software). A license of Adobe Photoshop can be purchased from authorized non-profit Adobe retailers for around $300. Adobe Photoshop uses the Extensible Metadata Platform (XMP), which incorporates fields from the Dublin Core Metadata schema. Taking into the consideration the technical background of the staff at GWBHS, we recommend using Photoshop's file explorer to tag images using the XMP metadata schema dialog boxes. The XMP schema can later be exported directly to Dublin Core, either via Photoshop or another XMP compatible software. Adobe Photoshop also delivers the tools necessary to reformat the master TIFF scans to JPEG for creating access copies, as well as edit the color parameters of the image for quality control.

The Epson scanning software comes pre-packaged with the Epson V750-M scanner, and can be used to capture objects in color or grayscale, at a variety of resolutions. While scanning, the GWBHS should consult the Digital Master Images chart compiled by Kit A. Peterson[12] to ensure that all recommended digital master image parameters are dialed into the capture interface. The final recommendation is monitor calibration software. One of the industry's leading software for monitor calibration is the Datacolor Spyder 3 ($175 retail), which is capable of accurately reproducing the color of a digital image object by measuring the output of the monitor. This software should be run on the monitor at every scanning session to provide the most accurate rendition of colors of the scanned master copy for the human eye. In combination with the scanning software and the professional scanner, the calibration software will help reduce color rendition disparities between the physical artifact and its corresponding master digital image.

Section 4.1.3: Image Capture Parameters

The digital master copy for preservation should be submitted and preserved in the Tagged Image File Format (TIFF). The Online Archive of California and Harvard University Libraries, in addition to the U.S. National Archives and Records Administration currently use the TIFF 6.0 (latest) format for long-term preservation of digital images in their repositories, and so we have also chosen to standardize on this uncompressed format. For photographs and other non-transparencies, we recommend master TIFF files should

[12] "Digital Master Images - Sample Technical Specifications for Photograph Collections." Library of Congress: http://www.loc.gov/rr/print/tp/DgtlMastersSamplSpecsSelctdRcmndFinal7_2004.pdf

conform to minimum 300 dots-per-inch (dpi) resolution (maximum 400 dpi), maximizing legibility of the image while also remaining storage-efficient.[13] [14] Pixel width and height for capture is also calculated using the U.S. National Archives and Records Administration standards.[15] While Harvard University Libraries note that the compressed JPEG 2000 format is sometimes used as an acceptable format for preservation due to its more economical file sizes (and therefore, ability to store), our model will only recommend TIFF images for the master copy.

In addition to the master copy, an access copy should be created directly from the master TIFF file. For the sake of ease of use for Web publication, we recommend the GWBHS create a compressed access copy in the JPEG container. The images should be compressed enough to remain quickly render-able in a Web environment, but should retain the 300 dpi resolution parameters of the master copy. Ideally, the final size of the access image will remain at about a megabyte in file size, but not much larger. This file will be saved alongside the master copy, and should reflect its version (i.e. JPEG) within the naming convention used for storage search and retrieval. The U.S. Department of Cultural Resources provides a valuable best practices document[16] that can offer suggestions for GWBHS in crafting their own naming convention procedures for file storage.

Color images should be scanned in using 24-bit Red-Green-Blue (RGB) color space, creating 8-bit depth for each color channel. Black and white images or documents should be scanned at 8-bit grayscale. It is recommended that images are scanned *without* auto-correction filters so that color, brightness, contrast, and other values can be accurately corrected by an experienced intern using advanced tools in Adobe Photoshop. However, experimentation with scanner settings and preset values may be necessary at first operation to ensure efficient workflow by reducing the amount (or scope) of post-capture editing. One primary advantage to using the monitor calibration software is to enable faster editing by allowing the editor to see color discrepancies caused at capture.

Section 4.1.4: Quality Control

In terms of quality control, effort should be taken to accurately monitor color post-capture. The reason we have suggested the Epson scanner is for its highly reputed attention to color accuracy. The scanner uses a very specific light source that is trained (via software) to

[13] Ibid.
[14] "Manuscript Digitization Demonstration Project." *5. Preservation-quality Images: Types and Resolution.* Library of Congress: http://memory.loc.gov/ammem/pictel/mddp305.htm
[15] "Digital Master Images - Sample Technical Specifications for Photograph Collections." Library of Congress: http://www.loc.gov/rr/print/tp/DgtlMastersSamplSpecsSelctdRcmndFinal7_2004.pdf
[16] "Best Practices for File Naming." N. Carolina Dept. of Cultural Resources: http://www.records.ncdcr.gov/erecords/filenaming_20080508_final.pdf

accurately render the color of the objects. Additionally, with the aid of monitor calibration software and Adobe Photoshop post-processing tools, the GWBHS can correct for color imbalances post-capture. This type of post-capture quality control requires expert attention, and should require an intern or vendor with advanced knowledge of color space editing via Adobe Photoshop software.

