GUIDELINES FOR THE CREATION OF DIGITAL COLLECTIONS

Digitization Best Practices for Images

This document sets forth guidelines for digitizing two-dimensional, non-textual materials for the CARLI Digital Collections. The issues described concern image quality, file formats, storage, and access.

This document was created by the CARLI Digital Collections Users' Group (DCUG) and is maintained and updated by the CARLI Created Content Committee (CCC).

For questions about this document, please contact CARLI at support@carli.illinois.edu

Introduction

Images may include such items as photographs, maps, plans, blueprints, drawings, paintings, and other two-dimensional visual media.

In many instances, images will contain or have accompanying textual material. Due to this dual nature, the digitization of images is very similar to the digitization of text. For further information on text digitization, consult the "Guidelines for the Creation of Digital Collections: Digitization Best Practices for Text" document maintained by Created Content Committee: http://www.carli.illinois.edu/sites/files/digital_collections/documentation/guidelines_for_text.pdf

The sections below provide guidance on the processes of creating digital images.

Image Collections

Although no universal standards for quality image capture exist and technical standards are constantly evolving, member institutions participating in the CARLI Digital Collections should adhere to the best practices adopted by recognized leading institutions when the collections' purpose would benefit from those actions.

This document does not address best practices for preserving digital files. Digital preservation resources can be found at: Digital Preservation Best Practices and Guidelines, made available by the North Carolina Department of Natural and Cultural Resources.

Digital Images

A digital image is a two-dimensional array of small square regions known as pixels. Images typically fall into one of three categories: monochrome (bitonal), grayscale, and color. In the case of a monochrome image, the brightness of each pixel is represented by a numeric value. Black and white are the most common pairing of pixel values, but any two colors may be used. Grayscale images typically contain values in the range from 0 to 255, where 0 represents black, 255 represents white, and values in between represent shades or intensity of gray. A color image can be represented by a two-dimensional array of Red, Green and Blue triples, where 0 indicates that none of that primary color is present in that pixel and 255 indicates a maximum amount of that primary color.

Creating Images

At least one high-quality digital master or archival image file should be created for each object photographed or scanned. From that master file derivative files can be created, such as:

- An access image (an image used for detailed on-screen viewing)
- A thumbnail image (for fast access during search, browse and retrieval). Often the thumbnail is automatically created by the digital asset management software (DAMS) used to display images.

The Federal Agencies Digital Guidelines Initiative (FADGI), defines two master image file types that can be created, along with access images.

Archival Master Image	Production Master Image	Access Image
• Represents as closely as possible the information contained in the original • Uncompressed, or lossless compression • Unedited and "use-neutral" • Serves as long term, sustainable resource • Can serve as surrogate for the original • Highest quality that can be produced by the digitizing organization • Large file size • Stored in the TIFF file format	• Produced from archival master • Uncompressed, or lossless compression • May be edited for technical corrections to the file, such as adjusting tone or stitching together multiple images of a large item (e.g. maps) into one file • Used to create derivative files for access • Large file size • 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 or JPEG2000 file format

From the Federal Agencies Digital Guidelines Initiatives (FADGI), *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files*. Publication Date: September 2016. May be accessed from: http://www.digitizationguidelines.gov/guidelines/digitize-technical.html

Master Images

Due to the stress of digitizing unique materials, a digital master should be generated for every object created. The digital master image represents as accurately as possible the visual information in the original object. If possible, both an archival master file and a production master file should be kept.

The archival master image's primary function is to serve as a long-term archival record for preservation and be used as infrequently as possible. From that file, a production master file can be created, and used to create derivative images (access and thumbnails) or for other purposes. The production master may also have to serve as the archival master file if digital storage is limited, or if the institution is not concerned with preservation of digital images. Digital master files are measured in ppi (pixels per inch). Master files are most often saved to a designated server or other long-term storage device. See the NSDA link provided above for more resources on digital preservation.

Master images should be scanned at the highest level of quality the digitizing institution can produce to avoid re-handling of any original materials. Archival master images should not be edited for any specific output or use, and should be saved as large TIFF files with lossless or no compression. Production master images may have minimal editing.

	File Format	Resolution in Pixels Per Inch (ppi)	Bit Depth per color channel
Master Image, Photographs and Prints	TIFF	400-600 ppi. Very fine details might require higher ppi. 600 is preferable for highest-quality archival master.	8- or 16-bit color (24- or 48-bit total)
Master Image, Oversized Items (e.g. maps, posters)	TIFF or JPEG2000 Large originals may be saved as tiles if necessary.	300-400 ppi. Very fine details might require higher ppi. 400 is preferable for highest-quality archival master.	8- or 16-bit color or grayscale, depending on original (24- or 48-bit total)

Values can vary outside of the ranges given depending on source material. Specifications taken from FADGI, *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files*. September 2016. http://www.digitizationguidelines.gov/guidelines/FADGI Federal Agencies Digital Guidelines Initiative-2016 Final_rev1.pdf

Creating digital master files:

- Guidelines for file size and resolution of digital master files will vary by collection based on end user needs, sizes and types of original objects, software specifications, available file storage space, etc.

