

FSU Digital Library Imaging Guidelines

FSU Digital Library Center

Version: 7/24/2019

File Formats

Digital preservation standards recommend that we have the raw or lossless master copies of the item in order to address possible digital preservation necessities such as migration or emulation. Capture of master images in TIFF or RAW formats allows derivatives of JPEGs, PDF, etc. to be created as needed. All images intended for use in the Digital Library will be captured as TIFFs unless otherwise stated.

TIFF

- Save all TIFFs with lossless LZW compression when possible in order to reduce file sizes and retain image quality. Raw files from digital cameras (.CR2 or .IIQ) should be processed and exported as TIFFs. It is important to upload the best quality images into the Digital Library.
- There is no need to manually save reduced-quality derivatives and thumbnail images for display in the Digital Library. These files will be automatically created during the ingest process.
- When digitizing materials for patrons or for other uses outside the Digital Library, it is sometimes preferred to provide JPEGs, PDFs or other derivatives. Patrons' file format preferences will typically be listed in the digital reproduction request form.

JPEG

- JPEG files are not used for ingesting objects into the Digital Library and are usually reserved for internal purposes or patron requests
- JPEG compression should be adjusted based on the user's needs. In general, JPEG compression should be set to no lower than 10. Lower levels will show visible degradation in image quality. Only if low file size is a priority should compression levels lower than 10 be used.
- Take care not to edit and resave the same JPEG image multiple times. Each time the same JPEG file is opened, edited, and re-saved, the image is compressed further and image quality is reduced.

PDF

- When creating PDFs it is important to keep in mind the size and quality of the images being combined. For example, a normal PDF of a 300-page book will be exceedingly large in file size if it is created using high-resolution, full color TIFFs.
 - In many cases it is recommended to create PDFs using compressed grayscale or black & white JPEGs at a 50%-reduced resolution images. Depending on the amount of quality and file size needed in the PDF, these settings can be adjusted accordingly.
 - Many users simply want readable images at small file sizes, as opposed to our standard high-resolution images at large file sizes.
- It is sometimes necessary to create PDFs with editable and searchable textual data. This feature, OCR (Optical Character Recognition), can be applied to images with legible typescripts to embed the text data into the PDF file.
- OCR'd images can also be exported directly to text documents if preferred by the user.
 - Both ABBYY Finereader and Adobe Acrobat Pro X can be used to create PDFs
 - ABBYY Finereader is recommended due to its advanced customization and flexibility with OCR, editing, and saving options.
- OCR is automatically applied to PDFs during the ingest process, so it is unnecessary to perform manual OCR on PDF files intended for the Digital Library.

PNG

- PNG file are only used when creating thumbnails for collections in the Digital Library.
- All collection thumbnails must be saved as 200x160-pixel PNG or JPEG files at 72 PPI.
- Consult the Digital Archivist or Studio Manager for more information on uploading thumbnails to the Digital Library.

File Naming

Listed below are the common file naming conventions used in the Digital Library Center (DLC). These guidelines should be followed as closely as possible to allow for consistent workflows and organization of image files. The filenames of master images created by the DLC should be unique and provide enough information for staff to locate the origin of the images, regardless of where they are found in any directory. If these file naming conventions are not closely followed, individual images separated from their parent directory may become indistinguishable from images in a different directory.

There will sometimes be exceptions in which projects or materials do not exactly conform to the guidelines provided. If an object being digitized is not covered in this section, use best judgement to create a file naming scheme that closely aligns with the DLC's naming conventions below. When in doubt, consult the Project Manager or Digital Archivist.

These naming rules apply only for master images destined for the Digital Library. Alternative naming conventions may apply when digitizing patron requests or miscellaneous imaging jobs.

Naming Master Files:

