Western States Digital Standards Group
Digital Imaging Working Group

Western States Digital Imaging Best Practices

Version 1.0

January 2003

Funded by an IMLS grant to University of Denver and the Colorado Digitization Program

Created on: 2002-07-17
Last Modified: 2003-05-22

Acknowledgements

In the Fall of 2001, the University of Denver was awarded an Institute for Museum and Library Services (IMLS) grant to develop a multi-state collaborative initiative that demonstrates the ability to broaden access to a collection of widely dispersed digital resources by creating a virtual collection of digital resources using the topic Western Trails. Twenty-three institutions in four Western states were awarded mini-grants to create digital content and associated metadata. Additionally, the grant was to develop a set of Best Practices for Digital Imaging, with involvement of representatives from cultural heritage institutions in the four participating states. In June 2002 6 representatives from Western States met in Denver, Colorado to begin the task of reviewing and updating the existing Colorado Digitization Program best practice guidelines.

I would like to thank the following individuals for their contribution to the creation of this document: Steve Boss, University of Wyoming, Clee Library; Jill Koelling, Nebraska State Historical Society; Jennifer S. Kutzik, Colorado State University; Roel Muñoz, Denver Public Library; Mark Shelstad, American Heritage Center, University of Wyoming; Sarah Goodwin Thiel, Spencer Research Library, University of Kansas; Richard Urban, Colorado Digitization Program; Brian Pytlik Zillig, Electronic Text Center, University of Nebraska - Lincoln Libraries

Liz Bishoff
November 2002

Purpose

The purpose of this document is to offer guidance and to provide ***minimum*** digital imaging recommendations to institutions that are planning for or are involved in digitization projects. These guidelines are not intended to be used as the de facto standard for digital imaging, but rather as a guide for image capture, presentation, storage, and preservation. Inherent or unique characteristics of different source materials necessitate different approaches to scanning and conversion requirements for digital projects should be considered on a case-by-case basis (particularly for grant projects with specific requirements). The recommendations we make in this document are purposely broad enough to apply to a variety of institutions and collections and attempt to synthesize different recommendations previously made for specific institutions or projects.

This document addresses the more standard formats of text, photographs, maps, and graphic materials and is written for institutions that have the equipment and expertise to scan in-house. If you are planning to primarily scan oversize materials, bound materials, or materials in non-standard formats and sizes, you may consider outsourcing these materials to imaging vendors.

These guidelines have been developed in order to:

1. increase the interoperability and accessibility of digital collections across the cultural heritage community through the use of widely accepted standards and formats
2. ensure a consistent, high level of image quality across collections
3. decrease the likelihood of rescanning in the future by promoting best practices for conversion of materials into digital format and the long-term preservation of these digital resources.

Because technology and industry standards are constantly improving and changing, we view this as a continually evolving document. We welcome your comments and suggestions.

Scope

What is addressed in this document:

- Scanning and file format recommendations for:
- Text, photographs, maps, and graphic materials
- Suggested hardware configurations
- Software considerations
- Quality control, file naming, scanner and monitor calibration, targets and color bars, storing images, and recording and verification of CD-ROMs.

What is not addressed in this document:

- Scanning and file format recommendations for:
    - Audio, Video/Moving Images, 3-D Objects, Born-digital items.
      These are available in other CDP documentation and through other projects specifically dealing with these types of material.
- Workflow issues
- Metadata and access standards
- Selection of collections for digitization
- Systems and network architecture

Revisions

The Western States Digital Imaging Workgroup will review and update the Digital Imaging Best Practices on an annual basis. The workgroup will also supplement the Digital Imaging Best Practices with individual case studies that expand on basic requirements included in this document. The Colorado Digitization Program will serve as the host institution for the workgroup's efforts and all documents and related materials are available through the CDP website at http://www.cdpheritage.org.

Pixels Per Inch (PPI) vs. Dots Per Inch (DPI)

Digitization guidelines and hardware manufacturer's documentation frequently use the measurement of "Dots Per Inch" or DPI when discussing optical resolutions for images and hardware. As Dots Per Inch (DPI) more accurately refers to output devices (how many dots of ink per inch does a printer put on the paper), the Digital Imaging Workgroup decided to standardize on the technically correct Pixels Per Inch (PPI) when discussing scanning resolution and on-screen display. For more information on digital measurement units see the Digitization Glossary.

General Principles

1. Scan at the highest resolution appropriate to the nature of the source material.
2. Scan at an appropriate level of quality to avoid rescanning and re-handling of the originals in the future--**scan once**
3. Create and store a master image file that can be used to produce derivative image files and serve a variety of current and future user needs
4. Use system components that are non-proprietary
5. Use image file formats and compression techniques that conform to standards within the cultural heritage community
6. Create backup copies of all files on a stable medium
7. Create meaningful metadata for image files or collections
8. Store media in an appropriate environment
9. Monitor and recopy data as necessary
10. Document a migration strategy for transferring data across generations of technology
11. Anticipate and plan for future technological developments
12. Scan an original or first generation (i.e., negative rather than print) of the source material to achieve the best quality image possible.

Related Documents

The following documents available at the Colorado Digitization Program website may be helpful to review along with the Digital Imaging Best Practices:

- Questions to Ask Before You Begin a Digitization Project
- Digitization Glossary
- Legal Issues to Consider When Digitizing Collections
- Western States Digital Imaging – Case Studies
- Western States Dublin Core Metadata Element Set Best Practices

Project Planning

Successful projects should include careful planning before implementing a digitization initiative. This planning should consider how digitization fits into the institution's overall strategic plan, technology plan, and project workflows. For more information on project planning see the Project Planning, Strategic Planning, and workflow resources on the Colorado Digitization Program web site.

Documentation

Documentation of the choices your project has made can be a key factor in the long-term success of digitization efforts. Good documentation can offset the impact of staff turnover and allow future staff an ability to deal with digital collections created by their predecessors. Among the items to consider documenting:

- Critical assumptions for decisions (funding, costs, staffing, available technology and skills, intellectual property concerns)
- Local guidelines and benchmarks for image quality and resolution
- Resources that contributed to local practice guidelines
- Types of metadata captured (see Metadata below)
- File naming schemes
- Sustainability plans and procedures (storage, archiving, refreshing media, etc.)

Staffing

In practice, many digital imaging projects will not have unique staff working on the project, but will utilize existing staff from other areas in the organization, student assistants, or volunteers. It may benefit the project coordinator(s) to look at "transferable skills" that project staff members already possess that would be useful in any digitization project. Sufficient time for training, and opportunities to receive further education and training, should also be provided. Cataloging and digitizing are labor intensive and may require a skill level above that of minimum wage employees. If your current staff does not have the necessary skills, training should be included in the planning.

Digitization projects require a combination of skills from a variety of staff with different areas of expertise. The following areas and skills may be important to any digitization project:

- Project management skills
- Knowledge of cataloging, registration methods, or metadata schema.
- Familiarity with conservation methods
- Understanding of photographic techniques and methods
- Subject matter specialists (curators, archivists, scholars, librarians, faculty, etc.)

