



American Heritage Center

Digitization Committee Mission

Approved 06 June 2003

The charge of the committee is to provide oversight and coordination of digitization activities at the AHC. In this context, digitization activities are defined as those projects in which the AHC creates, or participates in the creation of, scholarly content made available in electronic form. Projects may include the development of appropriate indexing and navigational tools, metadata, interfaces, and other delivery/access tools for resources requiring digitization or already in digital form.

The committee will develop internal guidelines, establish standards and best practices for digital projects, and, where possible, adopt strategies for interoperability among projects mounted by the AHC and those mounted by other organizations within and outside of the University of Wyoming. The committee will advocate and provide information on the staff resources, and hardware and software needs, and monitor the AHC's scanning and other use fees.



American Heritage Center

Digitization Guidelines

Approved 06 June 2003

INTRODUCTION TO DIGITIZATION AND BEST PRACTICES

This document provides recommendations for selection of materials, image quality, file formats, how to describe the materials being digitized, access mechanisms, and other capture and storage issues when converting paper, photographic and other physical materials into digital form.

By adopting community shared good practice, the AHC can ensure the broadest use of its materials, today and in the future, by audiences they may not even have imagined and by future applications that will dynamically recombine digital objects into new resources. The National Initiative for a Networked Cultural Heritage (NINCH) has proposed six core principles for digitization best practices. They are:

- 1. Optimize interoperability of materials.** Digitization projects should enable the optimal interoperability between source materials from different repositories or digitization projects.
- 2. Enable broadest use.** Projects should enable multiple and diverse uses of material by multiple and diverse audiences.
- 3. Address the need for the preservation of original materials.** Projects should incorporate procedures to address the preservation of original materials.
- 4. Indicate strategy for life-cycle management of digital resources.** Projects should plan for the life-cycle management of digital resources, including the initial assessment of resources, selection of materials and digital rights management; the technical questions of digitizing all formats; and the long-term issues of sustainability, user assessment, digital asset management and preservation.
- 5. Investigate and declare intellectual property rights and ownership.** Ownership and rights issues need to be investigated before digitization commences and findings should be reported to users.
- 6. Articulate intent and declare methodology.** All relevant methods, perspectives and assumptions used by project staff should be clarified and made explicit.

In addition to the NINCH guidelines, some basic principles and considerations governing digitization:

- Materials should never be digitized more than once and digitized at the highest level reflecting the material's intellectual content.
- Once digitized, materials shall be stored in an archivally sound manner and accompanied by appropriate administrative, structural, and descriptive metadata.
- Is there staff available to fulfill digitization request?

- Is the amount of time required to digitize reasonable when balanced against other departmental priorities?
- Does the AHC own the equipment and software necessary to complete the project?
- Has the AHC ever worked with this kind of material/technique before?
- Can the request be fulfilled within the stated timeline?
- Is there any question about whether or not the material can be digitized?
- Does metadata exist and is there expertise available to create the required metadata for the original or digital items?
- What kind of care and planning is required to digitize valuable or fragile material?
- How much storage space will the project require, and will archiving and file migration needs. How will it impact staff and computer resources extensively?
- Except insofar as surrogates can lessen or eliminate risk to documents, digitization is not a preservation medium.

SELECTION CRITERIA

The creation of digital versions of print, photographic, archival, and other materials may be undertaken for a variety of purposes, ranging from in-depth research to popular interest, and from large-scale projects to conversion of a few items for an individual's use. It is expensive to select, create, and maintain digital resources; the costs of image capture accounts for only one-third of the total expense. It is therefore important to assure during the selection process that issues of technical feasibility, intellectual property rights, and institutional support are considered along with the value of the materials and the interest of their content. The criteria listed below are neither exhaustive nor prescriptive, and their relative importance will vary depending on the purpose for which digitization is proposed.

The value of the materials' content and the benefits derived from access to digital versions justify the expenditure of time and effort of carrying out a digitization project. The content should have sufficient intrinsic value to ensure ongoing use by a defined constituency on and/or beyond the University of Wyoming for a significant period of time.