1. Filename begins with either **"FSU_"**, **"FSU_MSS"**, or **"FSU_HUA_"** (items from Heritage Protocol and University Archives)
 - a. *No other special characters or spaces are allowed*
 - b. Images beginning with "FSU_" refer to items from an FSU entity. Items from Special Collections & Archives should begin with this prefix
 - c. Images beginning with "FSU_HUA_" refer to items from Heritage Protocol and University Archives
2. The next element is the call number or manuscript number followed by an underscore
3. The number sequence is the final element in the filenames. This sequence should be named according to the number of pages in the object as seen in the examples below:
 - a. Manuscripts with 10 to 99 pages should have a two-digit sequence with a leading zero after the manuscript number: **"_01" up to "_99"**
 - i. An additional numerical suffix will be required after each filename if a folder contains several multi-page manuscripts: **"01_01"** and **"01_02"** for a two-page manuscript in the very front of its folder.
 - b. Books with 100 to 999 pages should have a three-digit sequence with leading zeros after the call number: **"_001" up to "_999"**
4. Numbering the files is important so that they can be properly sorted into their correct order. This also helps in proper METS file creation.
 - a. These file name numbers will not match the page numbers of the actual item since you are calling the front cover 001, something a publisher would never do. The file numbers are created so that the images from the front cover to back cover are in numerical order.
5. All letters in the filenames should be **capitalized**

File Naming Examples:

- **Book of Nonsense**

Call# PR4879.L2 N62 1894

350 pages

1. **The correct file name and sequence should be:**

- a. FSU_PR4879L2N621894_001
- b. FSU_PR4879L2N621894_002
- c.
- d. FSU_PR4879L2N621894_349
- e. FSU_PR4879L2N621894_350

- **Captain Hugh Black Papers**

MSS #1964-022

Box 1802

Folder 2

3 handwritten letters

1. **The correct file name and sequence should be:**

- a. FSU_MSS_1964022_B1802_F2_001
- b. FSU_MSS_1964022_B1802_F2_002
- c. FSU_MSS_1964022_B1802_F2_003

[Note: with manuscript collections, the MSS#, Box# and Folder# is the equivalent of a book's call number in the filename. Within a folder, the items should be digitized from first item on top to last in back.]

- **Florida State University Negatives**

Binder date: 1995-49

Photographer's initials: KTM

Negative number: 12a

1. **The correct file name should be**

- a. FSU_HUA_199549KTM_012A
2. Most negatives being scanned at the DLC will be part of Heritage Protocol & University Archives and that will be reflected in the file name as "FSU_HUA_."
 3. The majority of negatives are organized in binders and sleeves. Each sleeve usually has descriptive information which includes the date, title of event and photographer's initials.

4. To minimize the length of the filenames, we will include the date, abbreviated descriptor, photographer's initials, and individual negative number.

Photography of Materials

- Set lens apertures to F8 when photographing most material on the Phase One IQ180 or ATIZ BookDrive. This setting will produce the sharpest images of relatively flat objects. Adjust the shutter speed and ISO as needed to obtain a sharp, properly exposed image.
- Slower shutter speeds will increase the risk of blurred images due to movement of the objects, cameras, or platforms. The BookDrive setup is more at risk of creating motion blur due to the unsteady nature of the cradle.
 - Care must be taken to keep all objects still when photographing objects.
 - Carefully raise and lower the BookDrive's v-platen between shots to avoid capturing blurry images.
- There will be cases in which a higher aperture is needed to increase the depth-of-field of an object. This exception mostly applies to the IQ180 but can also be used with the BookDrive
- Objects that are not completely flat, such as wrinkled pages, peaks of folded material, or curved artifacts, may require higher apertures to 'level' the focus of the image.
- The aperture of the cameras should never be set higher than f16. The focus of the image will be even at this aperture, but will overall much softer than images shot at f8.
- When digitizing large books on the IQ180, it is sometimes beneficial to set the focus point at the center spread (halfway through the book) with a higher aperture.
 - After setting the initial focus and aperture, it is possible to shoot an entire book without re-focusing.
 - However, it is recommended to periodically check the focus and composition of the images while digitizing.
 - The thicker the book, the more often re-focusing will be necessary.
- When photographing material with any lenses in the DLC it is important to leave a wide enough margin around the desired composition to prevent focus fall-off from affecting image quality.
 - Most photographic lenses, including even expensive, high-end lenses, are not as sharp in the corners as in the center.
 - Focus fall-off is more pronounced at lower apertures but levels out at higher apertures.
 - It is sometimes necessary to zoom out further than expected to avoid focus fall-off.

- Image quality will be reduced as the ISO is increased. It is not recommended to raise the ISO past 100 on the IQ180 and 500 on the Canon Mark II 5Ds.