- Database development and administration skills
- Computer programming skills
- Web design and development skills
- Artistic/graphic design skills

By nature, digitization projects require a team approach, and bring together diverse sets of skills from different areas of the organization, perhaps more than any other project. Administration, technical services staff, cataloging specialists, the information technology department, subject specialists, curators, librarians, preservation/conservation staff, faculty, and others may all be involved.[1]

Sample project staff and their roles:

- Project manager
- Selector, Conservator, Preparations technician
- Cataloger, Metadata analyst
- Scanning technician or Photographer, Quality control technician
- Programmer or other Database developer to integrate metadata and images
- Systems administrator, Network administrator
- User Interface Developer or Designer

Training

Many organizations around the country offer workshops and training on digital imaging, and many conferences are held each year addressing imaging issues.

- The Colorado Digitization Program offers workshops and seminars in digital imaging and metadata creation.
- The Northeast Document Conservation Center hosts a "School for Scanning" several times a year for digital project managers. Information on the school can be found on their website.
- AMIGOS Bibliographic Council gives several workshops a year on digital imaging to institutions in the Southwest and at conferences.
- Online publications such as RLG Diginews and D-Lib list conferences on digitization around the country.
- Cornell University has developed an online training tutorial
- Several professional organizations often host workshops and conferences on issues related to digitization:
    - American Library Association (ALA),
    - Library and Information Technology Association (LITA) ,
    - Association for Library Collections and Technical Services (ALCTS),
    - Society of American Archivists (SAA),

- [American Association of Museums (AAM)](),
- [Museum Computer Network (MCN)]()
- [Archives & Museum Informatics (A&MI)]()
- [CIMI Consortium]()
- [National Initiative for a Networked Cultural Heritage (NINCH)]()
- [Institute for Museum and Library Services]()

▪ Non-profit and educational institutions can also qualify for discounted training through corporate training firms. These firms can provide basic instruction in common image editing and web authoring software.

In-house or outsource?

Every organization should carefully consider the pros and cons of outsourcing digitization projects or conducting them in-house. Following are some points to consider for both strategies[2]:

In-house pros:

- Development of digital imaging project experience by "doing it" (project management, familiarity with technology, etc.)
- More control over the entire imaging process as well as handling and storage of originals
- Requirements for image quality, access, and scanning can be adjusted as you go instead of defined up front
- Direct participation in development of image collections that best suit your organization and users

In-house cons:

- Requires large initial and ongoing financial investment in equipment, staff
- Longer time needed to implement imaging process and technical infrastructure
- Limited production level
- Staffing expertise not always available
- Institution must accept costs for network downtime, equipment failure, training of staff, etc.
- Need to enforce standards and best practices

Outsourcing pros:

- Pay for cost of scanning the image only, not equipment or staffing
- High production levels
- On-site expertise
- Less risk
- Vendor absorbs costs of technology obsolescence, failure, downtime, etc.

Outsourcing cons:

- Organization has less control over imaging process, quality control
- Complex contractual process: image specifications must be clearly defined up front, solutions to problems must be negotiated, communication must be open, and problems must be accommodated
- Vendor many know more than client or may presume a level of understanding on part of library/museum/archives that they may not have
- Lack of standards with which to negotiate services and to measure quality against
- Originals must be transported, shipped, and then also handled by vendor staff
- Possible inexperience of vendor with library/archival/museum/historical society communities

Costs

It is difficult to predict just how much a digital imaging project is actually going to cost, and little hard data on the cost, cost effectiveness, and costs over time of digital projects is readily available. Generally, capture and conversion of data often comprises only 1/3 of the total costs, while cataloging, description, and indexing comprise 2/3 of the total costs. Upfront and ongoing costs can be significant, and economic advantage may be better realized through collaborative initiatives or cooperative/regional digitization initiatives, where costs, resources, goals, and expertise can be shared. Initial investment in equipment, staff training, capture and conversion, handling, storing, and housing originals, producing derivative files, CD production, cataloging and building the image database system, and developing Web interfaces are all possible areas of cost for any digitization project. However, the costs of a project do not end after conversion. Some on-going costs that an institution must commit to include the costs of maintaining data and systems over time, including media migration costs and infrastructure costs.

Digitization Toolbox

Hardware

Computers

Critical to the success of any digitization initiative is the purchase of a computer with a balance of reliable components, speed and storage that will increase productivity and overall effectiveness. Projects planning the purchase of computer hardware should consider the following general principles:

- Purchase a computer dedicated solely to digitization initiatives
- Purchase as much Random Access Memory (RAM) as your budget allows. More memory allows the computer to more quickly process large amounts of image data.
- Purchase computers with processors optimized for image manipulation (Pentium III and higher processors included special instructions for processing Photoshop functions.)
- Purchase computers that support high-speed data input through serial connections, USB 2.0, or IEEE 1394 "Firewire"
- ISO 9660 compliant CD-RW burner[3]

For the selection of specific equipment it is recommended to review trade publications, such as *PC Magazine* and *Macworld*. Publications and review websites geared towards graphic arts professionals can frequently offer additional information relative to cultural heritage digitization projects.

Display

The investment in a large display monitor (19" – 21" viewable) will also increase productivity of your project by providing more "screen real estate" to view and evaluate images. At this writing cathode ray tube (CRT) monitors still held a cost advantage over liquid crystal display (LCD) monitors.

For institutions considering technical color calibration (see Color Calibration) high quality monitors with standardized hardware and software calibration that are compatible with input devices (scanners, cameras, etc.) should be strongly considered.

Each type, size, and quality of monitor can interpret and display colors and tonal values differently. Without careful and frequent color calibration (see Color Calibration) projects should exercise caution when adjusting and manipulating images since you may be introducing color and tonal biases of your equipment into your images. Even with careful calibration there is no guarantee that images will appear the same on other computers, particularly when delivered across the Internet to a wide variety of end-users.

Scanners

The purchase of a scanner will have the greatest impact on quality of images for the majority of digitization projects. Recent developments have increased the challenges in selecting a quality scanner by increasing variety and availability while reducing the costs of equipment. What scanner is right for your project depends on numerous factors including overall project goals, format, size, and condition of materials to be scanned and available budget. Several technical factors will also influence your purchase including available optical resolution, bit depth, size of scan area, speed, connectivity, and ability to handle different formats and materials in your collection.

Optical Resolution

Most scanners use a grid-like array of light sensors that translate light into the 1s and 0s of your digital image. The number of sensors in the array determines the ***optical resolution*** of a particular device. The optical resolution is normally expressed in scanner specifications as “dots per inch” (DPI) or “pixels per inch” (PPI). The optical resolution of any equipment you purchase should exceed the maximum resolution needed to accurately capture the types of material in your collections (see Guidelines by Source Type). For example, flatbed scanner with an optical resolution of 1200 dpi has sufficient optical resolution to scan an 8x10” print at 600 dpi, but insufficient optical resolution to scan a 2x2 slide at 2000 dpi.

Many models of scanners are advertised with very high resolutions that represent the ***interpolated resolution.*** To increase the resolution the equipment uses a mathematical algorithm to “guess” what color and light values exist in the spaces that the light sensors can’t see. Make sure to select equipment based on its ***optical resolution*** and not the ***interpolated resolution*** since scanners with adequate optical resolution will produce more accurate scans.

Scanner specifications often include the size of the array (1600 x 3200). The first value measures the optical resolution of the array and the second value represents the capacity of the array to capture information as it moves across the scan area (how many pixels does the array move before taking another sample). If the second number is smaller than the first number the samples are interpolated. For most professional quality scanners the second value will be higher than the optical resolution.

