Digital Preservation Plan Mariposa Museum and History Center

Created by Diane Mello, University of Washington

Contributor Miranda Fengel, Curator, Mariposa Museum and History Center

> Submitted to Nic Weber, PhD LIS 598 - Digital Preservation Winter 2020

> > Date: 03/11/2020

The Mariposa Museum and History Center (MMHC) participates in the creation and maintenance of the emerging global digital library by digitizing and sharing electronic information. This document was created to describe the rationale and strategies for digital preservation of the MMHC archive of photographs.

1. Rationale for digital preservation plan.

The MMHC Digitization Plan will be developed in accordance with the objectives of the MMHC Strategic Plan, Collection Development and Accession policies. The plan is intended to provide an accession and digitization framework to select, create, maintain, monitor, and provide access to electronic resources which support the scholarly and informational needs of the county of Mariposa, California. Through this commitment and process, the museum makes these resources available to the wider community, thereby contributing to the educational, cultural, and economic well-being of the region, their partners and beyond. By expanding and enhancing access to published and unpublished materials of potential interest, especially those unique to the county of Mariposa California, MMHC pursues its mission which supports the discovery, collection, preservation, interpretation, exhibition and demonstration for material culture, visual objects and photographs that best illustrate the history of Mariposa County.

Digitization is the process by which an electronic representation of a tangible item is created. Transferring materials from tangible to digital format protects the original item from damage while providing public access to materials that might otherwise be restricted. In addition, the implementation of a digital preservation plan supports the goals towards creating a sustainable and enduring collection. There is unique and specific information within printed materials and documents such as letters, photos, postcards, etc. However, many of these items may exist in poor condition, and/or are too fragile for frequent use. The museum's plan will digitize materials of great interest and items whose present format obstructs ease-of-use. Through the museum's discovery system, users will have the ability to access digitized information remotely. By digitizing the unique documents, researchers no longer have to travel to the location where the document is held and the collection gains prominence and relevance among interrelated collections.

Currently, digital photographs are processed and stored through Past Perfect 5 (PP5), a collection management software, on the museum desktop hard drive and maintained by the curator. PP5 allows assignment of permissions related to roles, responsibilities and full administrative control to set up its relational database for browse and keyword functions to search the collection. Currently, access to the digitized images for research and education require onsite visits and use of the museum desktop workstation.

In adopting the plan; security, discoverability, and access are prioritized. To ensure prevention of loss for the images, digital photograph formats are stored on an external server or portable drive offsite. This step mitigates the risk of loss of digital objects and supports retention of the physical integrity, identity, and authenticity of the digital files. Administrative, technical and descriptive metadata for electronic images will be determined by administrators, archivists and designated community partners. The implementation of a partnership with Information Technology staff ensures knowledge and expertise regarding interfacing with PP5 folders and files data, fixity, network setup, server support, storage, maintenance, and redundancy. In addition, MMHC benefits from established partnerships with the University of California online library, museum and archive digital collections communities such as Calisphere and Online Archive of California which offers remote access to the photographic collection.

2. Statement of organizational commitment

Fundamental to a successful long-term preservation of MMHC digital collection files, several entities must engage and remain committed to the establishment and development of the photographic archive.

As a non-profit, the Curator, Board of Directors, staff, volunteers, scholarly partners, designated community members, and information technology staff benefit from their collaborative and collectively defined archive.

Currently museum staff includes a Curator who administers the collection and supervises two staff members and numerous volunteers/docents. Collections intended for cultural heritage may

be subject to review by a designated community, its members and/or an acquisition committee made up of scholars and designated community members who share expertise of the archaeological, biographical, ethnological, and history of the object.

Regular recruitment and training for archival staff and volunteers support the continuous sharing of historical knowledge and archival processes thus sustaining the museum's ability to operate with consistent standards and without interruption.

3. Statement of Financial Commitment

A realistic budget for initial set-up and long-term archive maintenance is developed by the MMHC Curator and Board. The foundational step to collaborate and include Information Technology staff (IT) prioritizes preservation, an enduring collection and scalable requirements of the digital process and infrastructure including digital housing, staff, archival standards and processes and timelines.

Currently the museum relies on grants, direct funding, in kind donations and donation of time by volunteers. MMHC's Curator will collaborate with the Board, IT and online collection partners to define the benefits, costs and requirements of scalable digital storage and remote public access. IT staff will assess the museum, its local area network and establish online and local access points to the collection by the public. Financial resources will be prioritized for IT support and archival staff processing hours, online collection partnerships, and data and server storage. Initial time investments will be required of administrators, staff and volunteers to train and learn the preservation process therefore long- term financial commitments for supporting the costs will be required to sustain a long-term preservation plan.

4. Preservation Actions and Quality Control (See table).

The MMHC digitization process consist of two primary functions:

- 1) Preserve information to secure its long-term sustainability and persistence.
- Provide access to the archived information consistent with needs of the archive's primary users or designated community.