1. **Value.** Many factors contribute, but certainly they include intellectual content, historic, and physical value.
2. **Demand.** To justify the effort and expense, there should be a reasonable expectation that the product will have immediate utility.
3. **Non-Duplication.** There is no identical or similar digital product that can reasonably meet the expressed needs.
4. **Collaborative Potential.** Part of a collection split among a number of institutions that could be united online as a virtual collection, or contribution to development of a "critical mass" of digital materials in a subject area.
5. **Enhancement of Intellectual Access.** Enhancement of intellectual control through creation of new finding aids, links to bibliographic records, and development of indices and other tools.

Intellectual Property Rights Criteria

Intellectual property rights must be managed in accordance with applicable laws, and any necessary restrictions to access must be able to be implemented through current institutionally supported mechanisms. Considerations include whether:

- The work or collection is in the public domain.
- The goal/purpose of digitization lies within fair use limits.
- The AHC owns the legal right to make and disseminate digital copies.
- The AHC can get clearance from the holder of the rights.
- Privacy issues can be addressed.

For both published and unpublished material, see Peter Hirtle's *When Works Pass into Public Domain*, Web site from Cornell at <<http://cidc.library.cornell.edu/copyright/>>.

Preservation Criteria

While digitization does not in itself constitute preservation, there are preservation aspects to be considered, both in terms of the original materials and in terms of the files that will be created. These are:

- Digital surrogates allow significant reduction in handling of fragile materials, access to materials that cannot otherwise be easily used, and protection of materials at high risk of theft or mutilation.
- The condition of originals allows them to be digitized safely, or possibility of scanning photographic intermediaries instead of the originals.
- If the digital resources are deemed to be of long-term value, provision must be made for archiving and maintenance of images through time and technological change.

Technical Feasibility Criteria

Potential projects should be evaluated as to whether it is technically possible with current equipment and software to capture, present, and store images in ways that meet user needs. Considerations include:

- The degree to which a digital version can represent the full content of the original.
- Understanding of how people will use the digital versions and the level of image quality that that implies.
- Whether the materials will display well digitally.
- The capacity for accessing images from current institutionally supported platforms and networked environments, and delivering them with reasonable speed.
- Anticipation of future users with better equipment, to avoid a need to rescan in a few years.
- Authentication of images, if appropriate.
- Staff and resources to support programming, user interface design, and search engine development to assure that the project can fulfill the functions for which digitization is planned.

Intellectual Control Criteria

Potential projects should be evaluated as to whether appropriate intellectual control can be provided for the original materials and the digital versions:

- The degree to which the materials are organized/arranged in a way suited to online use.
- Cataloging, processing and related organizational work already accomplished or to be accomplished as part of the project.

- Staff and resources to support creation of appropriate metadata relating to document identification, technical capture information, provenance, and easy navigation within the information resource.

DIGITIZING ITEMS FROM AHC COLLECTIONS

All AHC digitization activities are required to follow some basic procedures. A checklist must be completed for every project, and the information may be collected online and/or stored centrally. Materials must be digitized according to the specifications set in this document, taken from the Western States Digital Imaging Best Practices Guidelines <http://www.cdpheritage.org/westertrails/wt_bpscanning.html> and other sources.

Quality Control

To assure high quality digital files, each image will be evaluated twice. The scanning technician will examine the histogram at the time of capture and creation of the master file. At random, the master files are examined for:

- File format
- Resolution
- Color
- Tone
- Detail of reproduction and resolution
- Noise
- Compression effects

Files failing to meet the above criteria will be rescanned. No adjustments in PhotoShop or other image manipulation software will be allowed to correct mistakes made in the scanning process.

Once the master files pass quality control, the thumbnail and reference files will be created. The technician will need to make some adjustments to contrast and possibly cut back frame edges when creating these trimmed-down files. Because of the manipulations necessary in PhotoShop to create images for on-screen viewing, the thumbnails and reference files will undergo a more limited set of quality control procedures. When the technician is satisfied that the files will pass quality control, a random selection will be made and checked for quality control by a designated person. Following that check, the files will be moved onto the final storage media and metadata creation.