Point and shoot digital cameras frequently measure their optical resolution in “megapixels” or the total number of pixels captured by the array. (A digital camera that captures an image measuring 2048 pixels by 1536 pixels captures a total of 3.1 megapixels (2048 x 1536 =

319,488 pixels) . At this writing, point-and-shoot digital cameras do not offer sufficient optical resolution for capturing archival quality master images, whereas digital scan-back cameras (offering arrays up to 8,000 x 10,000 pixels) can produce images at recommended resolutions. For more information see Digital Cameras below.

Bit Depth / Color Depth

A "bit" is the most elemental value in the digital world that measures either the presence or absence of an electrical charge – the 1s (on) and 0s (off) that make up digital information. When applied to digital imaging "bits" are the basic pieces of information sent from the sensors in an array to be assembled and recorded as a digital image.

The amount of information that a sensor in an array can capture is represented by the "bit depth" The greater the bit depth the more information about the source is captured by the array, resulting in a more accurate digital representation of the original. A bit depth of 8 can capture enough information to represent 256 colors or shades of grey. A bit depth of 24 captures over 16 million colors or shades of grey.

Since higher bit depths capture more information they impact the overall file size of an image. Scanners generally sample at a higher bit depth and sample down to a lower bit depth for final output. Sampling at a higher bit depth aids in reducing noise, extends the possible tonal range of the image, and allows the scanner to capture a larger density range without loss of detail. Many scanner models available at this writing have bit depths between 36 and 48 bits and output a 24 bit image.

As with resolution any equipment that you purchase should meet or exceed the bit depth required by the types of materials being digitized. (see Guidelines by Source Type). When comparing specifications note that some scanner specifications divide the total bit depth into the color channels (Red, Green, Blue) used in digital capture and display (8 bits Red + 8 bits Green + 8 bits Blue = 24 bits).

Optical Density

The optical density measures the "brightest bright" and the "darkest dark" that a piece of equipment can capture. A scanner or camera's optical density will impact the overall quality of the image's tonal dynamic range (see below) and ability to capture highlights and shadows in an image. Scanners that are able to capture higher bit depths generally are able to capture higher optical densities.

Newer equipment specifications include the measurement for the "darkest dark" or maximum density (dMax). Equipment with a higher dMax is able to capture deeper shadows. A higher dMax is particularly critical for quality capture of transparencies, negatives, and slides that require an external light source to be captured.

Density Range of Scanning Equipment[4]	
Equipment	**Maximum Density range**
24 bit color flatbed scanners	2.2 – 2.6
30 bit color flatbed scanners	2.8 – 3.2
34 – 36 bit color flatbed scanners	3.3 – 3.6
Desktop drum scanners	3.3 – 3.7
High-end drum scanners	3.4- 3.8
Film/Transparency scanners	2.2 – 4.2

Density Range of source materials[5]	
Source	**Density range**
Newsprint	.09
Printed material	1.5
Coated stock	1.5 – 1.9
Normal Photographic prints (C-Type)	1.6 - 2.0
High contrast photographic prints (R-type cibachrome)	2.0 - 2.3
Negative film	2.8
35mm slides	2.8 – 3.0
Transparencies	3.0 – 4.0

Speed & Connectivity

An important factor to consider for the most efficient projects is the speed that equipment can capture images and transmit them to the host computer. Most scanners include specifications on scanning speed. Higher scan speeds and transfer rates will reduce the exposure of sensitive materials to light, particularly when using digital scan-backs and cameras on a copy stand.

To ensure efficient transfer of image data select equipment (both scanner and computers) that use high speed data transfer standards, such as Universal Serial Bus (USB) 2.0, Small Computer Serial Interface (SCSI) cards & cables, or IEEE 1394 "Firewire." Avoid equipment that uses slower methods such as parallel ports, or USB 1.0.

Types of Scanners

Flatbed Scanners

Flatbed scanners are one of the most popular types scanners used in libraries and archives and are suitable for scanning papers, flat photographs, and other printed materials. An important consideration when selecting a flatbed scanner is the size of the scan area. Most consumer models are limited to a scan area of 8.5” x 11” but professional grade models are available with a larger scan area. Some models of flatbed scanners are available with accessories such as transparency adapters (for slides and negatives) and automatic document feeders. The quality of scans from transparency adapters varies greatly and projects with large numbers of slides or negatives should consider a dedicated slide and film scanner. Automatic feeders are not recommended for original materials because of the danger of damage but can increase efficiency of scanning contemporary documents such as archival finding aids.

Slide Scanners

Slide and film scanners are specifically designed to digitize transparent materials such as slides and 35mm film. Transparency scanners generally produce a higher quality scans over flatbed transparency adapters due to higher dynamic tonal ranges and optical resolutions. Projects with a large number of slides should consider the advantages of slide scanners and accessories such as automatic slide feeders.

Drum Scanners

Drum scanners are most frequently used by pre-press and graphic design professionals working with contemporary materials. Because materials are affixed to a drum rotating at high speed around a sensor, drum scanners are not recommended for cultural heritage materials, particularly materials that are fragile or brittle. Drum scanners do produce high resolution scans with high color fidelity and dynamic ranges and are suitable for scanning surrogate negatives and transparencies. The cost of drum scanners will be a limiting factor for most projects.

Wide-format Scanners

Wide-format scanners were developed to digitize large format materials such as engineering drawings and architectural blueprints and are frequently found in municipal engineering departments or local blueprint shops. Materials are drawn over the scanning sensor through a pair of drums. Due to the danger of mechanical damage (ripping, tearing) these types of scanners are not recommended for cultural heritage materials.

Digital cameras

Institutions with large format materials such as maps, posters, 3-dimensional objects and artwork may need to consider digital cameras for capture of these materials. At the time of this writing most consumer and professional digital cameras do not have sufficient resolution for archival capture of cultural heritage materials. Lens used in these types of cameras are designed for capturing three-dimensional scenes and may introduce distortions to flat materials.

For the creation of high resolution, high quality digital images using a camera many cultural heritage institutions employ a "digital scan back" camera. A digital scan back consists of a scanning array that attaches in place of a film holder to a 4" x 5" view camera body. Digital scan back cameras are ideal for original items that cannot be put onto a more traditional scanner—three dimensional objects, photos or artwork bound in albums, large photographs, artwork or maps. Projects considering digital scan-backs should consult with manufacturers to select the correct lenses and lighting needed for the types of material being scanned.

At present, the cost of digital scan backs are beyond modest sized projects and best practice has been to scan negatives of oversize materials created using traditional photography.

Software

Scanner Software

The last link between your hardware and your computer hardware is the software that controls the scanner or camera and passes information to computer storage or image editing software. Higher end scanners normally come bundled with software that allows the operator to manually adjust resolution, tonal dynamic range, and color channel values. Consumer model scanners frequently include pre-set software that does not allow careful adjustment of these values and can result in poor quality images.

Scanner software should also be able to output image files in the file format appropriate to the materials being digitized (see File Formats below), lower end scanner software frequently limits file format choices and may not produce files adequate for high quality master images.

Frequently scanner software offers additional features for image manipulation at the time of scanning. Best practice is to carefully compare the results of these processes with those of your image editing software to select which offers better quality. Whether or not to apply these processes (such as descreening, color correction, etc.) also depends on the goals of the project. Projects interested in capturing accurate representations of an object's current condition may not wish to apply certain processes, whereas projects that wish to capture the artist's original intent may make adjustments so that the digital file more accurately reflects the original.