- Each library should develop specific scanning guidelines based on individual collection needs and requirements.

- Where possible, scanning guidelines for creation of digital master files should follow the specifications outlined in the FADGI *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files*. http://www.digitizationguidelines.gov/guidelines/FADGI Federal_Agencies Digital Guidelines Initiative-2016 Final_rev1.pdf

- CARLI member libraries using CONTENTdm **should not** upload full resolution TIFF files to the CARLI server. Archival image file storage is the responsibility of each contributing institution and must be managed locally. The CONTENTdm Project Client can automatically convert TIFF files into JPEG2000 or JPEG display images.

Derivative Images

From the master file, derivative files can be created for general use. Derivative images can be used to create access image used in DAMS systems and/or for creating reproductions. They may be edited and enhanced or converted to different formats. For each master image, two derivative files are often created: an access image (for more detailed onscreen viewing) and a thumbnail image (for searching and browsing). In the case of collections using CONTENTdm, the software is configured to automatically generate access and thumbnail images from a master file.

Recommendations

	File Format	**Resolution in Pixels Per Inch (ppi)**	**Bit Depth per color channel**
Access Image	Any, normally JPEG or JPEG2000	400-200 ppi	1-bit bitonal, 4- or 8-bit grayscale, or 8-bit color (24-bit total)
Thumbnail Image	Any, often JPEG or GIF	100-200 ppi	1-bit bitonal, 4- or 8-bit grayscale, or 8-bit color (24-bit total)

Some recommendations based on 1-. 2-, and 3- star specifications from FADGI, *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files*. September 2016.
http://www.digitizationguidelines.gov/guidelines/FADGI Federal Agencies Digital Guidelines Initiative-2016 Final_rev1.pdf

File Naming Conventions

Each digital object in a collection should be assigned a unique identifier. Unique identifiers should follow a consistent naming format to ensure ongoing identification and retrieval of digital files.

Guidelines for file names will vary by collection based on local needs and specifications. Each library should develop specific file naming conventions based on individual collection needs and local requirements.

Examples from the CARLI CONTENTdm collections include the following:
ACWC0002
bru004_03_nF
Images_W-04.jpg
Lor137_01.jpg

More information on best practices for file naming conventions may be found on the National Institute of Standards and Technology website at: https://www.nist.gov/system/files/documents/pml/wmd/labmetrology/ElectronicFileOrganizationTips-2016-03.pdf.

Monitor Calibration

Monitors used for image editing and color correction should be calibrated according to the following specifications[1]:

- Set to 24 millions of colors
- Set monitor Gamma at 2.2
- Color temperature at 5000 degrees K. The FADGI recommendation is to "Adjust the illumination and color temperature of the monitor to provide the best approximation of white in the viewing environment to the digital representation of white on the monitor."

Monitor calibration software can be selected and purchased by member libraries and will vary depending on local budgets, equipment and software specifications.

Technical Metadata

In the interest of preservation and reproduction, it is helpful to capture technical metadata in the creation of the digital image file. Digital cameras and scanners can automatically capture this information and embed it in the object file. NISO Standard Z39.87-2006 (R2017) (Data Dictionary - Technical Metadata for Digital Still Images) is widely accepted for use in the management of technical metadata. Among the attributes that can be described by the technical metadata are the following[2]:

---

[1] From FADGI *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files*. http://www.digitizationguidelines.gov/guidelines/FADGI Federal_Agencies Digital Guidelines Initiative-2016 Final_rev1.pdf

[2] From Best Practices for Technical Metadata: http://guides.library.illinois.edu/DigitizationMetadata/bestpractices

- file format
- file resolution (pixels per inch)
- dimensions (image dimension or size in inches or centimeters)
- bit-depth (e.g., 8-bit, 16-bit, 24-bit, etc.)
- color space of digital image (e.g., RGB, CMYK, or grayscale)
- information about method of digitization: scanner or digital camera brand, name, and model number
- information about software used to manipulate or compress the image, including the software name and version

Collection of technical metadata will vary according to equipment used and local policy. Inclusion of technical metadata is not required for items in CARLI Digital Collections, but it can be included. For more information, see Summary of Data Dictionary Elements, NISO Standard Z39.87-2006 (R2017) pp. 94-101, which lists mandatory containers and elements for technical metadata: https://www.niso.org/publications/ansiniso-z3987-2006-r2017-data-dictionary-technical-metadata-digital-still-images