Please also reference the WSDIBP document for modes of capture, spatial resolution, bit depth, and format.

Text Collections

	Master	Access	Thumbnail
File Format	TIFF	JPEG	JPEG or GIF
Bit Depth	1 bit bitonal 8 bit grayscale 24 bit color	1 bit bitonal 8 bit grayscale 24 bit color	1 bit bitonal 8 bit grayscale 8 bit indexed color (GIF) 24 bit color
Spatial Resolution	600 ppi	150 dpi	72 dpi
Spatial Dimensions	100% of original	600 pixels along the long dimension	150-200 pixels across the long dimension

Photographs

	Master	Access	Thumbnail
File Format	TIFF	JPEG	JPEG or GIF
Bit Depth	8 bit grayscale 24 bit color	8 bit grayscale 24 bit color	8 bit grayscale 8 bit indexed color (GIF) 24 bit color
Spatial Resolution	3000 to 5000 pixels across the long dimension	150 dpi	72 dpi
Spatial Dimensions	100% of original	600 pixels along the long dimension	150-200 pixels across the long dimension

Maps

	Master	Access	Thumbnail
File Format	TIFF	JPEG	JPEG or GIF
Bit Depth	8 bit grayscale 24 bit color	8 bit grayscale 24 bit color	8 bit grayscale 8 bit indexed color (GIF) 24 bit color
Spatial Resolution	300 ppi or 3000 pixels across the long dimension, whichever is greater	150 dpi	72 dpi
Spatial Dimensions	100% of original	600 pixels along the long dimension	150-200 pixels across the long dimension

Graphic Materials

	Master	Access	Thumbnail
File Format	TIFF	JPEG	JPEG or GIF
Bit Depth	8 bit grayscale 24 bit color	8 bit grayscale 24 bit color	8 bit grayscale 8 bit indexed color (GIF) 24 bit color
Spatial Resolution	3000 pixels across the long dimension	150 dpi	72 dpi
Spatial Dimensions	100% of original	600 pixels along the long dimension	150-200 pixels across the long dimension

Audio and Video Collections

Approaches to audio and video digitization may depend somewhat on the quality and format of the original. Extremely low fidelity recordings may not benefit from high frequency capture. Higher end audio standards, such as the standard for DVD Audio, are still emerging. Therefore, in keeping with a use-neutral and technology-neutral approach, the following standards shall be applied. For further information, see the Library of Congress AV prototyping project:

<<http://www.loc.gov/rr/mopic/avprot/audioSOW.html>> and the Colorado Digitization Program's Digital Audio Guidelines: <<http://www.cdpheritage.org/resource/audio/index.html>>.

Note that a digital master must be generated for every object digitized, whether or not service files are also generated.

METADATA

Digitized materials must be accompanied by metadata which describes the content and nature of the material, the process of digitization, its source, and any structural data necessary to describe relations between parts or to other objects. Metadata should be stored and transmitted so that future users can find and access the digital material. Certain basic procedures should be adhered to:

- Metadata should be stored according to the AHC's existing procedures for accessing content, i.e. Ferret or the Colorado Digitization Program's Heritage database of Dublin Core records.
- Certain core data should be captured for every object digitized. This data is expressed in terms of unqualified Dublin Core elements, but if another standard is chosen (such as EAD), their equivalents should be used.

See the Western States Dublin Core Metadata Elements for further information <http://www.cdheritage.org/westertrails/wt_bpmetadata.html>.

LONG-TERM STORAGE

Storage Media and Location

Online storage (in file systems) is preferred and the AHC has approximately 240 GB of storage available on the server shared with the UW Libraries, named Minirva. CD-ROM can be used for short-term storage of images if necessary.

Audio CD, DVD Video and DVD Audio should be used only in limited situations. In these cases, the data is stored in a format that will require special equipment for access. While this equipment is widely available, and expected to be widely available for some time, it adds a layer to the data access process, and should be avoided if at all possible. Note that the CD/DVD media have a limited physical life, so these storage procedures must be reviewed periodically for quality checks.