Image Editing Software

The function of scanner drivers and plug-ins offers a limited array of features for the manipulation of images. Projects should consider acquiring professional image editing software for the creation of surrogates for delivery via the web, print publications, or for in-house uses such as exhibits.

When selecting image editing software projects should consider the following features:

- Ability to work directly with scanner software through TWAIN or other plug-ins
- Support for common non-proprietary file formats (see File Formats below)
- Tools for controllable image optimization (color adjustment, tonal adjustments, color spaces)
- Features for the optimization of images for web delivery and automatic creation of HTML templates
- Ability to convert color spaces (RGB to CMYK for print output)
- Usable documentation and reliable technical support
- Ability to extend functionality through custom plug-ins
- Ability to create action sets or macros for frequently applied functions.
- Ability to process images in automatic batches
- Other project specific needs and goals

Projects should consider the costs of implementing the software beyond its initial costs. Does your computer hardware exceed the minimum requirements of the software? Do I have the staff with skills to use the software or the funds to provide training? Can I afford future upgrades to the software? Does the software feature automated processes that can increase efficiency and reduce staffing costs?

Digital Asset Management

Projects are increasingly recognizing the need for software to manage the large number of digital files created. A number of vendors provide out-of-the-box solutions for the creation of image metadata, surrogate file creation, workflow management, intellectual property and rights management. Currently software is available for both stand-alone and network management of digital files or as add-on modules to existing library and collections management software.

Projects considering digital asset management software should ask these questions:

- How does it fit with the overall project goals and workflows?
- Is it scaleable?
- Does my current technology infrastructure support the software?
- Does the system work with standardized file formats and metadata schemes?
- Startup, maintenance, and staffing costs?
- Is it possible to easily extract and migrate metadata and files?
- Does it support multiple users and offer user security levels

Workspace

Providing a comfortable, safe, and secure workspace for a digitization project can increase productivity and quality of images by reducing operator fatigue and potential damage to collections. Proper climate control and security are important if collections will stay in the lab for extended periods of time A workspace for digitization should offer a controlled lighting source to maintain consistency and quality of images. Changes in room lighting can affect how images are represented on computer monitors and may introduce challenges to accurate calibration.

Projects considering outsourcing digitization are recommended to visit vendor facilities to ensure that the workspaces provided meet the criteria above.

Storage and Preservation of Digital Images

The significant resources institutions devote to the creation of digital collections has increased awareness of the need for careful planning for the storage and long-term preservation. Successful digitization projects should include planning and documentation for the responsible sustainability of these collections.

The Western States Digital Standards group plans on developing additional best practices for the long-term preservation of digital assets in the coming year. Please visit the Colorado Digitization Program website for additional information.

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 or three characters can be an alphabetical unit-specific identification and the remaining characters a numeric digital object identifier. For example: am000421.tif (Art Museum, digital object number 421, TIFF).

Other naming conventions may also employ a protocol that includes institutional acronym, collection identifier, part designator and file extension, separated by an underscore. Such as the Laramie Historical Society, Accession Number 9800, Box 5, Folder 2, item 26. For example, lhs_9800_b5f2i26.tif.

Other protocols may include the collection number and folder number as part of the path name, i.e. lhs/9800/b5/f2/i26.tif. In this protocol any derivative copies (i.e. jpeg, gif) could be stored in the same directory if desired.

Optical Media Storage

Optical media include CD-ROM (Compact Disc – Read Only Memory), CD-R (Compact Disc – Recordable), and DVD-ROM (Digital Versatile Disc – Read Only Memory). All employ a laser to read data from a metallic coating over the disc, with a clear acrylic coating covering the metallic layer for protection. Optical discs are a popular medium for storing digital masters, transporting images, and as a publication medium.

Issues of cost, convenience, speed of retrieval and security factor into decisions regarding optical discs as an image storage and retrieval medium. It is recommended that the digital assets be stored on CD-ROMs that conform to ISO 9660, the 1988 standard for volume and file structure of CD-ROMs for information interchange. Audio CD, and DVD Video and Audio are not recommended at this time as a reliable storage medium.

Some institutions may wish to use optical media for storage media due to cost considerations. When acquired in bulk,650 MB CD-R discs can cost as little as 35 cents each, or .0005 cents per megabyte of storage. For example, an uncompressed 32 MB TIFF file (800 x 600 pixel black and white textual documents scanned at 600 dpi, 8 bit grayscale) will therefore cost approximately 1.6 cents for storage. Thus, approximately 20 32 MB TIFF files can be accommodated on the disc.

Optical discs can provide cost-efficient storage, but they incur staffing costs when accessing and managing the digital assets stored on them. The main drawbacks are slow recording and reading speeds, for even with a fast CD burner, creating CDs is a slow process, and a CD player needs to be available for access. While the CD burner and players are widely available, having to locate the disc on which the image is stored, loading the disc, and then locating the image adds a layer to the data access process.

Digital collections stored on CD ROM work well for small collections. Using this media as a long-term solution however presents major challenges. Storage and retrieval costs escalate as the CD ROM collections become larger and more challenging to manage.

CD-ROMs have a limited physical life span and the images stored on them are vulnerable due to physical deterioration, mishandling, improper storage and obsolescence. Studies have indicated that the physical lifetimes of the media to be wide ranging – anywhere from 5 to 100 years; but obsolete equipment for reading the media poses an equal if not greater threat. Proper conditions can prolong the life of the discs, and optimal environmental levels are 72° Fahrenheit and a Relative Humidity between 20 and 50 percent.

Both adhesive labels and permanent ink markers can cause early failure of CD-ROMs through chemical interaction with the CD's foil. Best practice is to not write or label CDs directly on the body of the CD. Some projects have placed small identification numbers in the central plastic "hub" of the CD.

Projects considering CD ROM for storage of master images should pay attention to the burning speed of the media they purchase and the maximum speed of the CD burner. Like the hardware used to create them, CD-Rs have a maximum burn speed that should not be exceeded even if the hardware is capable of higher speeds.

Projects should also use the verification utilities in most CD burning software to ensure that there are no errors in a disk before removing master images from temporary storage. Burning multiple copies of from master images (rather than copying a previously burned CD) also decreases the possibility of errors or disk failure.

Online Storage

Storing scanned images on "live" servers is an option for those wishing to archive high-resolution TIFF images on a stable platform that offers sustainability and easy storage and retrieval. To prevent the loss of data projects need to properly configure hardware and software, develop responsible backup and disaster recovery policies and procedures, and create realistic plans to deal with technological obsolescence.

RAID Arrays

A Redundant Array of Inexpensive (or Independent) Disks (RAID) is a collection of disk drives that, if configured properly, can act as a single storage system. These configurations are designed to enable a system to operate when an individual drive fails and prevent the loss of data. Currently, there are about 10 types of RAID configurations. Each configuration has its own unique strong and weak points. Some configurations are best suited for rebuild speed while others have are designed to maximize disk capacity and others are well suited for fault tolerance.

Network Attached Storage (NAS)

Network attached storage devices are servers that are optimized for file sharing rather than running applications. A NAS can offer flexible storage solution to institutions that already have a local area network and server for user authentication. As storage needs increase additional NAS devices can be added to your network for access and storage of digital images. NAS devices used for digitization initiatives should use a RAID configuration to protect against data corruption or loss.

Outsourced Storage

Projects who have created images for local and web access may wish to place their large, high-resolution masters into a "dark archive" provided by an outside vendor. Several vendors are developing service plans for the preservation and archiving of digital resources.