The following steps will support these functions:

- Set up the folders and files in Past Perfect 5. Define administrative and descriptive metadata using Dublin Core metadata schema.
- Verify security and standards for additions/deletions for archive.

- The Curator has administrative status.
- Staff and volunteers have 'add' only status for tasks and require approval.
- Prepare image: Based on MMHC policy, refer to museum's selection criteria and eligibility prior to addition/deletion to archive.
- **Submission Information Package:** Ingest digital object(s) with verifiable information via Past Perfect 5 software interface. Check for fixity of image.
- Archival Information Package: Create record of image. Include descriptive, administrative, and technical metadata.
- **Dissemination of Package:** Metadata for digital objects will be distributed to PP5 relational database with item-level metadata where appropriate.
- Administration: Since PP5 does not include monitor capabilities such as Checksum or Fixity, IT will perform preservation audits and backups periodically to verify integrity of content and metadata. Redundant Array of Inexpensive Disks or Redundant Array of Independent Disks (RAID) drives will also be purchased and maintained offsite to ensure integrity of files and these will allow Checksum and fixity processes.
- Digital content will be made available online in formats appropriate for online delivery, which may differ from formats used for preservation (e.g. lower resolution images, rather than downloadable files). Metadata for digital material will be added to the library catalog at least the collection level, with item-level metadata where appropriate.

Digitization specifications have been developed for different material types to provide a method of evaluating the quality of images produced, to estimate the data storage for access files (on-line) and master files (offline), and to assist in maintaining a certain level of quality control. Inherent differences in document types dictate different approaches to scanning. Due to the desire to image documents in a consistent manner, all scanning is done in 8-bit grayscale or 24-bit color.

Media formats provided for scanning include B&W and color photographic prints, slides, and negative transparencies. The digital master files are produced at 400-600 dpi based on the size of the image format. Images are scanned to TIFF for archival copy.

Scanning resolution adjusted to produce files with pixel arrays of 6000 pixels across the long image dimension by the proportional number of pixels for the specific photo format,

or at 400-600 dpi, depending on the scanning requirements. Scanner shall provide a true optical resolution of at least 6000 pixels across the long dimension of the image. Interpolating to a higher resolution from a lower resolution scan shall not be permitted. File Format: Uncompressed TIFF or PDF.

Suggested Storage - IT staff will support storage and backup processes: Preservation master image files will be stored on hard drive systems with data redundancy, such as (RAID) drives. Additionally, another set of images with metadata will be stored on an LTO-6 tape format and will be archived on Amazon's Glacier long term archival solution. Regular backups of the images onto tape from the RAID drives will be performed with checksums stored alongside image files. Regular backups of local archives will be copied over to Amazon Glacier for long term archival.

Attribute	Level 1: Protect Data	Level 2: Know Data	Level 3: Monitor Data	Level 4: Repair Data
Storage and Geographic Location	Maintain one high resolution copy (offsite) Maintain two low resolution copies (onsite). Create electronic, natural and human error disaster plans.	Document storage system and media including hardware, software and OS. Identify all services, etc needed to use.	Monitor obsolescence for storage systems and media. Maintain a copy of the image with a disaster plan in a different geographic location.	Locate at least three copies in geographic locations. Have a disaster plan which keeps files and metadata on an accessible media or system.
File Fixity and Data Integrity	Check file fixity on ingest. Create fixity if it is not provided.	IT performs periodical checks, backups and audits. Maintain logs.	Run periodic backups, check sums and audits. Maintain a log of fixity info and be able to supply audit on demand. Assign IT tasks to identify corrupt data and virus checks.	Respond to specific event or activity with an ability to replace/repair data.
Information Security	Assign read, write, move and delete permissions to	Maintain logs and document permissions to content. (People	Maintain logs of individual activity with preservation	Regularly audit security activity logs.

Preservation Action and Quality Control

	individual files. Ensure no one person has write access to all copies.	and codes)	actions, additions, deletions and file activity.	
Metadata	PP5 creates inventory of content and storage- Verify. Ensure backup of descriptive data without collation of inventory.	Review PP5 storage of administrative, transformative date and events.	Review and store standard technical and descriptive data.	Review and store the preservation data.
File Formats	Define naming standards of digital file format.	Record and review inventory of file formats in use in PP5	Monitor file formats for obsolescence issues.	If necessary, perform format migrations, emulation and /or similar activities to restore records and images.

5. Metadata Creation

The descriptive, administrative, technical and structural metadata refers to the elements required for discovery and identification of a resource. It describes bibliographic and physical attributes of the material while associating access points. These points are structured in one or more standardized and published schemes such as Dublin Core or MARC and are compatible with Past Perfect 5 collection management system.

