



# University of Colorado Digital Library Digitization Best Practices Version 1.0

---

## Contents:

- I. [ABOUT THIS DOCUMENT](#)
  - II. [INTRODUCTION](#)
  - III. [DIGITIZATION: AN OVERVIEW](#)
  - IV. [MINIMUM REQUIREMENTS FOR DIGITAL IMAGING AT THE UNIVERSITY OF COLORADO](#)
  - V. [MINIMUM REQUIREMENTS FOR DIGITAL AUDIO AT THE UNIVERSITY OF COLORADO](#)
  - VI. [CU DIGITAL LIBRARY'S DIGITIZATION BEST PRACTICES WORKING GROUP](#)
- 

## I. ABOUT THIS DOCUMENT

Collections in the University of Colorado Digital Library must follow appropriate digitization standards. This document offers an introduction to digitization, provides links to resources containing more information, and describes the recommended digitization parameters for collections in the CU Digital Library. The document is maintained by the University of Colorado's Digitization Best Practices Working Group: Lynn Lickteig (College of Architecture & Planning), Holley Long (University Libraries' Systems Department), and Elaine Paul (Department of Art and Art History).

## II. INTRODUCTION

The University of Colorado Digital Library collections can contain a variety of digital media including images, text, audio, and moving images. The CU Digital Library has adopted minimum guidelines for imaging and audio digitization specifications. Adherence to these guidelines ensures the quality, consistency, and longevity of these valuable resources.

## III. DIGITIZATION: AN OVERVIEW

### Digital Imaging

According to Cornell University Library's *Moving Theory Into Practice* tutorial, digital imaging is defined as: "electronic snapshots taken of a scene or scanned from documents, such as photographs, manuscripts, printed texts, and artwork. The digital image is sampled and mapped as a grid of dots or picture elements (pixels). Each pixel is assigned a tonal value (black, white, shades of gray or color), which is

represented in binary code (zeros and ones). The binary digits ("bits") for each pixel are stored in a sequence by a computer and often reduced to a mathematical representation (compressed).

(<http://www.library.cornell.edu/preservation/tutorial/intro/intro-01.html>) This definition introduces the basic concepts described in the digital imaging guidelines.

Digital images come from many sources, including photographs, maps, textual documents, artwork and are created in one of two ways. They can be scanned from analog material. Alternately, images can be "born digital", i.e. originally created in a digital format. Photos taken with a digital camera or Microsoft Word documents are examples of "born digital" materials.

Digital images are composed of pixels, or picture elements. Increasing the number of pixels improves image quality. Resolution is a measure of the number of pixels used to render an image and is expressed as "pixels per inch" or ppi. Resolution can also be measured by the total number of pixels on the longest side.

Each pixel also has a tonal value: black, white, shades of gray, or color. The number of tones available to create the image is called bit-depth. For example, an image scanned in 8 bit-depth can use up to 256 colors or shades of gray.

Digital images are saved in a file format, the structure by which data is organized in a file. Common file formats include TIFF (Tagged Image File Format), JPEG (Joint Photographic Experts Group), GIF (Graphics Interchange Format), BMP (Bit-Mapped), and JPEG-2000. Some file formats discard redundant information to reduce the file's size, a process known as compression. Best practices suggest saving digital images in an uncompressed file format, such as TIFF.

## **Digital Imaging Terms**

*Bit Depth* the tonal or signal resolution; determines maximum number of shades of gray or colors in a digital file.

*Color Mode* refers to whether the image is black and white, grayscale, or color. Grayscale images consist of a single channel and can be 8-bit (256 levels) or 16-bit (65,536 levels). Color images consist of 3 or more grayscale channels that represent color and brightness information and may be either 8-bits or 16-bits. Common color modes are RGB, CMYK, and LAB color.

*Compression* a process that eliminates redundant data to create a smaller file size.

*Dimensions* the size or measurement of an image's height and length, recorded in inches or centimeters.

*File Format* a structure for encoding the information in a data file.

*Pixel Array* a measurement of the spatial resolution or the amount of information in an image file expressed as the number of pixels on each dimension of the image.

*Resolution* a measurement of the spatial resolution, written as *pixels per inch* or “ppi”. The term “dpi” refers to printer resolution or *dots per inch* and is often used interchangeably for ppi.

## Digital Audio

The Collaborative Digitization Program’s *Digital Audio Best Practices* describes audio digitization as follows: “... an analog recording is played back through an electronic device, and the variations of the electric current generated by the device are sampled at very fast time intervals. The amplitude of the current, corresponding to the amplitude of the original sound wave, is recorded as a number at each sampling point.” (<http://www.bcr.org/dps/cdp/best/digital-audio-bp.pdf>)

## Digitization Standards

### Digital Imaging

- Digital Imaging Best Practices (Collaborative Digitization Program) <http://www.bcr.org/dps/cdp/best/digital-imaging-bp.pdf>
- Guidelines for Digital Images (California Digital Library) [http://www.cdlib.org/inside/diglib/guidelines/bpgimages/cdl\\_gdi\\_v2.pdf](http://www.cdlib.org/inside/diglib/guidelines/bpgimages/cdl_gdi_v2.pdf)
- Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images (National Archives and Records Administration) <http://www.archives.gov/preservation/technical/guidelines.pdf>

### Digital Audio

- Digital Audio Best Practices (Collaborative Digitization Program) <http://www.bcr.org/dps/cdp/best/digital-audio-bp.pdf>
- Sound Directions: Best Practices for Audio Preservation (Indiana University, Harvard University) [http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/sd\\_bp\\_07.pdf](http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/sd_bp_07.pdf)

## Selected Resources

Cornell University Library. *Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems*. 2007. [http://www.icpsr.umich.edu/dpm/dpm-eng/eng\\_index.html](http://www.icpsr.umich.edu/dpm/dpm-eng/eng_index.html)

Cornell University Library. *Moving Theory Into Practice: Digital Imaging Tutorial*. 2003. <http://www.library.cornell.edu/preservation/tutorial/contents.html>

Joint Information Systems Committee. *TASI Technical Advisory Service for Images*. 2002. <http://www.tasi.ac.uk/>

Kenney, Anne R., and Oya Y. Rieger. *Moving Theory Into Practice: digital imaging for libraries and archives*. Mountain View, CA: Research Libraries Group, 2000.

National Initiative for a Networked Cultural Heritage. *The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials*. 2002. <http://www.nyu.edu/its/humanities/ninchguide/>

National Library of Australia. *PADI Preserving Access to Digital Information*. <http://www.nla.gov.au/padi/index.html>

#### **IV. MINIMUM REQUIREMENTS FOR DIGITAL IMAGING AT THE UNIVERSITY OF COLORADO**

The following minimum guidelines are excerpted from NARA's *Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images* (<http://www.archives.gov/research/arc/digitizing-archival-materials.pdf>). Reading NARA's document before selecting digitization parameters is strongly recommended. While the following guidelines represent minimum specifications, digitizing materials to the highest feasible standards increases the longevity and quality of the digital objects.

##### **Minimum Guidelines for Text Documents, Photographs, Illustrations, Artwork, Maps, and Oversized Items**

|  | Resolution | Bit Depth | Color Mode | Archival File Format |
|--|------------|-----------|------------|----------------------|
| Items that are black & white or shades of gray | 300 ppi    | 8-bit     | grayscale  | TIF                  |
| Items containing color                         | 300 ppi    | 24-bit    | RGB color  | TIF                  |

### Minimum Guidelines for Slide or Negative Film

|  | Pixel Array  | Resolution | Dimensions  | Bit Depth | Color Mode | Archival File Format |
|--|--|------------|---|-----------|------------|----------------------|
| Rectangular format original<br><br>Black and white | 3,000 pixels across the long dimension<br><br>4,000 pixels across the long dimension (preferred) | 300 ppi    | 10 inches on the long dimension<br><br>or<br><br>exact size of original for smaller objects | 8-bit     | Grayscale  | TIF                  |
| Rectangular format original<br><br>Color           | 3,000 pixels across the long dimension<br><br>4,000 pixels across the long dimension (preferred) | 300 ppi    | 10 inches on the long dimension<br><br>or<br><br>exact size of original for smaller objects | 24-bit    | Color      | TIF                  |
| Square format<br><br>Black and white               | 2,700 pixels across the long dimension   | 300 ppi    | 10 inches on the long dimension<br><br>or<br><br>exact size of original for smaller objects | 8-bit     | Grayscale  | TIF                  |
| Square format<br><br>Color                         | 2,700 pixels across the long dimension   | 300 ppi    | 10 inches on the long dimension<br><br>or<br><br>exact size of original for smaller objects | 24-bit    | Color      | TIF                  |

While reference to the pixel array of a digital image has become a common standard for referencing the size of the digital file, the NARA guidelines state that “to specify a desired file size, it is always necessary to provide both the resolution and the image

dimensions.” Although pixels per inch (ppi) are commonly used to refer to the size of digital files created from text or other document originals, for 35mm film and slides it is acceptable to refer to the overall megabytes of the file size, as in a “50MB TIFF file”. If the above guidelines for scanning original slides and negatives are followed, the resulting preservation master file will be approximately a “20MB TIFF” file.

### **Minimum Guidelines for “Born Digital” Image Files**

Recommendations for the minimum file size of “born digital” files are obviously dependant on the technical capabilities of the device which created the original digital files. As of 2005, NARA recommends that a minimum of a 6 megapixel camera with a minimum pixel array of 3,000 pixels on the long dimension be used. Born digital files that have been created on cameras of less than 6 megapixels or at less than 300 ppi may still have educational or aesthetic merit, particularly if the digital files represent difficult to obtain or unique “one of a kind” materials.

### **Other Considerations**

#### *Derivative Files*

Once the digital preservation master file is created, smaller derivative or access digital files may be created at the collection manager’s discretion from the master file utilizing different file formats or technical settings for a variety of purposes, from reference or study images on web sites to classroom PowerPoint or Keynote presentations. The idea is to “scan once” as a means of time savings and to ensure that a high quality image is created from which all other copies may be made in the future.

While the file format of the derivative or access files may not concern our Luna Insight implementations, since by default they are converted from the TIFF files to JPEG 2000 files when ingested by the Insight software, NARA generally recommends GIF file format over JPEG for derivative images.

#### *Optical Resolution versus Interpolation and Digital Zoom*

NARA guidelines state that digitization projects should use “digital cameras and scanners that produce records with true optical resolution.” Many digital cameras and scanners offer extra digital zoom features (in addition to true optical zoom) or some form of interpolation which creates higher resolution files from lower resolution files. For purposes of the preservation of digital files NARA recommends turning off interpolation features on digital cameras and scanners.

## **V. MINIMUM REQUIREMENTS FOR DIGITAL AUDIO AT THE UNIVERSITY OF COLORADO**

The following minimum guidelines are excerpted from the Collaborative Digitization Program’s *Digital Audio Best Practices* (<http://www.bcr.org/dps/cdp/best/digital-audio-bp.pdf>).

Reading CDP's document before selecting digitization parameters is strongly recommended. While the following guidelines represent minimum specifications, digitizing materials to the highest feasible standards increases the longevity and quality of the digital objects.

|                 | Sample Rate                                      | Bit Depth | Archival File Format |
|-----------------|--|-----------|----------------------|
| Spoken language | 44.1 kHz   | 16-bit    | WAV, AIF             |
| Music           | 44.1 kHz<br>(minimum)<br><br>96 kHz<br>(optimal) | 16-bit    | WAV, AIF             |

## VI. CU DIGITAL LIBRARY'S DIGITIZATION BEST PRACTICES WORKING GROUP

The following people have contributed to these best practices. Please contact any one of them with questions.

Lynn Lickteig  
[Lynn.Lickteig@colorado.edu](mailto:Lynn.Lickteig@colorado.edu)  
 Director of Visual Resource Center  
 College of Architecture & Planning  
 University of Colorado Denver\Boulder

Holley Long  
[Holley.Long@colorado.edu](mailto:Holley.Long@colorado.edu)  
 Systems Librarian for Digital Initiatives  
 Systems Department  
 University of Colorado at Boulder Libraries

Elaine Paul  
[Elaine.Paul@colorado.edu](mailto:Elaine.Paul@colorado.edu)  
 Director of Visual Resources Center  
 Department of Art and Art History  
 University of Colorado at Boulder