Projects considering outsourcing storage of master images should consider the following general principles:

- Cost of ingestion and ongoing maintenance fees
- Metadata requirements for outsourced resources
- Does your workflow meet ingestion criteria (online harvesting vs. batch ingestion from hard media)
- Digital object preservation provided (bit stream integrity, virus protection, media migration, disaster recovery, file format migration)
- Accessibility requirements

Refreshment

Regardless of which storage technique your project chooses, refreshment of media, servers, and file formats is one of the most important aspects of a sustainable project. Refreshment is the transfer of digital files from one storage media to another to ensure that the files remain retrievable as technology advances. Do you have 5 1/4 inch floppy disks with files that you cannot retrieve because you do not have a computer with a 5 1/4 inch drive? This is a common occurrence and will continue to be a problem as computer technology evolves. To avoid this problem in the future, transfer files to new media as it becomes widely available. The next big change in storage technology, CD-ROM to DVD, is already here. Most new computers are now coming with DVD drives as standard equipment, rather than CD-ROM drives. Although current DVD drives will read CDs, as newer, faster DVD drives are built, they will no longer be able to do so. Considering the amount of time and money put into your digitization project, it is worth the effort and cost to refresh your files. Do not let more than five years elapse before refreshing your data. Longevity of the storage media is not as important as the ability to access the information.

Metadata

The creation of quality metadata is a key component for the responsible management and long-term preservation of the digital files produced by your project. Metadata is the term used to describe traditional descriptive cataloging applied to digital files in addition to information needed retrieve, access, and manage those files. Frequently metadata creation begins with pre-existing descriptive cataloging, finding aids, or accession records that are extended by adding information about the digital files, other projects will be creating metadata from scratch as part of the project's overall workflow.

There is no one standard for metadata creation that meets all the needs of all types of collections and repositories, however most common metadata schemes include the following sets of information:

Descriptive Metadata	Metadata that describes the intellectual content of a resource.
Administrative Metadata	Metadata that includes information about ownership and rights management.
Structural Metadata	Metadata that describes relationships between multiple digital files, such as page order in a digitized book or manuscript.
Technical Metadata	Metadata that describes the features of the digital file, such as resolution, pixel dimensions, compression, etc.

Projects using this guide should also review the *Western States Dublin Core Best Practices and Element Set* as an example of how to apply metadata to their projects.

Concern over long-term preservation and management of digital resources has placed an increased emphasis on the capture of technical metadata. Projects concerned with protecting their investment in digitization are recommended to review the NISO's *Data Dictionary: Technical Metadata for Digital Still Images* (Z39.87) (http://www.niso.org)

Intellectual Property Concerns

One of the largest concerns of new digitization initiatives is the protection of a institutions intellectual property. For more information about legal issues surrounding copyright see the Colorado Digitization Program web site. Projects using this document need to adapt the guidelines for minimum resolution and image quality to meet the institution's risk tolerance level.

Watermarks: what works and what doesn't

There are two types of watermarks currently in use for digital image files, visible and electronic.

Visible watermarks are applied on top of the image, very much like a seal is applied to an official document. Often these watermarks are the name of the institution who owns the file, that institution's official seal, or some other identifying logo. In most cases a visible watermark is applied to surrogate files for use on the internet in image manipulation software packages like Adobe Photoshop. This step can be done as part of a batch process to save time. In all cases, visible watermarks cover a portion of the image file.

Visible watermarks do not stop users from downloading files, however, they can be difficult to remove depending on the complexity, size, and color value of the design. The biggest drawback to visible watermarks is the obstruction to parts of the image, making use of that image file less appealing to some researchers.

Electronic watermarks are imbedded in the image file and they are invisible. They are usually a numeric code licensed by an electronic watermarking firm. The numeric code is specific to the institution that owns the files. Electronic watermarks are usually applied as part of the filter mechanism in programs like Adobe Photoshop. In some cases, upon very close inspection, the file will appear grainy after an electronic watermark is applied. This phenomenon is difficult to see, however, without comparing the watermarked file to the original.

Electronic watermarks do not stop users from downloading files and they can be easily overcome through manipulation. For example, if a digital file holding a watermark is posted online at 300ppi, a user can quite easily download the file, put it into Photoshop, drop the resolution to 72ppi and enough of the watermark information is removed to make it no longer detectable. This problem is resolved by posted watermarked images at low resolutions.

The advantage to using watermarks, visible or electronic, is that the institution is making a concerted effort to identify the image files as property of that institution. Watermarks do not prevent theft, but they do help prove the intent of the institution is to protect its' collections.

Guidelines for Creating Digital Images

These guidelines provide the **minimum** qualities we feel are necessary for achieving an acceptable level of image quality.

We recognize that all collections differ in the ways they are used and accessed and that institutions have differing purposes and clientele, which will likely have an impact on how and for what purposes and reasons collections are digitized. These are not hard and fast recommendations for every collection and every institution. As a rule, the key to quality scanning is not to scan at the highest resolution possible but to scan at a level that matches the informational content of the original.

Decisions on image quality and resolution should be based on the needs of users, how the images will be used, and the nature of the materials you are scanning (dimensions, color, tonal range, format, material type, etc.). The quality and condition of the original (such as the quality of the shooting or processing technique in the case of photographs) impacts on the resolution at which you scan and the resulting quality of the digital image.

Master Image File

Good digital imaging projects scan a high-quality "master" or archival image and then derive multiple versions in smaller sizes or alternative formats for a variety of uses. There are compelling preservation, access, and economic reasons for creating an archival-quality digital master image: it provides an information-rich, unedited, research quality surrogate, and ensures rescanning will not be necessary in the future. A high-quality master image will make the investment in the image capture process worthwhile. Since user expectations and technology change over time, a digital master must be available and rich enough to accommodate future needs and applications. ***The master image should be the highest quality you can afford***; it should not be edited or processed for any specific output; and it should be uncompressed. Intensive quality control should be applied in creating master image files.

Derivative Image Files

Derivative files are created from the master digital image, and are used in place of it, usually for general Internet or network access. Derivative files typically include an **access image**, which is sized to fit within the screen of an average monitor or other delivery mechanism and a **thumbnail image**, which is small enough to load quickly and linked to the larger access image. With the proper image editing software it is not necessary to subject source materials to multiple scans as derivative files can be created from the high quality master images.

We recommend that three versions of an image be created:

Master Image	Access Image	Thumbnail Image
• Represents as closely as possible the information contained in the original • Uncompressed • Unedited • Serves as long term source for derivative files • Can serve as surrogate for the original • High quality • Very large file size • Used for creating high quality print reproductions • Usually stored in the TIFF file format	• Used in place of master image for general web access • Generally fits within viewing area of average monitor • Reasonable file size for fast download time; does not require a fast network connection • Acceptable quality for general research • Compressed for speed of access • Usually stored in JPEG file format	• A very small image usually presented with the bibliographic record • Designed to display quickly online; allows user to determine whether they want to view access image • Usually stored in GIF or JPEG file formats • Not always suitable for images consisting primarily of text, musical scores, etc.; user cannot tell what content is at so small a scale

Modes of Capture

Most imaging equipment offer three modes for capturing a digital image:

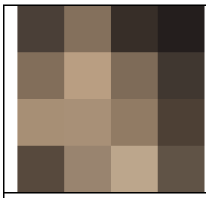
- ***Bitonal (black and white, line art)*** – One bit per pixel representing black and white. Bitonal scanning is best suited to high-contrast documents such as printed text.
- ***Grayscale (black and white photograph)*** – Multiple bits per pixel representing shades of gray. Grayscale is suited to continuous tone documents, such as black and white photographs.
- ***Color*** - Multiple bits per pixel representing color. Color scanning is suited to documents with continuous tone color information.