Cropping Images

- When scanning books, all pages should be cropped at the exact same pixel width and height. This helps enhance the reading experience for our users. Front covers, back covers and inserts can be cropped larger than the main pages of books.
- When digitizing most books, set the crop box to just inside the covers and pages so that no background is visible in the image. Make sure to leave a wide enough margin around the page's main content (usually text-blocks or images) so that no valuable information is discarded.
 - Not only is this best practice (as it recreates the page as close to the original as possible), if you crop something too closely, often text will be lost in the display in Islandora.
- Flat objects which are not intended to be displayed in the Book Viewer module, (loose manuscripts, flat art, maps, or images with abnormal edges) have different cropping requirements.
 - When digitizing flat objects, or objects with content extending to the edge of its material, it is recommended to leave a small and even margin around the outside edges of the objects. This gives users more context about the objects while ensuring that important information was not removed from the images when cropped.
 - With an even border extending past the edges of the objects, the background will be visible. It is important to use a clean background so that dust, hair, scratches or other markings are not included in the final image.
 - Use only a solid white, black, or neutral gray background to digitize these objects. All images within a collection or project should use the same background to ensure consistency.
 - The choice of background will usually be decided based on the color or overall brightness of the object being digitized.
 - It is important to show contrast between the object and the background.
 - Do not use a black background when digitizing a black object, for example. The same applies for a white object on a white background.

- Using the same color background for an object of the same, or similar, tone can make it difficult for users to see where the object ends and the background begins.

Imaging Guidelines

- The standards listed below are the minimum imaging recommendations for each object type, however it is understood that exceptions will often occur. Use these to start with and consult any special instructions given to you for a digitization project and then adjust as needed. If a project does not have specific requirements or instructions for image quality, then refer to the imaging recommendations below.
- It is sometimes necessary to exceed the minimum imaging guidelines. Discuss as needed and use best judgment for the item to be presented to the user in the best possible way.
 - Some researchers may require images scanned at extremely high resolutions to scrutinize the finest details of an object.
 - Images will sometimes be used for wall-sized prints, which will require objects to be digitized at resolutions higher than those listed the guidelines.
- Depending on the material being digitized, intended use of the images, equipment limitations, and other factors, it is sometimes impossible to meet these minimum imaging recommendations. In these cases, digitize the objects using the method that will produce the highest quality image under the limiting circumstances. Certain objects should not be digitized by equipment that might cause additional damage.
 - For example, do not digitize a fragile rare book on a flatbed scanner, even if it can produce higher quality images than the Atiz BookDrive.
- Note the optical resolution of the Epson XL10000 is 2400 PPI and no images should be scanned at a higher resolution. Objects scanned at a resolution higher than 2400 PPI will be interpolated and the image quality will not be improved. This will result in a larger file size with no added benefit.
- If you have a unique item not covered in these guidelines, speak with the project leader or consult the imaging guidelines from similar institutions. Several of these guidelines can be found in the [DLC's Digital Imaging Resources](#) folder in Sharepoint.

Digital Imaging Terms

Image Resolution (PPI or DPI)

Resolution is often referred to as dots per inch or DPI. This term is appropriate when describing printer resolution. In common usage, the terms PPI and DPI are used interchangeably; however, since image files are composed of pixels, PPI is a more accurate term for digital images and has been used in this document.

Materials are generally not scanned at a fixed resolution but instead at a fixed number of pixels across the longest side of the image. The appropriate resolution is determined by dividing the desired number of pixels (e.g. 4000) by the number of inches of the long side of the object (e.g. 10" for an 8x10 photo).

Using the example above: $4000 \text{ pixels} / 10 \text{ inches} = 400 \text{ pixels per inch}$

In this case an 8x10" print is scanned at a resolution of 400PPI to create an image with 4000 pixels on the long edge.

Pixel Dimensions

The horizontal and vertical measurements of an image expressed in pixels. The pixel dimensions of an image can be determined by multiplying both the width and the height by the PPI.

Example: an 8" x 10" document scanned at 300 PPI has the pixel dimensions of 2,400 pixels by 3,000 pixels.

Color Depth

The color depth (or bit depth) of an image refers to the number of colors or shades of grey that can be represented by a pixel. At high color depths more data is available to help create accurate images. Digitizing images at a 48-bit color depth (or 16-bit grayscale) can provide more versatility in color and tone corrections but with much larger file sizes. Since storage is limited, the DLC creates color images at 24-bit RGB and black & white images at 8-bit grayscale unless otherwise stated.

Color Correction

All images being digitized for the Digital Library should have color corrections applied via Capture One. When using object-level targets to perform color corrections on the Atiz BookDrive or IQ180, make efforts to avoid unwanted color-casts from untreated walls, ceilings, or even brightly-colored clothing. When light is reflected off of non-neutral colored surfaces, it will often show up on the

items being digitized, including targets. This makes color-corrections difficult since the swatches on the target are no longer accurate.

Color targets should be used at the beginning of every batch of similar objects. It is not always necessary to include object-level targets in every image if the contents are alike. For example, including a color target at the beginning of book digitization will be sufficient for applying color corrections to the entire book via batch-processing. However, if photographing a collection of pottery shards of various shapes and colors, it is necessary to include a target in each image. These targets will mainly be used for color/exposure adjustments of raw files and should be cropped out before exporting the final image.

Image Quality

Image files may be adjusted after photography using image processing software, such as Capture One or Adobe Photoshop. It is preferred to get a good image directly from the camera or scanner and to do as little adjustment after photography. This will help reduce the amount of time spent on post-production.

When editing images do not rely on “auto correct” features. Most automatic color correction tools are set to standard tone and color distribution that is not likely to match your images. This is particularly true for scans of text documents, maps, plans, etc. Before scanning images with the Epson Scan software be sure to turn off any auto-exposure, auto-color, auto-tone, or similar features. Similarly, do not use any auto features in image editing software such as Adobe Photoshop. These features will attempt to ‘guess’ how the image should look and will often provide inaccurate results. It is recommended to match the object as close as possible in the camera or scanner and to utilize targets to manually achieve accurate color and tone if needed.

Digital Imaging Specifications

Because there are far too many formats and document characteristics for comprehensive discussion in these guidelines, the tables below provide scanning recommendations for the most typical document types and photographic formats found in most cultural institutions.

Textual documents, graphic artwork, maps, plans and oversized material	
Format	Recommended Image Parameters
Clean, high-contrast documents with printed type and without significant markings, handwriting or colors.	8-bit grayscale mode - 300 PPI for documents with smallest significant character of 1.0 mm or larger

Mostly refers to clearly printed typescripts or similar material	NOTE: Regardless of the approach used, adjust scan resolution to produce a minimum pixel measurement across the long dimension of 3,000 lines for 8-bit files
Documents with poor legibility, handwritten annotations or other markings, low inherent contrast, staining, fading, or photographs Maps, plans and other oversized material are included in this category.	24-bit RGB mode - 400 PPI for documents with smallest significant character of 1.0 mm or larger NOTE: Regardless of the approach used, adjust scan resolution to produce a minimum pixel measurement across the long dimension of 4,000 lines for 24-bit files
The spatial dimensions of all images should be sized at 100% of the original object: For example, an 8" x 10" object scanned at 400 PPI should be 3200 x 4000 pixels.	

Notes about textual objects, artwork, maps and oversized material:

- Textual documents with clearly printed type and no significant markings or colors can be scanned as 8-bit grayscale images. This will reduce the file sizes of the images and will not discard any important information.
- When digitizing translucent documents such as carbon-copies, thin papers, or papyri, place an opaque white background behind the object to increase readability. The same technique can be used on two-sided documents to prevent ink from showing through the paper
- Maps are diverse, and flexibility should be given when selecting a resolution. It is common that smaller maps have smaller details, but large maps can often have tiny elements. Digitizing large maps at high resolutions may create exceptionally large files that are difficult to handle without benefit unless the details are also small.
- Maps with large details may be adequately digitized at 400 PPI, though maps with very small details may require 600 PPI resolution or higher. Large file sizes may be worth the extra detail clarity for some items.
- Images must be sized and saved at 1:1 scale to the dimensions of the original page. For example, the scale on a map may make sense only if the reader knows the physical dimensions of the original

Photography: Film and Negatives	
Format	Recommended Image Parameters
<p>Format range: 35mm negatives and positive slide film up to medium format negatives</p> <p>Size range: 1.5 x 1 in. to 4 x 5 in.</p>	<p>Pixel Array: 3000 pixels across the long dimension of image area, excluding mounts and borders</p> <p>Resolution: Scan resolution to be calculated from actual image dimensions - approx. 2400 PPI for 35mm originals</p> <p>Dimensions: Sized to match original object</p> <p>Bit Depth: 24-bit RGB mode for color and monochrome Desaturate black & white negatives after scanning in color mode</p>
<p>Format range: Medium format negatives and larger</p> <p>Size range: 4 x 5 in. and larger</p>	<p>Pixel Array: 4000 pixels across the long dimension of image area, excluding mounts and borders</p> <p>Resolution: Scan resolution to be calculated from actual image dimensions - approx. 400 PPI for 8"x10" originals</p> <p>Dimensions: Sized to match original object</p> <p>Bit Depth: 24-bit RGB mode for color and monochrome Desaturate black & white negatives after scanning in color mode</p>

Photographic Prints
Recommended Image Parameters
<p>Pixel Array: 4000 pixels across the long dimension of image area, excluding mounts and borders</p> <p>Resolution: Scan resolution to be calculated from actual image dimensions – approx. 400 PPI for 8"x10" originals and ranging up to the appropriate resolution to produce the desired size file from smaller originals, approx. 570 PPI for 5"x7" and 800 PPI for 4"x5" or 3.5"x5" originals</p> <p>Dimensions: Sized to match original object</p> <p>Bit Depth: 24-bit RGB mode for color and monochrome Desaturate black & white negatives after scanning in color mode</p>
<p>Please note that some photographs, especially oversized prints, may need to be digitized at higher resolutions</p>

Aerial Photography: Film and Negatives

Recommended Image Parameters

Pixel Array:

6000 pixels across the long dimension of image area, excluding mounts and borders

Resolution:

Scan resolution to be calculated from actual image dimensions – approx. 2700 PPI for 70mm originals and ranging down to the appropriate resolution to produce the desired size file from larger originals

Dimensions:

Sized to match original object

Bit Depth:

24-bit RGB mode for color and monochrome

Desaturate black & white negatives after scanning in color mode

Please note that some aerial film and negatives may need to be digitized at higher resolutions to capture the smallest details

Aerial Photography: Prints

Recommended Image Parameters

Pixel Array:

6000 pixels across the long dimension of image area, excluding mounts and borders

Resolution:

Scan resolution to be calculated from actual image dimensions – approx. 600 PPI for originals approx. 8"x10" and ranging up to the appropriate resolution to produce the desired size file from smaller originals, approx. 800 PPI for 5"x7" and 1200 PPI for 4"x5" originals

Dimensions:

Sized to match original object

Bit Depth:

24-bit RGB mode for color and monochrome

Desaturate black & white negatives after scanning in color mode

Please note that some aerial prints may need to be digitized at higher resolutions to capture the smallest details

Photography Film and Prints Notes

- Scans are to be scaled to 100% of their source document dimensions. In all cases, reproduction to original size is ideal. For photographic prints, size measurements for determining appropriate resolution are based on the size of the image area only, excluding

any borders, frames, or mounts. However, in order to show that the entire record has been captured, it is good practice to capture the border area in the master scan file.

- The recommended scanning specifications for photographs support the capture of an appropriate level of detail from the format, and, in general, support the reproduction, at a minimum, of a high-quality 8"x10" print of the photograph.
- Some scanners are not good at inverting the data from negative film. It is sometimes recommended to scan the image as a positive and then through Photoshop or another professional software, invert the image to produce the final image.
- If black-and-white negatives are stained or discolored, it is recommended to make color RGB scans of the negatives and using the channel which minimizes the appearance of the staining/discoloration when viewed as a positive. The image can then be converted to a grayscale image.
- Always wear archival gloves when handling photographic film or prints. Even clean hands and fingers can leave smears or fingerprints on the surfaces of reflective objects. The same applies for framed objects and other objects enclosed in glass.
- Aerial photographs have a lot of fine detail, often require a high degree of enlargement, and may require a higher degree of precision regarding the dimensional accuracy of the scans (compared to textual documents or other types of photographs)

Objects and Artifacts
Recommended Image Parameters
<p>Pixel Array: 3000 pixels across the long dimension of image area. These parameters will vary depending on the objects being photographed and the equipment being used.</p> <p>10 to 16 megapixel 24-bit RGB mode image</p> <p>If scanning photographic copies of objects and artifacts, see recommended requirements in the appropriate photo charts above.</p>

Resources

NARA Digitization Guidelines

<http://www.archives.gov/preservation/technical/guidelines.pdf>

FADGI Digitization Guidelines

http://www.digitizationguidelines.gov/guidelines/FADGI%20Federal%20%20Agencies%20Digitization%20Guidelines%20Initiative-2016%20Final_rev1.pdf

Proposed Digital Imaging Standards and Best Practices - Indiana Memory and LSTA Digitization
Projects

http://www.in.gov/library/files/dig_imgst.pdf