The following standard file types for different types of digital objects will be stored on the server:

- GIF (Graphics Interchange Format)
- JPEG (Joint Photographic Experts Group)
- TIFF (Tag Image File Format)
- PNG (Portable Networks Graphic)
- MPEG (Motion Pictures Experts Group)

File Naming Conventions

Systematic file naming is important for system compatibility, interoperability, and to demonstrate ownership of the digital asset. General practice indicates using a convention with an eight-character file name with a three-character extension to accommodate different systems; the characters are alpha-numeric, lowercase, and do not utilize spaces, tabs, or any characters reserved for system use (i.e. \ / ? * |, etc.).

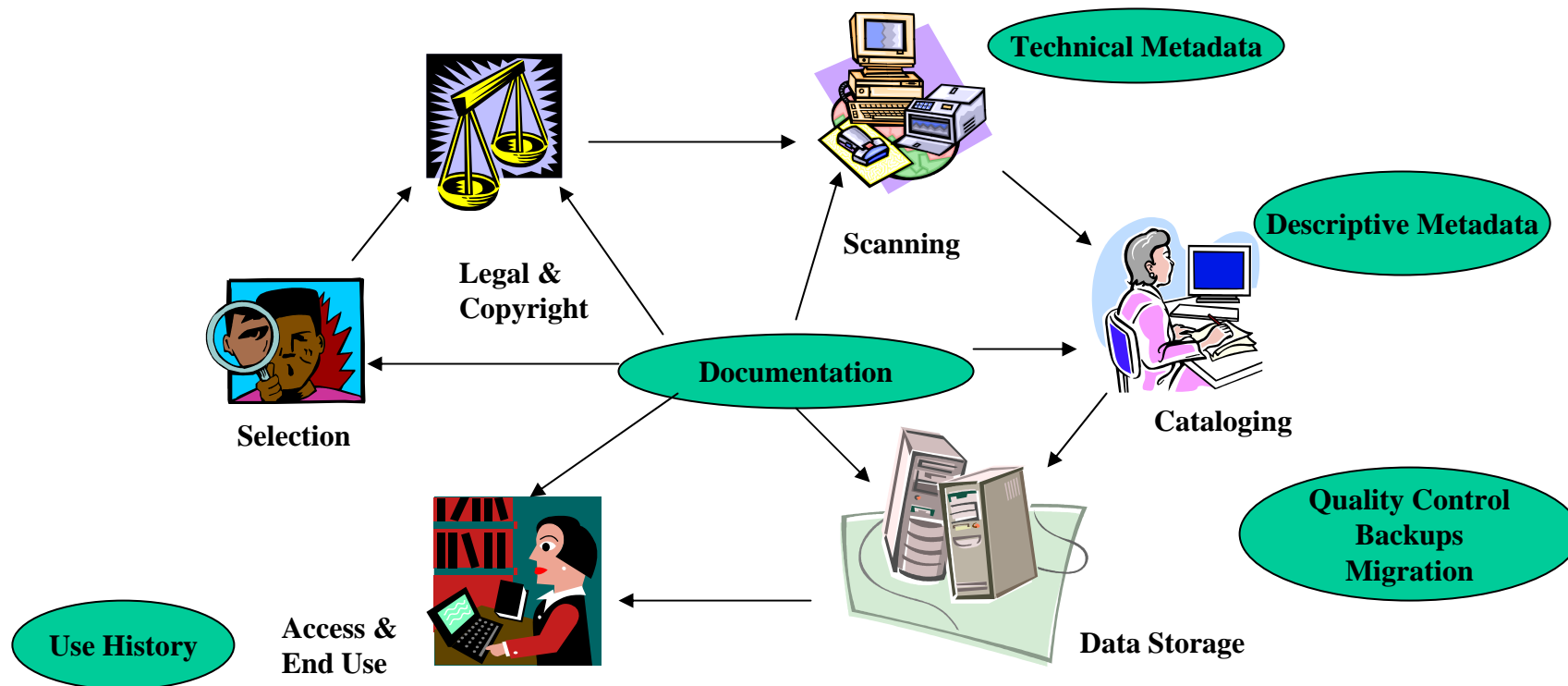
The first two characters will be an alphabetical unit-specific identification and the remaining characters a numeric digital object identifier.