These three modes of scanning also require some subjective decisions. For example, a black and white typed document may have annotations in red ink. Although bitonal scanning is often used for typed documents, scanning in color may be preferable in this case, depending on how the image will be used. Manuscripts, older printed matter, and sheet music may be better served by scanning as continuous tone in grayscale or color to bring out the shade and condition of the paper and the marks inscribed on it. Projects interested in capturing the current condition of source materials should consider scanning in color.

Spatial Resolution

Spatial resolution measures the frequency at which individual pixels or points are sampled and is commonly referred to as "dots per inch" (dpi) or "points per inch" (ppi). Higher resolutions take more frequent samples of the original and contain a more accurate representation. Since higher resolutions are capturing more information, files sizes also increase.

There is no one "perfect" resolution to scan all collection materials. Spatial resolutions should be adjusted based on the size, quality, condition, and uses of the digital object. See Guidelines by Source Type for specific spatial resolution targets.

			
4 ppi	72 ppi	150 ppi	300 ppi

Detail of Serene Donner in a Picture Hat, 1910
courtesy Ira M. Beck Memorial Archives, University of Denver

Funded by an IMLS grant to
The University of Denver and
The Colorado Digitization Program

3 inches

5000 pixels on the
long dimensions
or
~ 1600 ppi

2 inches

Serene Donner in a Picture Hat, 1910
courtesy Ira M. Beck Memorial Archives, University of Denver

These two examples demonstrate that spatial resolution depends on size and quality of the original. The smaller 3" x 2" print above requires a higher resolution than the larger 8" x 10" print below. The print above also suffers from oxidation and its condition warranted increasing the resolution beyond the minimum requirement.

10 inches

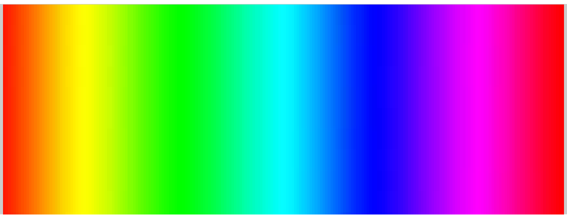
8 inches

3000 pixels on the long dimension
or ~300 ppi

Kayaking the Poudre River, 1968
courtesy Larimer County Digitization Project

Bit Depth (Color Depth)

Bit depth measures the number of colors (or levels of gray in grayscale images) available to represent the color/gray value in the original work. (A bit is the basic digital building block with a value of either 1 or 0). Every pixel sampled is assigned a value that corresponds to the color/shade it represents. The higher bit depths (the more information you capture) increases the number of colors/shades available and correspondingly increases the file size.

1 bit (bitonal)	2 colors/shades of gray, usually black or white		
8 bit	256 colors/shades of gray		
24 bit	16 million colors/shades of gray		

File Formats

Master Images

Master digital images should be stored in a file format that supports the fidelity and long-term preservation of the image. The master object file format should be:

- Non-proprietary / open source
- Uncompressed
- Ability to capture technical metadata as part of the file structure

The format most frequently used for master digital images is the Tagged Image File Format (TIFF).

Delivery/Access Images

In order to be delivered across a network in a timely fashion, large high-quality master images are normally converted to smaller lower-quality access images. Master images might also be converted into a format used to create a print publication that provides access to the collection.

File Compression

Lossy

In lossy compression, a certain amount of information is discarded during the compression process. Although the discarded information may be invisible to the human eye, a loss of quality occurs. Lossy compression formats also introduce *generational loss* – each time a lossy image is manipulated or edited the quality of the image decreases. Generational loss is a primary reason that master images are not stored using lossy compression.

The Joint Photographic Experts Group (JPEG) format is most frequently used for access images requiring lossy compression. The JPEG compression algorithm was designed for the continuous tone images.

		
JPEG High Compression Low Quality 5 kilobytes	JPEG Medium Compression Medium Quality 10 kilobytes	JPEG Low Compression High Quality 82 kilobytes
Detail of Serene Donner in a Picture Hat, 1910 *courtesy Ira M. Beck Memorial Archives, University of Denver*		

Converting images from a bit depth of 24 bit to a bit depth of 8 is also considered a lossy compression method as color information is discarded.

Lossless

Lossless compression results in a file similar to the original image, with no loss of information. The Graphic Interchange Format (GIF), Tagged Image File Format (TIFF) and Portable Network Graphic (PNG) formats support lossless compression.

Proprietary Formats

Proprietary formats are controlled or owned by a particular entity that licenses the format for use by others. Proprietary formats often require special plug-ins or software for viewing that the owner provides. Proprietary formats are not recommended for master images because licensing requirements may prevent the long-term access and preservation of images. Examples of proprietary file formats include: Photoshop (psd); Kodak PhotoCD (pcd), and Portable Document Format (pdf), and Encapsulated PostScript (eps).

Guidelines by Source Type

Image Quality Calculator

In addition to the guidelines below some projects have adapted formulas used for the quality of microform materials to calculate image quality based on the smallest detail that needs to be legible to the eye.

See the Image Quality Calculator at the University of Illinois for additional examples and information: http://images.library.uiuc.edu/projects/calculator/

Text			
	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 across the long dimension	150 – 200 pixels across the long dimension

Alternative format

PDF (Portable Document Format) from Adobe is an alternative file format for creating and displaying files on the web. PDF format provides the ability to construct complex multi-page objects such as books, journals, or diaries. See http://www.adobe.com for further information. Requires Adobe Acrobat software to create and manipulate files. The Adobe Acrobat viewer is free to download so users can view documents on their computers.

Other considerations

When scanning text documents, spatial resolutions should be based on the size of text included in the document and resolutions should be adjusted accordingly. Documents with smaller printed text may require higher resolutions and bit depths than documents that use large typefaces. Image Quality Calculators (see above) can provide help in selecting the best resolution for scanning text documents. Projects planning on applying Optical Character Recognition (OCR) may wish to test pages at several resolutions to find the most satisfactory results. Images that produce the best results for OCR may not be be pleasing to the eye and may require separate scans for OCR and human display.

Projects with large amounts of textual materials, particularly hard-to-read materials such as manuscripts, should consider providing transcriptions of the materials in addition to the digital image. Access to textual material can be further enhanced through SGML/XML markup schemes such as the Text Encoding Initiative (TEI). As re-keying text can be cost prohibitive, projects considering transcriptions should investigate including Optical Character Recognition (OCR) software in their toolkit.

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 across the long dimension	150 – 200 pixels across the long dimension

Alternative formats

Many imaging projects are using the proprietary Kodak PhotoCD (pcd) format for converting their photographic images. For more information on Kodak PhotoCD can be found at Kodak's site and in an article in RLG Diginews, *Using Kodak PhotoCD for Preservation and Access* at For long-term preservation consider converting the proprietary PhotoCD format into TIFF format.

Portable Network Graphics (PNG) is an image format designed to replace the GIF format. It offers a smaller file size than GIF but does not lose any information to compression. It is not yet widely supported.