Dublin Core Term	Definition	Note/Guideline	Example
Title	The name given to the resource	Capitalize words in Title field. Include date if applicable.	Covered Bridge Over South Fork of Merced River, Wawona, Yosemite, 1955
Creator	The person(s) or organization(s) primarily responsible for the intellectual content of the resource; the author or originator	The photographer's name. Format: Last Name, First name	Adams, Ansel
Subject	The topic of the resource; also, keywords, phrases or	Use the Library of Congress Subject heading.	BridgesCalifornia BridgesYosemite RiversCalifornia

 Table: Example Data Dictionary using DublinCore

	classification descriptors that describe the subject or content of the resource.	http://id.loc.gov/autho rities/subjects.html	RiversYosemite
Description	A textual description of the content.	Describe what is depicted in the image or summary of the photograph. Describe condition of physical item. Copy any metadata written on front or back of photo	A black and white image of the covered bridge over the South Fork of the Merced River with a stagecoach parked in front of the entrance. The Merced river is in Mariposa county.
Publisher	The organization responsible for making the resource available in its present form.	Yosemite Research Library, PO Box 577, Yosemite National Park, CA 95389 unless otherwise stated	Yosemite Research Library, PO Box 577, Yosemite National Park, CA 95389
Contributor	Person(s) or organization(s) in addition to those specified in the CREATOR field, who have made a significant intellectual contribution to the resource, but on a secondary basis.	Contributors may be photographer's assistant, people who added to the image. (artist, craftsperson). Format Last name, First name	Sexton, John
Date	The date the resource was made available in its present form.	Use format YYY-MM- DD. If date is unknown or approximate use YYYY.	1955-08-06
Туре	The resource type, such as home page, novel, poem, working paper, technical report, image, audio, etc. In this case: Image,	Choose from: Image, Document	Image
Format	The representation of the resource, such as glass plate negative, etc.	List digital surrogates and original format. Note technical information: scanner used and dpi. Choose from photograph, glass plate negative, etc.	Original photographic prints were scanned as 300 dpi TIFF files on a NAME OF SCANNER. 72 dpi JPEG files were then added to the CONTENT database

			at the MMHC collection.
Identifier	A string or numbers or letters used to uniquely identify the resource. This might be the file name if digital, or a unique ID or catalog number if physical,	Use item number	123
Source	The collection or work, either print or electronic, from which the resource is taken.	Collection number. Also used only if image is taken from a larger work like a bound volume.	AA 50
Language	The language(s) of the intellectual content of the resource.	List multiple if needed. Check with supervisor if you identify a language you do not know	English
Coverage	The geographic locations and time of the resource, if applicable	List area: Area and time period.	Mariposa County, CA. Mid 19th Century
Relation	The relationship to other sources.	Use if image is taken from a larger work like a bound volume, or if other images are closely related	AA 50 is a copy of the glass plate negative, etc.
Rights	A rights statement or a link to a general copyright notice, or a rights-management statement for MMHC.	Use MMHC general rights statement.	For permission to publish please contact Mariposa Museum and History Center (MMHC) (209) 966-2924. MMHC does not claim to control the copyright for all pictorial materials in its collections. MMHC will provide any information it has available to the user in this regard.

6. Roles and Responsibilities

The Curator of MMHC and the Board of Directors contribute to the administration, oversight and decision of the museum, its collection and the digital preservation and lifecycle of digital content in the MMHC archives.

The Curator is qualified to negotiate between informational technologists, archivists, board members, designated community members, staff, volunteers and service providers for the financial, technical and preservation needs of the collection. This includes managing an archive committee and understanding staff skillsets, areas of future learning, and technology to maintain and support workflow and an enduring, sustainable collection.

Information Technology staff (Network Engineers, System Engineers) is qualified to maintain and administer computer networks and related computing environments including systems software, applications software, hardware, and configurations. In addition, protecting data, software, and hardware by coordinating, planning and implementing network security measures.

Archivist skills include the ability to authenticate and appraise historical documents and archival materials. They will create and maintain accessible, retrievable computer archives and databases, incorporating current advances in electric information storage technology. The archivist for MMHC will develop a working knowledge of the collection management program - Past Perfect 5 and provide facilitation of the digital processes to administrate, deposit, ingest, search and create access of the collection. In addition, the archivist will collaborate with the Curator, Board of Director and designated community to define metadata needs for MMHC.

7. Training and Education

The staff at MMHC will commit to ongoing learning and development while observing the best practices of digital preservation. In order to best steward the museum's presence in national, international, consortia and local governing bodies, the following areas of awareness will be maintained: current preservation practice, information lifecycle management, changes in information storage management and systems maintenance, and MMHC copyright standards.

8. Monitoring and Review

The MMHC archive will demonstrate compliance with the Online Archive Information System (OAIS) with IT support through monitoring and review procedures as stated in the preservation action and quality control guidelines.