Another important consideration is the value of investing in a file validation software solution that periodically checks the integrity of files. Such a tool should run after every backup to ensure a successful backup of files, and that no data was corrupted or lost during the transfer. One of the reasons we do not recommend a cloud option for file backup is because file integrity cannot be ascertained by GWBHS on a consistent basis within the cloud.

Section 4.1.5: Metadata and Digital Object Description

The registration of scanned object metadata should be completed using the XMP schema developed by Adobe. Using Adobe Photoshop, GWBHS can input specific descriptive metadata for each of the digital objects using the File Browser and the XMP dialog boxes.[17] The XMP schema included all the required fields contained in the Dublin Core metadata schema, and can be exported to Dublin Core (and theoretically, many other XML schemas) using Photoshop or other XMP-compatible software. In addition to the XMP schema, EXIF data - image data created at capture including date, time, exposure and other data - can also be accessed using Adobe Photoshop.

The types of descriptive metadata we recommend GWBHS input post-capture follow directly from the U.S. Library of Congress document entitled *Common and Useful Information Elements for Cataloging Pictorial Materials*.[18] This document is especially useful as a model for the GWBHS because it includes item-level and collection-level examples of actual authored metadata for a collection. This metadata guide should be used for item-level descriptions within the XMP fields of the Adobe Photoshop software. In addition to the types of metadata outlined in the Library of Congress' document, we also recommend that the GWBHS pay close attention to the condition of the physical artifacts, and mark a detailed description of the condition in the "General Notes" (or equivalent) metadata field.

As we mentioned briefly before, digitizing the artifacts at the museum also gives GWBHS the opportunity to properly catalog its holdings. During metadata tagging, each item should

[17] "About Metadata." Adobe: http://www.adobe.com/digitalimag/pdfs/about_metadata.pdf

[18] "Common and Useful Information Elements for Cataloging Pictorial Materials." Library of Congress: http://www.loc.gov/rr/print/tp/Common%20Information%20Elements.pdf

be assigned a unique catalog number, which can then be saved in the "Acquisition and Appraisal" (or nearest equivalent) metadata field. This catalog information should also be kept in a separate spreadsheet to accompany the physical collections separate of the digital collections.

Section 4.2: Digital Object Management

The previous section (4.1: *Digital Object Control & Standardization*) detailed the importance of robust metadata description, and image standards, and display consistency for end-users among other control mechanisms. These are essential for a well-crafted digitization policy and successful endeavor, however, their long-term preservation can only be guaranteed through efforts taken to promote and care for sustainability beyond the moment of digital capture and manipulation for first use.

The National Digital Information Infrastructure and Preservation Program (NDIIPP) at the Library of Congress developed a seven-point analysis schema[19] to determine what factors need to be considered in order to manage digital objects' waning longevity. These seven points will be listed below, and will be followed by a paragraph-form assessment of long term needs necessitated by the digital image metrics recommended in *Section 4.1* of this plan.

1. *Disclosure* - the extent to which complete details and resources for technical integrity validation both 1.) exist and 2.) can be accessed by the individuals responsible for the creation and management of digital content.
2. *Adoption* - the level to which a given format is already being utilized by any of: creators, distributors, or other consumers.
3. *Transparency* - the degree to which the digital image is viewable by the user without manipulation or advanced tools, and is readable with minimal technological aid.
4. *Self-Documentation* - Self-documenting digital objects will be easier to sustain over time and less vulnerable to disaster than data objects stored separately from their metadata, which is required to render the data or understand an object's context.
5. *External Dependencies* - rely on specific hardware, operating systems, or software for using or displaying digital objects, in addition to expected complexities associated with dealing with varying dependencies in technical environments.

[19] "Sustainability of Digital Formats: Planning for Library of Congress Collections." National Digital Information Infrastructure and Preservation Program (NDIIPP). http://www.digitalpreservation.gov/formats/sustain/sustain.shtml

6. *Impact of Patents* - their ability to sustain (or destroy) the value of digital objects over time is unknown and difficult to assess as of yet.
7. *Technical Protection Mechanisms* - should be put in place to guarantee digital content over time and provide service to users and relevant communities. It is expected that beyond digitization, the staff, finances, and attention also get put toward migrating (or replicating) content on new media when older versions are determined to be obsolete.

With sustainability as a requirement for preservation, it is just as important for the GWBHS to allocate attention, personnel, funding, and potential help for digital efforts in the same way that the policy planning procedures will require help.

Section 4.2.1: Data Management Personnel

Specifically, a part-time intern (whose experience level meets the needs of the GWBHS file creation and management policies) should stand to carry out the analysis and upkeep of evolving image standards so that digitized collections at the GWBHS stay in touch and up to date with nationally recognized standards. The *Digitization Manager* should oversee all major decisions that fall within the expected scope, and should be prepared to handle decisions that fall outside of it (any unforeseen circumstances). The *Assistant Manager*, however, should also be held responsible for making these types of decisions in the event that the primary manager is unavailable.