Other considerations

Photographs can present many scanning challenges. We recommend scanning from the negative (or the earliest generation of the photograph) to yield a higher-quality image. Another consideration is whether to scan sepia-tone photographs as color or black and white images. We

Funded by an IMLS grant to
The University of Denver and
The Colorado Digitization Program

recommend scanning them as color images to create a better image, although this will increase the file size.

Another consideration with photographs is whether to scan the backs of photographs as separate image files if there is significant information on the back of the photo (which may be of interest to users) that may not be included elsewhere. If a scanned image of the verso of the photograph is available, the digital image may serve as a more successful surrogate for the original.

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	3000 pixels across the long dimension	150 dpi	72 dpi
Spatial Dimensions	100% of original	600 pixels across the long dimension	150 – 200 pixels across the long dimension

Alternative formats

The GeoExpress (formerly MrSID, MultiResolution Seamless Image Database) format by LizardTech, Inc. allows for the compression, storage, and retrieval of large digital images. Files are stored in proprietary format and are compressed with a "wavelet" compression algorithm that also provides a "zoom in" capability in the browser software that provides little loss in image quality. LizardTech provides viewers to those who wish to download and manipulate .GeoExpress images, but the technology can be used to deliver a portion of the image requested as a standard JPEG, with no viewers required.

Other considerations

When scanning maps, spatial resolutions should be based on the size of smallest detail included in the map and resolutions should be adjusted accordingly. Documents with smaller details may require higher resolutions and bit depths than maps that use larger details and typefaces. Image Quality Calculators (see above) can provide help in selecting the best resolution for scanning maps.

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 across the long dimension	150 – 200 pixels across the long dimension

Artwork / 3 Dimensional Objects

Standards for artwork are not well defined. Usually artwork imaging projects involve scanning from photographic surrogates such as 35mm slides, in which case recommendations for transparent photographs should be followed. For large format artwork, outsourcing to a vendor with an overhead digital camera or large flatbed scanner suitable for scanning large documents is recommended.

If you do choose to distribute master images over the web for access by users, you may want to consider digital watermarking or some kind of copyright/ownership mark, possibly embedded in the image itself, as master image files are of a quality that can be used for commercial reproduction. The access and thumbnail files are for web display only, and are not of a quality suitable for reproduction.

Quality Control for Images

Many projects being digitization considering resolution the most important factor in determining image quality. In fact, numerous other factors play as important a role in the final outcome of a digitization project. Original condition of materials, quality and maintenance of equipment, staff training, external lighting, are some factors that can influence the quality of images.

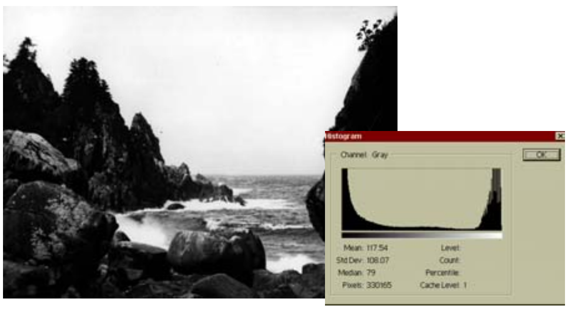
Tonal Dynamic Range.

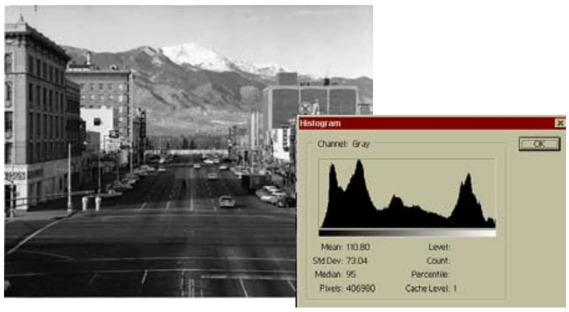
One of the most significant factors affecting image quality is the Tonal Dynamic Range – the color space an image occupies between pure white (255) and pure black (0). The tonal dynamic range can be displayed in professional level TWAIN drivers and image editing software such as Photoshop. The histogram graphically displays the number of pixels in the image at selected values from white to black. Reviewing histograms at the time of scanning can ensure that all of the image's information is being captured and not lost due to "clipping and spiking."

Clipping & Spiking

Clipping and spiking result when the white and black points are not set on true white and black during the set up of the scan. If white and black are improperly set, everything above or below those points is "clipped" or registers as the same tone.

Spiking on the ends of the histogram usually indicates clipping. This problem also shows up in the image itself as blockage and pixelization in the shadows and blowouts in the highlights. Acceptable spikes can occur if the edge of the original negative has lost emulsion, for example, or the sky holds no detail and is one tone in the original. Such instances, however, are rare.





(Left) Detail in the image on the left has been lost due to improperly adjusted white and black points. The histogram shows the "clipping and spiking" associated with incorrect points. The image on the right shows properly adjusted white and black points with no clipping or spiking in the histogram. ***(Left) Rocky coastline on Forrester Island, 1920; courtesy Denver Museum of Nature and Science. (Right) Downtown Colorado Springs, 1964; courtesy Pikes Peak Library System.***

White & Black Points

Optimum placement of whites and blacks is best observed through the histogram, although the image itself will be examined without the histogram on the screen. It is important to look at the number value assigned to the brightest highlight and the darkest shadow. Highlights should not read a number value higher than 247 and shadows should not be less than 7 or 8. If these numbers are exceeded the scan must be redone. This is particularly vital if the original image has a short dynamic range. The white and black points must not be set on 0 and 255, as this will stretch the dynamic range of the image, creating gaps in the histograms, and thus unusable scans.

These adjustments need to be made at the time of scanning since adjusting images after the fact in image editing software introduces interpolation of the dynamic range (guessing what the points are in the image) and frequently results in gaps in the histogram rather than continuous curves.

Color Management

Color management can be one of the most challenging and costly aspects of the digitization workflow. Each piece of hardware in the path from source to digital file can introduce biases of color and tone. The goal of color management systems is to accurately and predictably compensate for each of these biases across the entire system from scanner to print output or monitor display.

Due to the cost of hardware, software, staff time and expertise required for effective color management this guide recommends it be used only by projects concerned about color fidelity (such as reproductions of works of art). Please see the Color Management Case Study for additional information about implementing color management in your project.

Projects not undertaking color management should be aware that equipment will introduce color biases into any digitized materials. Attempting to make digitized materials “look good” on un-calibrated equipment may introduce these biases into the master images.

Projects without a color management system should use available tools to perform basic monitor calibrations:

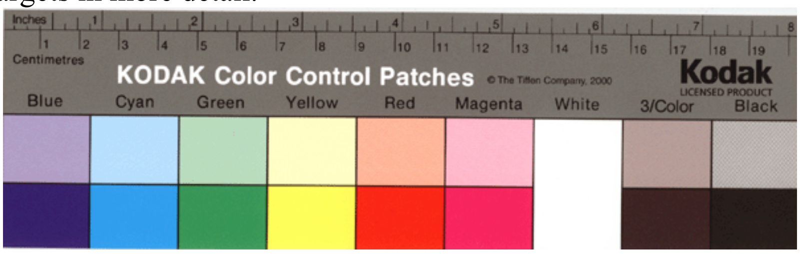
- Set to 24 millions of colors
- Set monitor Gamma at 2.2 (including Macintosh computers, that by default are set at 1.8 gamma)
- Color temperature at 6500° K

The Denver Public Library - Western History & Genealogy Department provides instructions for calibrating a monitor to view digitized collections. Including similar instructions with targets will help users adjust their equipment for optimal viewing of online resources.