Example: ah000421.tif (American Heritage Center, digital object number 421, TIFF).

Refreshment, Migration

Refreshment of media, servers, and file formats is crucial for ongoing sustainability of the digital items. Refreshment is the transfer of digital files from one storage media to another to ensure that the files remain retrievable as technology advances. Migration is the process of moving from the use of one operating environment to another operating environment and can involve hardware, software, and file formats. It is recommended that the AHC develop a plan for the migration and refreshment of the data and its implementation.

Digital Task Mapping





American Heritage Center

Digitization Checklist

Approved 06 June 2003

Note: This document outlines the production process for historical collections at the AHC and not every collection requires all of the steps listed; some collections may require other steps that are not listed. In practice, many of the operations are carried out in parallel and not sequentially.

It is strongly recommended that collections be digitized after the collection has been processed and a MARC record has already been created.

I. Select a collection for digital conversion

A. Name of initiator/requester (if any):

B. Brief description of item(s) to be digitized, include Collection Name, Accession Number, Box and Folder Location:

C. Analyze Collection

1. Determine scope or extent of digitization (entire or subset?)
2. Assess the status and quality of access (finding aid, catalog record)
3. Assess best format (full text conversion, scanned page images)
4. Assess the physical condition and readiness for scanning
5. Assess restrictions and copyright (if any)

II. Plan the Approach to Digitization

A. Develop method and resource plans for collection preparation & digitization

1. Develop plan for required processing (if any)
2. Develop preservation treatment plan (if any)
3. Complete evaluation of physical condition with recommendations
4. Determine formats for capture, archiving and presentation
5. Determine physical size (number of characters, images) and special production requirements

B. Determine repository requirements

1. Verify file naming sequence
2. Estimate required storage space
3. Develop restriction plan & implementation (if any, note at the item level)

III. Produce Digital Collection

A. Image Capture

1. Preparation

- a. Prepare targets
- b. Prepare scanning instructions specific to collection

2. Capture

- a. Scan item, items or collection
- b. Review images for quality
- c. Coordinate rework

3. Archive images on server in directories as specified by naming scheme

B. Text Capture

1. Prepare keying instructions specific to collection or batch
2. Mark up and key text
3. Review completed text for quality
4. Coordinate rework
5. Archive files on server in directories as specified by naming scheme

C. Audio and Video Capture

1. Create preservation and working copy
2. Determine sample rate and resolution
3. Select digital audio file format
4. Analyze storage requirements
5. Acquire temporary storage space for conversion
6. Perform analog to digital conversion
7. Edit digital files removing "dead air" at cue-up points
8. Perform quality review
 - a. Inspect graphic waveforms for truncation and peak
 - b. Audition percentage of sound files
9. Archive files on server in directories as specified by naming scheme

IV. Create Access

A. Complete Metadata Record using Metadata Worksheet

Digitization Committee

Workflow – Scan Operator



1. Check the form on the J drive for scanning requests.
2. Pull materials to be scanned from the Reference holding shelves or the scanning cart.
3. Scan materials according to AHC Digitization Guidelines to create master image with resolution and file format appropriate to the material type. Save the files according to the following protocol:
 - Use an eight-character file name with a three-character extension. The first two characters of the file name will be ah (always in lowercase) for the AHC, followed by a six digit sequential number, i.e. ah000001.tif
4. Test the first 10 images of each batch to ensure quality. Following the quality check, 10% of the remaining images will be checked. Five attributes need to be verified:
 - A. Color
 - B. Tone
 - C. Detail of reproduction
 - D. Noise
 - E. Compression effects
5. Use Photoshop to create access and thumbnail copies from the master image in batches.
6. On the scanning workstation pc, under the My Documents directory is a sub-directory labeled ah, with sub-sub-directories for master, access, and thumb. Place each appropriate image in these sub-sub-directories. The scanning workstation is only a temporary holding location for scanned images; please monitor the storage capacity of the pc's hard drive.
7. Complete a Dublin Core record for each item scanned.
8. Move the files to the server, named NAS2 for storage. Under the main AHC folder there will be a folder for the master, access, and thumb copies and place each file in its respective folder.
9. Complete a log of the date, time and numbers items scanned, and any comments each time scanning is performed.
10. Return materials back to the Reference holding shelves, the scanning cart, or any photograph orders to the Photo Archivist.