Section 4.2.2: Backing-Up Data

Once a policy and personnel have been allocated to the effort and digitization has begun, the implementation of long-term data management metrics must be considered even at the first steps of the process. All added digital content should be backed-up with an appropriately sized external hard-drive (see Section 4.1 for suggested hardware and software specifications) large-enough to keep digital object metadata and images (master and access copies) for the entire collection (including all expected phases of digitization).

At least two full data back-ups should exist. One should be kept on-site, updated daily, and used as necessary. A second back up should be kept off-site, so that in the event that an emergency or disaster occurs within the museum, a preserved copy of collections content will still exist. Similar to the on-site back up, the secondary should also save collection data at any time when content is added. However, outside of regular check-ups (to see that the hard-drive is fully operational), the secondary back up should act as the preservation copy of the museum's digital collections: not to be touched, removed, or accessed except in emergency situations.

Section 4.2.3: Digital Storage

To keep the information retrieval process as consistent and fluid as the rest of the effort, naming conventions consistent with the images and their description should be implemented within the computer mainframe storage and back-up storage environments as well.[20] This is important for several reasons. First, it will reduce any potential confusion if files are warped or lost in one system, but saved in another: a user or staff member can much more easily identify information whose tags and names they recognize from previous experience. Further, a basic outline and intellectual structure surrounding naming conventions in operation within the system should be articulated within a word processing document (i.e. Microsoft Word) and implemented as a part of the digitization policy: to preserve the digitization process itself, but also for the purposes of passing the framework forward to future workers, interns, and GWBHS volunteers.

Section 4.2.4: Quality Assurance

Although a software solution to the cataloging and online-access problem may be the best option today, stronger possibilities could become available in the future and warrant changes. As these technologies develop, not only will the existing systems become obsolete, but descriptive standards will also require transformation. It is essential that the GWBHS align its goals to keeping up with system and network developments so that safety, security, and digital object preservation can all be maintained for users in years to come (more information on necessary *evaluation* metrics will be discussed in the section that follows).

Purchasing software and hardware on a yearly basis is unnecessary (and likely out of budget), but the GWBHS should reevaluate the needs and current standards for digital objects every one to two years. It is our belief that eventually, software systems will be developed to structure XML metadata directly within a system, but that XML descriptive standards will not disappear or fade out any time soon. Thus, the *Digitization Manager* should consistently advertise the willingness of the GWBHS to work with student-interns who can share information about cutting-edge development in the field, and return information about the needs of the museum to their programs, colleagues, and other potential aid.

Technological upgrades should be assessed and carried out by technologically adept personnel. Otherwise, potential shutdowns and user-orientated problems could occur. By preparing for potential issues before they occur, and including plans to deal with unforeseen

[20] As briefly noted in earlier sections, we recommend the use of Dublin Core description when cataloging digital collections. Typically a "creator" or collection "title" is used as the main identifiers for collections. The GWBHS should use these as the storage naming conventions in a digital hard-drive, categorized more broadly within its appropriate digitization phase.

circumstances within policy guidelines in the digitization policy drafted from this document, the GWBHS will defray against the potential damage that disaster situations could otherwise cause.

Further, it will be important to ensure the capabilities of software and up-to-date equipment over time. Otherwise, software could be discontinued, funding could run out, disaster (natural, malicious, or otherwise) could destroy the value and ability of the collection to be digitized, or other similar problems could endanger the process. If value-loss occurs within the life span of a digital object, the item will have to be re-digitized or moved to a queue for later value-assessment (as time and funding permits beyond the scope of the project).

Section 4.2.5: Metadata Migration

We have briefly introduced and discussed the application of metadata to digital objects throughout this report. Metadata is essentially digital description. In the digital environment, this information has to be coded so that users can see it when accessing the GWBHS digital collections through remote online access.

There are software programs today that do this on a basic level, and we understand that the GWBHS is somewhat familiar with the practices of Past Perfect. As we stated in the introductory section on *Context for the Project*, our aim in this project plan is not to recommend platforms or software that will render the digital images viewable through the GWBHS website. Rather, we are interested in the process of digitization alone, which includes ensuring that image capture, associated metadata, and policies regarding labor, organizing physical collections, and attention paid to legal metrics (such as copyright and intellectual property) are attended to.

However, as a general note: regardless of what system is chosen for patron access to the digital objects (which, through our conversation Bill and Buzz of the GWBHS, was determined to likely be PastPerfect), it is highly likely to require replacement (due to obsolescence) within a decade. Digital repositories that care for permanent or semi-permanent digital object data have to care for those materials with this in mind. Thus, the GWBHS needs to constantly be watching for developing software and newer metadata-capturing practices. An easy way to do this is to watch and communicate directly with similar institutions in the Southeast Michigan area, build a network of professionals by allowing some GWBHS to go to conferences on cultural heritage preservation and related topics, and to watch and follow the initiatives that others take, including how frequently they do so.

For naming conventions and description standards, we are recommending that the GWBHS describe all digital objects using the minimum fields set out by the Dublin Core metadata schema (explanation provided in **Appendix E**), whether applied through a compliant software or through raw XML coding. Many museum software programs currently employ entry fields that translate directly to the required Dublin Core fields (including PastPerfect), so high-level expertise is not necessary for utilizing the standard. Dublin Core is a well-known framework with 15 recommended fields, however, other (better) description standards also exist, and in the case that the GWBHS collection becomes extremely high-volume with thousands of daily users whose needs push for further description, other standards may need to be considered.[21]

Section 4.2.6: Image Management

Over time, the quality and format of images themselves can also change. Today, the international standards for images are JPEG and TIFF. Other generally accepted formats include JPEG 2000 (JP2), and PNG. Each of these formats cannot be accepted across all platforms and systems. Therefore, all digital objects in the GWBHS repository should be displayed using a single consistent format (specific recommendations are listed in the previous section: 4.1). When that format begins to shift to an improved standard, the GWBHS digitization committee will have to assess how (and whether or not) to implement those changes: metadata and objects can be transferred in house and by hand, but the software and hardware of the museum might not be able to render the newer standards. Thus, the costs at large will need to be considered throughout this process. It is worth keeping in mind that some components of a project will grow beyond what was expected, and thus, will also become more expensive. Whether or not those added expenses can fit within the project budget will have to be assessed on a case-by-case basis.

Although the costs of actual image storage space are likely to remain relatively low, the costs of frequently improving, upgrading, and replacing technological equipment can be taxing. Looking at these changing standards very early in the process can give staff a much larger window of time during which to learn and master newly identified standards before their required implementation into the facility at large.

Section 4.2.7: Managing Collection Scope

Although standards and new innovations may change the requirements for the digitization process, the evolving needs and mission of the GWBHS can also necessitate changes. In

[21] An example Dublin Core description record for a digital photograph was provided in *Section 4.1*, and can illustrate a helpful example in undergoing the process. If the GWBHS would like any more examples beyond that, they can be easily found and retrieved online.

addition to evaluating whether or not equipment meets the needs of the collection at large, the digitization team should assess whether or not the digital records meets the needs of the user community. When that answer becomes "no," user feedback should be sought for aid in determining which content needs to be considered for de-accessioning, and what subsequent "new" items would warrant further accessioning in order to validate shifting audience needs. This will be discussed in greater detail in the following section on *Policy Reevaluation*.

Section 5: Policy Reevaluation

As an overall recommendation, the GWBHS should review and reevaluate the digitization policy every one to two years (although annual evaluation is recommended). Current standards should always be reflected in within policy documentation, but the evolving goals and needs of changing user audience should play a key role in policy development.

Throughout this paper we have discussed and attempted to reiterate the importance of policy reevaluation. This will logically take place at the end of each digitization phase, but can be required in additional places throughout the project. A reevaluation process should involve all workers involved with the digitization process (the "digitization team") at its first level. Any changes that have arisen and need to be reflected within the overall policy should be drafted as an addendum to the original policy. After the digitization team has agreed upon and approved said changes, the amended policy should be presented to the GWBHS board and ratified as the newest version of the policy. This process, although tedious, will ensure that all employees of the GWBHS maintain the same motivation to continue the process, share and develop new project goals, and understand the necessity and both preservation and user accessibility throughout the process.

In evaluating digitization, several aspects of the process should be considered:

- ***Project Assessment***
  The project manager should monitor the quality of the files that are being produced by the digitization team, and give them quality feedback on any problems with their workflow or end-product files. The *Digitization Manager* should provide feedback on employees' work output early (and often) within the process. The project manager should also make note of several use metrics once digital materials are released to the public. Types of materials with the highest levels of access, the types of users utilizing the artifacts, high-traffic web-use times, and levels of access and materials circulation

should be measured. This information could directly translate to the types of user-demands determined to be major criterion within the selection process.

- ***File & Metadata Management***
  As we have discussed earlier, files are likely to require migration to newer and more improved formats within their life span. Migration will ensure the preservation quality of digitized images, and will continue to deliver the accessibility of online collections to GWBHS patrons. When migration is deemed necessary and is undertaken by the digitization team, the newly determined file format should be reflected as the required standard within policy documentation, and added during the annual evaluation period (if not already explicitly listed within the policy).

  Rich metadata records can be described in extreme detail in several existing metadata schemas. Because the GWBHS is a small institution with a targeted clientele, complex algorithms and XML schemas are not realistic or necessary options. Thus, we have recommended the Dublin Core framework to be used within or outside of museum-digitization software. However, in the future, successive editions of appropriate software may cease to exist, and the museum will have to select a new software program to render the images viewable for the online community. Metadata may require updating and some manual editing in different environments, and all changes made to existing records should be recorded with explicit detail in a review document and submitted as an appendix to the overall GWBHS digitization policy.

- ***Selection***
  We described the importance of careful selection for digitization earlier in this project plan; however, a long-term selection consideration may not require *additions* to the repository, but *deletions* from it. In the instance that future analyses of digital collections find that some items and artifacts are out of scope, the GWBHS should carefully consider de-accessioning non-applicable items. This process could become vital for the continued mission of the museum if and when the space in the online storage environment begins to run out.

- ***Audiovisual Materials & Born-Digital Material***
  Digital technology improves rapidly, and during annual evaluation, the current state of multimedia and image standards will be measured. Audiovisual material is not the primary concern at the moment, but when other priorities have been cared for, they could eventually become important for the GWBHS to preserve.

In considering the preservation of their own materials, the GWBHS may consider creating or preserving other born-digital materials that reflect their institutional growth. Such creations (i.e. the GWBHS and other culturally-important websites in the region) could be considered valuable for archiving in the future. The GWBHS can capture this with web crawling software.

- ***Job Descriptions & Training***
  Job and labor descriptions should be amended as necessary at each annual policy reevaluation meeting. The nature and duties of the intern positions can, and likely will, change over the length of the project. The GWBHS should also record and reevaluate training processes offered to members of the digitization team. As phases and priorities of the institution evolve, the training will need to capture the most effective teaching mechanisms while adding new components that are relevant to modified policies and institutional needs. By keeping track of past successes and failures in training, the GWBHS can ensure that institutional knowledge of the digitization effort can be retained over time.

- ***Timeline & Costs***

  Reevaluation of the timeline is extremely important for the GWBHS to stay on-track and to understand the needs and abilities of its people, finances, and to balance those in accordance with user demands. A timeline is sure to require frequent readjustment, and all changes should be reflected within the digitization policy at large. In staying on par with transparency recommendations made earlier, we also believe that making the development and progress of the process visible to users will build a greater reputation and community trust in the repository.

  In tandem with the timeline are associated costs: just as a timeline requires flexibility, costs listed on the budget sheet may also change. An annual reevaluation will allow the GWBHS to project costs forward and to make sure that the project is still manageable. Newer vendors with different costs, interns with advanced capabilities and access to low-cost equipment, or other resources could all be identified in the future. The GWBHS should keep close tabs on evolving budgetary concerns, and schedule consistent meetings with the digitization team so that any necessary policy alterations can be identified and enacted as early as possible.

Conclusion

Within this plan we have aimed to give specific recommendations wherever possible, and include clear examples of drafted policies, examples of similar institutions' public efforts, and external links to more information. We hope that the GWBHS board will take many of these recommendations to heart, but by no means are they comprehensive or all encompassing, nor do they claim to be. The GWBHS will have to determine timelines, external vendors, *Digital* (and assistant) *Managers*, and designate appropriate interns and equipment—all while using our plan to structure a culturally fit policy on digitization. From our interactions and experiences with GWBHS staff, we have recognized a clear devotion to the greater West Bloomfield educational and user communities, to the sharing of cultural resources, and staff-wide dedication to the museum. These characteristics are what will help the museum to make well-informed decisions that are the most sensible for the institution, its mission, resources, and its community at large.

Worth noting, is that "there are no absolute rules for creating good digital collections. Every digital collection-building initiative is unique, with its own users, goals, and needs…Museums, libraries, archives, and schools have different constituencies, priorities, institutional cultures, funding mechanisms, and governance structures."[22] Thus, even the cited example policies and directed recommendations should be taken with a grain of salt. They are all thorough and well researched, but the GWBHS staff should carefully select bits and pieces of example policies, while touching on all of the specific needs we asserted within this project plan, to find a comfortable plan that fits their needs.

In concluding this document, we would foremost like to reiterate the necessity of policy development. Clear planning is essential for creating a comprehensive project that will restructure not only the GWBHS collections, but also the institution and its relationship with the community at large. Structured selection procedures will ensure that the most appropriate digital items are chosen for successive digitization phases, and appropriate job descriptions will further guarantee that those processes are being undertaken by persons skilled enough to produce an end-result worthy of long-term preservation.

Outside of preparation and policy development, specific measures should be taken to purchase and utilize appropriate equipment for digitization, beginning with phases that bring 2D content to digital form, and ending with outsourced 3D materials. Consistent metrics for including metadata, rendering those objects visible in the digital environment, and creating multiple copies of images (preservation-quality and user-accessible images) should be followed closely. We have carefully considered the budget (with some

[22] National Information Standards Organization http://framework.niso.org/node/7 ("how to use" section)

understanding that portions of the GWBHS savings through the generous bequest received last year would be allocated to digitization efforts) in making technical and quality assurance recommendations. Other recommendations were made according to professional, nationally recognized standards (not tailored to the GWBHS). Thus, the museum, in its aim to digitize collections comprehensively for preservation and access, will have to do the same: consider its own needs in tandem with professional standards.

Note of thanks:
Throughout this process, we have greatly appreciated the GWBHS' flexibility and dedication to instituting preservation practices. We have graciously been provided with timely feedback, extensive documentation, and honest expectations. In return, we hope—in the form of this project plan—to have provided the GWBHS with some of the baseline tools it needs to begin a digitization effort. In due time, we look forward to being the first among many to access the museum's wealth of historical information in its digital form.

Appendices

Appendix A: Intern Job Descriptions

Appendix A.1.Title: Digitization Intern

Formal information:
The position of Digitization Intern at the Greater West Bloomfield Historical Society is a seasonal, full-time, and unpaid position. The intern will be reporting to members of different inter society committees, as well as the project manager. The Greater West Bloomfield Historical Society is digitizing their collection to ensure widespread online accessibility and to meet the needs of the modern researcher.

Duties:
- Management and organization of archival digital files.
- Scanning and digitization of 2-dimensional objects.
- Saving files into access copies from original scanned format.
- Accessioning 2-dimensional objects into a digital space.
- Light cataloging of collection.

Minimum qualifications:
- Experience with Adobe Photoshop.
- Experience with editing of digital images.
- Effective communication skills, both written and spoken.
- Ability to handle sensitive museum materials.
- Ability to learn and use computer software necessary for performance of job duties.
- Ability to maintain focus while performing repetitive tasks.

Desired qualifications:
- Familiarity with operation of flatbed scanner.
- Familiarity with OAIS model.

Appendix A.2. Title: Metadata Creation Intern

Formal information:
The position of Metadata Creation Intern at the Greater West Bloomfield Historical Society is a seasonal, full-time, and unpaid position. The intern will be reporting to members of different inter society committees, as well as the project manager. The Greater West Bloomfield Historical Society is digitizing their collection to ensure widespread online accessibility and to meet the needs of the modern researcher. The Metadata Creation Intern will ensure continuity in metadata and continued accessibility through the creation of XMP records.

Duties:
- Creation of valid XMP records.
- Connection of XMP/Dublin Core records to digital photo files.

Minimum qualifications:

- Effective communication skills, both written and spoken.
- Experience with creation of XMP records.
- Ability to learn and use computer software necessary for performance of job duties.
- Ability to maintain focus while performing repetitive tasks.
- Proven experience working with XML metadata schemas, especially Dublin Core.
- Ability to handle sensitive museum materials.
- Experience with Adobe Photoshop

Desired qualifications:
- Research skills.

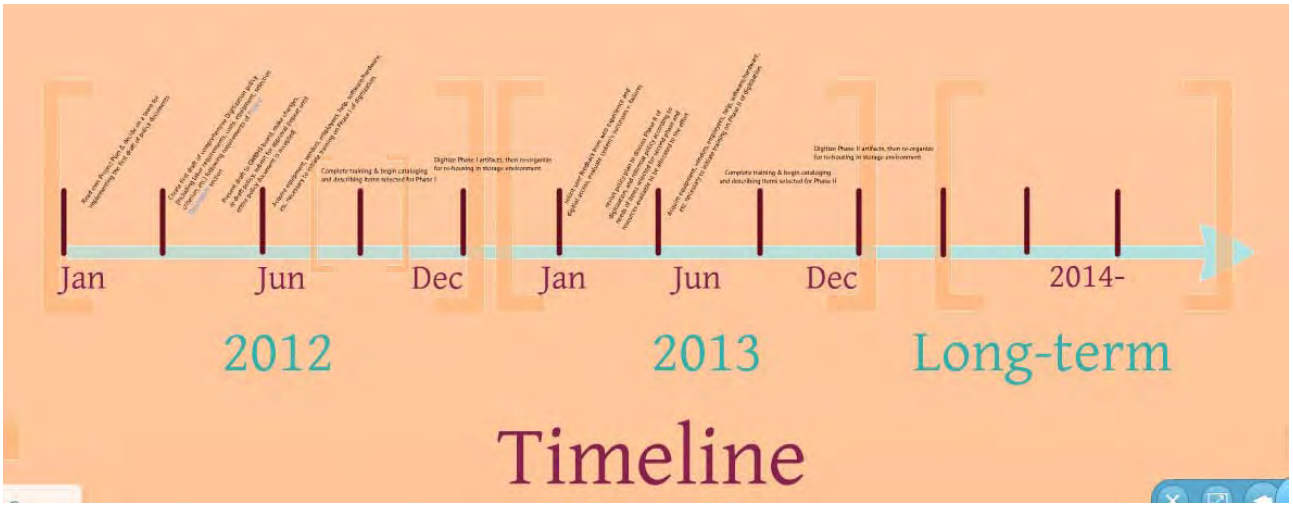
Appendix B: Budget for Project

Cost Analysis Worksheet

Item	Quantity	Unit	Cost rate	Total	Notes and assumptions
SUPPLIES AND EQUIPMENT					
Epson V750-M Scanner	1	1	$850.00	$850.00	
Computer (if needed)	1	1	$700.00	$700.00	
Adobe Photoshop CS5	1	1	$300.00	$300.00	Non-profit License
Monitor Calibration Software	1	1	$175.00	$175.00	Monitor calibration software ensures that digital files' blacks and whites/colors
Total Supplies & Equip Costs				**$2,025.00**	
SERVICES					
Digitization Vendor				TBA	Completely dependent on how much of your collection you choose to digitize through this venue
Total Services Cost				**TBA**	
LABOR					
Intern	2	2	$0.00	$0.00	We have listed the interns as unpaid positions. If you would like to pay them, that is an option. Will still require training.
Volunteers	TBA	TBA	$0.00	$0.00	Will require training.
Technical Support	1	1	TBA	TBA	Having someone knowledgeable in IT on call would be good in case any part of the hardware fails.
Total Labor Cost				**TBA**	
INDIRECT COSTS					
Misc. Materials				TBA	Markers, materials, sleeves, papers, ect.
Poss. External Facility				TBA	2D digitization may require use of an external facility
Extra Building Costs				TBA	Costs of running the building extra hours (in terms of electricity, heating)
Total Indirect Cost				**TBA**	
Project Total				$2,025.00	Total does not include Services, a possible external facility, nor extra building functioning costs

Appendix C: Timeline

Pictured are recommended benchmarks for the GWBHS digitization project.



Appendix D: Adobe Photoshop Technical Specifications

System Requirements for Digitization Computer Systems (from Adobe Photoshop CS5 minimum system requirements)

Adobe Creative Suite family /

Adobe Photoshop CS5 / Tech specs

Overview | Features | **Tech specs** | Reviews | FAQ | Showcase | Extend | Buying guide

System requirements

Windows

- Intel® Pentium® 4 or AMD Athlon® 64 processor
- Microsoft® Windows® XP with Service Pack 3; Windows Vista® Home Premium, Business, Ultimate, or Enterprise with Service Pack 1 (Service Pack 2 recommended); or Windows 7
- 1GB of RAM
- 1GB of available hard-disk space for installation; additional free space required during installation (cannot install on removable flash storage devices)
- 1024x768 display (1280x800 recommended) with qualified hardware-accelerated OpenGL graphics card, 16-bit color, and 256MB of VRAM
- Some GPU-accelerated features require graphics support for Shader Model 3.0 and OpenGL 2.0
- DVD-ROM drive
- QuickTime 7.6.2 software required for multimedia features
- Broadband Internet connection required for online services and to validate Subscription Edition (if applicable) on an ongoing basis*

Mac OS

- Multicore Intel processor
- Mac OS X v10.5.8 or v10.6
- 1GB of RAM
- 2GB of available hard-disk space for installation; additional free space required during installation (cannot install on a volume that uses a case-sensitive file system or on removable flash storage devices)
- 1024x768 display (1280x800 recommended) with qualified hardware-accelerated OpenGL graphics card, 16-bit color, and 256MB of VRAM
- Some GPU-accelerated features require graphics support for Shader Model 3.0 and OpenGL 2.0
- DVD-ROM drive
- QuickTime 7.6.2 software required for multimedia features
- Broadband Internet connection required for online services and to validate Subscription Edition (if applicable) on an ongoing basis*

**Adobe online services, including Adobe CS Live Services, are available only to users 13 and older and require agreement to additional terms and Adobe's online privacy policy. Online services are not available in all countries or languages, may require user registration and may be subject to change or discontinuation without notice. Additional fees or subscription charges may apply.*

Appendix E: Dublin Core Metadata Framework

A simple **Dublin Core Metadata Element Set (DCMES)** includes 15 metadata elements. As many fields as information is known for should be included for each record digitized within the GWBHS. For more information on how, when, and where to use metadata description, see the Dublin Core website: (http://dublincore.org/documents/usageguide/elements.shtml). Below, the 15 elements have been included with brief descriptions:

1. **Title** - *name given to the resource. Typically, a Title will be a name by which the resource is formally known.*
2. **Subject** - *The topic of the content of the resource. Typically, a Subject will be expressed as keywords or key phrases or classification codes that describe the topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.*
3. **Description** - *An account of the content of the resource. Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.*
4. **Type** - *nature or genre of the content of the resource. Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary. To describe the physical or digital manifestation of the resource, use the format element.*
5. **Source** - *A Reference to a resource from which the present resource is derived. The present resource may be derived from the Source resource in whole or part. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.*
6. **Relation** - *A reference to a related resource. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.*
7. **Coverage** - *The extent or scope of the content of the resource. Coverage will typically include spatial location, temporal period, or jurisdiction. Recommended best practice is to select a value from a controlled vocabulary. Where appropriate, named places or time periods should be used in preference to numeric identifiers (i.e. date ranges).*
8. **Creator** - *An entity primarily responsible for making the content of the resource. Examples of a Creator include a person, an organization, or a service. Typically the name of the Creator should be used to indicate the entity.*
9. **Publisher** - *The entity responsible for making the resource available. Examples of a Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.*
10. **Contributor** - *An entity responsible for making contributions to the content of the resource. Examples of a Contributor include a person, an organization or a service. Typically, the name of a Contributor should be used to indicate the entity.*
11. **Rights** - *Information about rights held in and over the resource. Typically a Rights element will contain a rights management statement for the resource, or reference a service providing such information. Often encompasses Intellectual Property Rights , Copyright. If the rights element is absent, no assumptions can be made about the status of these and other rights with respect to the resource.*
12. **Date** - *A date associated with an event in the life cycle of the resource. Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for*

encoding the date value is defined in a profile of ISO 8601 (*Date and Time Formats, W3C Note, http://www.w3.org/TR/NOTE- datetime*) *and follows the YYYY-MM-DD format.*

13. **Format** - *The physical or digital manifestation of the resource. Typically, Format may include the media-type or dimensions of the resource. Examples of dimensions include size and duration. Format may be used to determine the software, hardware or other equipment needed to display or operate the resource.*
14. **Identifier** - *Typically a reference to the item or collection using a string or number conforming to a formal identification system. Examples include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).*
15. **Language** - Language of the intellectual content of the resource should be recorded.

Bibliography

- North Carolina Department of Cultural Resources, "Best Practices for File Naming." 7 May 2008. Web: 11 Dec. 2011 http://www.records.ncdcr.gov/erecords/filenaming_20080508_final.pdf

- *Espon Perfection V700 Photo/V750-M Professional*. Tech. Epson. Web: 11 Dec. 2011 http://www.epson.com/cmc_upload/0/000/142/238/V700_V750_InfoSheetR1.pdf

- Fleischhauer, Carl, Louis H. Sharpe II, and D. Michael Ott. *Manuscript Digitization Demonstration Project. 5. Preservation-quality Images: Types and Resolution*. Library of Congress. Web: 11 Dec. 2011 http://memory.loc.gov/ammem/pictel/mddp305.htm

- Gertz, Janet. "Preservation and Selection for Digitization." *Northeast Document Conservation Center - Preservation Leaflets*. Web: 8 Dec. 2011: http://www.nedcc.org/resources/leaflets/6Reformatting/06PreservationAndSelection.php

- Hurst-Wahl, Jill. "Digitization Vendors at the SLA Conference, part 3." 10 Jun. 2007. Web: 11 Dec. 2011. http://hurstassociates.blogspot.com/2007/06/digitization-vendors-at-sla-conference_10.html

- "Joyner Library Policy Manual: Digital Collections Development Policy" East Carolina University. 5 Feb. 20009. Web: 8 Dec. 2011. http://digital.lib.ecu.edu/info/DigitalCDP.pdf

- National Information Standards Organization, "A Framework of Guidance for Building Good Digital Collections." 16 Apr. 2008; Web 07 Dec. 2011. http://framework.niso.org/node/7

- "NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials." National Initiative for a Networked Cultural Heritage (NINCH),

Humanities Advanced Technology and Information Institute (HATII), and University of Glasgow. Oct. 2002. Web: 9 Dec. 2011. http://www.nyu.edu/its/humanities/ninchguide/

- National Library of Australia. “Collection Digitisation Policy,” Aug. 2009. Web: 9 Dec. 2011. http://www.nla.gov.au/policy/digitisation.html

- Peterson, Kit A. *Digital Master Images - Sample Technical Specifications for Photograph Collections*. Tech. Library of Congress, June 2004. Web: 11 Dec. 2011. http://www.loc.gov/rr/print/tp/DgtlMastersSamplSpecsSelctdRcmndFinal7_2004.pdf

- Schewe, Jeff. "About Metadata." Adobe, 2004. Web: 11 Dec. 2001. http://www.adobe.com/digitalimag/pdfs/about_metadata.pdf

- Still Image Working Group. “Technical Guidelines for Digitizing Cultural Heritage Materials”24 Aug. 2010. Web: 10 Dec. 2011. http://www.digitizationguidelines.gov/guidelines/digitize-technical.html

- “Sustainability of Digital Formats: Planning for Library of Congress Collections.” National Digital Information Infrastructure and Preservation Program (NDIIPP). Web: 8 Dec. 2011. http://www.digitalpreservation.gov/formats/sustain/sustain.shtml

- “Trusted Repositories Audit Certification (TRAC): Criteria and Checklist,” OCLC, Center for Research Libraries (CRL) and the National Archives and Records Administration (NARA). Feb, 2007. Web: 9 Dec. 2011. http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

- Zinkham, Helena. *Common and Useful Information Elements for Cataloging Pictorial Materials*. Tech. Library of Congress, June 2004. Web: 11 Dec. 2011. http://www.loc.gov/rr/print/tp/Common%20Information%20Elements.pdf