Targets and Color Bars

Targets and color bars are used to measure system resolution, tonal range and color fidelity. Including targets in a digitization workflow allows color management systems to create profiles for each device or for later adjustment in projects not implementing color management during scanning. Targets are a way of predicting image quality, and help ensure that the scanning system you are using is producing the best quality image it can and is operating at a consistent level of quality over time. Targets for prints and transparencies exist and targets appropriate for the materials being scanned should be used (paper, film, transparency, etc.). Targets usually contain patches of color, black and white, or shades of gray for verifying tone reproduction. Resolution targets allow projects to measure the level of detail a particular piece of equipment can capture. Resolution targets can be helpful in evaluating equipment before purchase or assessing the quality of output from a vendor.

As equipment biases can shift over time best practice is to include targets on a regular basis throughout the course of the project. Some digitization projects are also scanning a color bar along with the original, to be included in the final digital image, to aid users in verifying accuracy in color reproduction. Please see NARA's Guidelines for Digitizing Archival Materials and Kenney and Chapman's *Digital Imaging for Libraries and Archives* for a discussion of calibration and targets in more detail.



Kodak Color Control Patches

A few of the commonly available targets are:

- Kodak Color Separation Guide
- Greytag-Macbeth Color Checker
- IT8.7 (1/2) color output target (reflective/transmissive)
- IEEE Standard Facsimilie Test Chart
- RIT Alphanumberic Resolution Test Object target
- USAF 1951 Resolution Target
- ISO Camera Resolution Chart

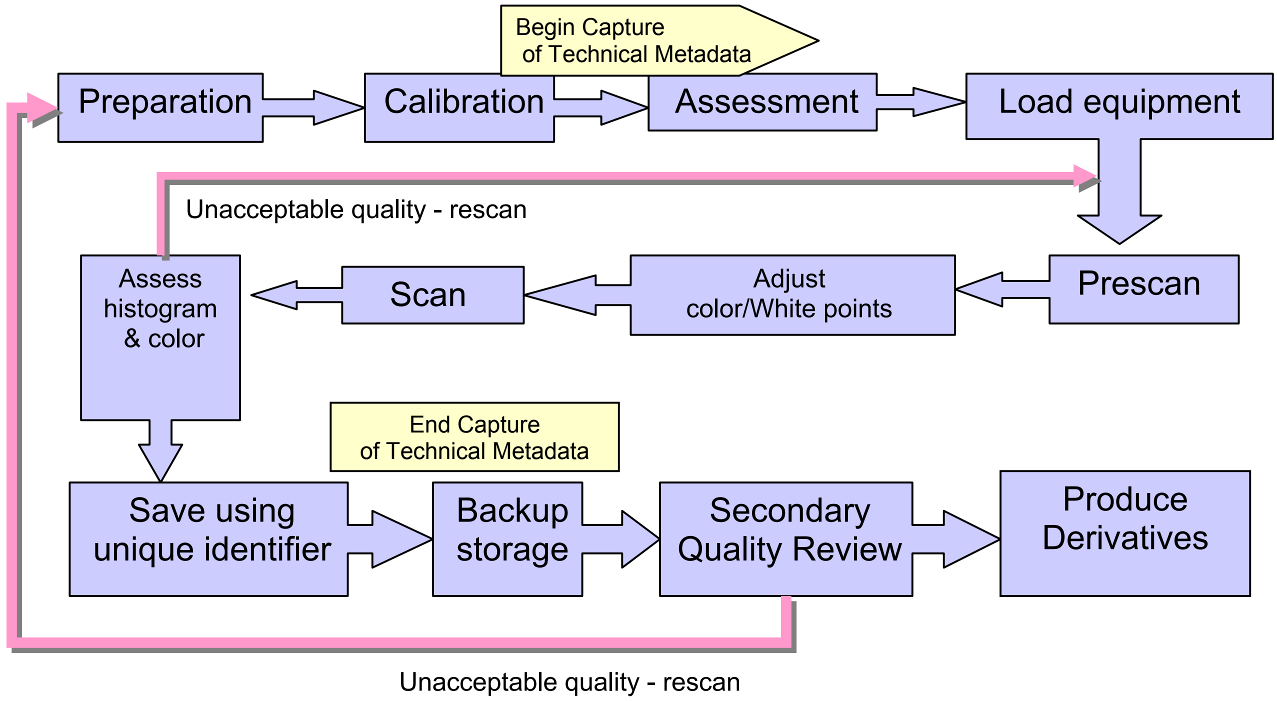
Quality Control Workflow

A quality control program should be conducted throughout all phases of the digital conversion process. Inspection of final digital image files should be incorporated into your project workflow. Typically, master image files are inspected via CD batch or online for a variety of defects. Depending on your project, you may want to inspect 100% of the master images, or 10% of the files randomly.

We do recommend that quality control procedures are implemented and documented and that you have clearly defined the specific defects that you find unacceptable in an image. Images should be inspected while viewing at a 1:1 pixel ratio or at 100% magnification or higher. Quality is evaluated both subjectively by project staff (scanner operator, image editors, etc.) through visual inspection and objectively in the imaging software (such as using targets, histograms, etc.).

Things to look for during visual inspection may include:

- Image not the correct size
- Image not the correct resolution
- File name is incorrect
- File format is incorrect
- Image is in incorrect mode (i.e., color image has been scaled as grayscale)
- Loss of detail in highlight or shadows
- Excessive noise especially in dark areas or shadows
- Overall too light or too dark
- Uneven tonal values or flare
- Lack of sharpness/Excessive sharpening
- Pixellated
- Presence of digital artifacts (such as very regular, straight lines across picture)
- Moiré patterns (wavy lines or swirls, usually found in areas where there are repeated patterns, such as half-tone dots)
- Image not cropped
- Image not rotated or is reversed
- Image skewed or not centered
- Incorrect color balance
- Image dull or no tonal variation
- Negative curve in the Look-Up Table
- Clipping black and white values (in histogram)


Begin Capture of Technical Metadata
Preparation
Calibration
Assessment
Load equipment
Unacceptable quality - rescan
Assess histogram & color
Scan
Adjust color/White points
Prescan
End Capture of Technical Metadata
Save using unique identifier
Backup storage
Secondary Quality Review
Produce Derivatives
Unacceptable quality - rescan

---

[1] Chapman, Stephen. *Handbook for Digital Projects: A Management Tool for Preservation and Access* http://www.nedcc.org/digital/intro.html

[2] Kenney, Anne and Steven Chapman. *Digital Imaging for Libraries and Archives*. Ithaca: New York, Department of Preservation and Conservation, Cornell University Library, June 1996.

[3] As of this writing the DVD-R format had not been standardized and increased the possibility of incompatibility across different hardware platforms.

[4] Density chart adapted from Technical Advisory Service for Images. *Scanners* http://www.tasi.ac.uk/advice/creating/scanners.html and Kenny & Reiger. *Moving Theory into Practice*, p. 39.

[5] Density chart adapted from Technical Advisory Service for Images. *Scanners* http://www.tasi.ac.uk/advice/creating/scanners.html and Kenny & Reiger. *Moving Theory into Practice*, p. 39.