**Digitization Committee
Proposal Form – Projects
August 25, 2003**



Statement of Purpose

The AHC Digital Committee wishes to solicit digitization proposals in order to further expand online access to AHC Collections. Departments and individuals are encouraged to describe, in detail, projects they feel best reflect the quality, diversity and breadth of AHC collections. The Digital Committee will evaluate proposals and establish necessary priorities within the context of available staffing, resources, and technology.

Setting Priorities

Priorities among all the proposals will be based in part on the following assessment criteria:

- **Project Description** – To what degree does this project address attempts to make available collections of distinction from the AHC?
- **Audience** – Does this project have the potential to reach out to a large or specially targeted audience?
- **Project Impact** – What is the size of the project relative to its potential impact?
- **Previous Projects** – Have similar or related projects been completed within or outside of the AHC?
- **Outside Funding** – Does this project have funding beyond the AHC?

Please complete the following information for the committee to evaluate the feasibility of the proposal and attach a short proposal summary, no more than one page in length. If necessary, please attach additional supporting information or documentation.

1. Title of project:

2. Desired outcome of digitization:

3. Scope of project:
 - a. Number of items:
 - b. Material type:
 - Photographs or negatives:
 - Textual materials:
 - Audio (please indicate format)
 - Video (please indicate format)
 - Other:

4. Legal and copyright status:

5. Existing access:
- a. Unprocessed:
 - b. Finding aid (please indicate if an EAD record is available):
 - c. Catalog record:
 - d. Does any descriptive metadata already exist and in what format?

6. Please indicate the collection's physical condition and any concerns:
- a. Fragile:
 - b. Good:
 - c. Excellent:
 - d. Other (please elaborate):

7. Timeframe:

8. Scholarly value/Audience:

9. How will digitization enhance the collection's value?

10. Relationship to other digital resources at the AHC or elsewhere:

11. What funding sources have been examined or pursued:

12. What are your plans for promoting the project?

- a. News release:
- b. Email announcement:
- c. Exhibit:
- d. Lesson plans:
- e. Other (please elaborate):

13. Will items continue to be added in an ongoing manner?

**Digitization Committee
 Proposal Evaluation Form
 August 25, 2003**



This form has been developed to help determine the feasibility of digitization proposals. The questions take into consideration organizational mission, informational value, audience, and resources.

1. The project supports the AHC's mission.
2. The project is within the scope of the AHC's resource capabilities for digitizing, creating metadata, migration, and storage.
3. Ownership and copyright of the collection have been verified. If the AHC does not have legal authority to make the records available or if the collection is not in the public domain, these permissions can be secured.
4. The project can be completed within the specified timeline.
5. The current physical condition of the collection does not prohibit digitization.
6. Digitization presents benefit in addition to more traditional means of access (i.e. a catalog entry or a finding aid).
7. There is a known demand for the information, or the intrinsic value of the material will ensure interest in a digital product.
8. The project has investigated collaborations or connections with other activities or materials at the AHC or elsewhere.
9. The project has investigated possible external funding sources.
10. The project incorporates user feedback.

Strongly Disagree						Strongly Agree
	0	1	2	3	4	5
	0	1	2	3	4	5
	0	1	2	3	4	5
	0	1	2	3	4	5
	0	1	2	3	4	5
	0	1	2	3	4	5
	0	1	2	3	4	5
	0	1	2	3	4	5
	/50					

TOTAL:
Approved: 42-50 points
Need additional information/discussion: 34-41
Rejected: 0-33

COMMENTS: