

Alabama Mosaic F.A.Q.

SCANNING

Updated: March 2009

Who will scan materials for our digital collection?

Each repository is responsible for creating its own digital collection. Some repositories have “out-sourced” their scanning, but AlabamaMosaic can not suggest vendors for such services.

What equipment will I need for a digital scanning project?

You will need a computer and a scanner. You may already have a computer that you can use. If you plan to add a new one, select the highest capacity you can afford. You should also add CD-R and CD-ROM drives for creating and copying CDs.

Most sites use flatbed scanners and these vary depending on the manufacturer. Most of today’s basic scanners provide adequate scanning quality for AlabamaMosaic. A large “bed” (at least 13x17) is desirable. Choose a scanner that offers a minimum of 600 Dots Per Inch (DPI) instead of “interpolated resolution” which means the scanner has built-in ability to enhance an image by changing the resolution.

Once you have scanned an item, you will need software to enhance it for public presentation. Adobe Photoshop™ is the most widely used commercial software for enhancing digital images. There are also free photo-editing products such as GIMP (www.gimp.org) and Paint.NET (www.getpaint.net). Both of these products are highly regarded open source choices for photo editing.

What scanning settings should we use?

AlabamaMosaic recommends using the scanning resolution that is most appropriate for the each item. Some materials are easier to digitize than others. The easiest are single-page, flat works printed on paper. Better scanning results are obtained if the item has high contrast print and uses a clean “crisp” type font. Copies on “onion skin” paper and handwritten documents may require many adjustments to obtain legible digital copies.

Your initial scan should be saved as a high resolution TIFF. This will become your master file to archive and should not be altered. Your master file can be copied to generate the smaller JPEG files that you contribute to AlabamaMosaic.

Some commonly used settings for an initial scan are:

- Photographs scanned at 300dpi
- Handwritten documents scanned at 400dpi
- Artwork scanned at 600dpi
- Maps scanned at 800dpi

This table is from "Scanning Standards and Guidelines," on the Ohio Memory website <http://www.ohiomemory.org/om/scanningreq.html>

Minimum Scanning Resolutions

Type of Original	Minimum Resolution and Pixel Depth
Published text and line art	300 dpi (dots per inch) bitonal or grayscale TIFF.
Hand-written materials	400 dpi grayscale or color TIFF.
35 mm negative and slide film	600 dpi grayscale or color TIFF.
Photographic Prints	Minimum 300 dpi, color or grayscale, depending on the source. Scanning from negatives may be preferable.
Oversized items or museum objects	It may be better to re-photograph the work or to capture the work with a high-end digital camera..

Can you help us determine the correct settings?

When your application to contribute to AlabamaMosaic has been accepted, you will be asked to send several sample images for review and advice. Do not scan a large number of items until you have received comments from this review.

Can I set or change the settings on my scanner?

Your scanner will arrive with the manufacturer's instructions for getting it operational. During the set-up, you will choose your "default" settings. Today's scanners include software that allows you to set scanning "instructions" for every item scanned. Each company designs its own software and interface, so instructions vary from machine to machine. You will need to read the manual and determine how to select Dots Per Inch (dpi— which may be labeled ppi or lpi). Your minimum setting for clean, high contrast type-font print on paper should be 300 dpi. When in doubt about a setting, choose the one that will achieve the highest quality scan. You will also have to choose how to save the digital file. AlabamaMosaic recommends that you save your initial scan as an unaltered TIFF so you can use this file to generate additional digital copies for other purposes.

When I create a TIFF for a master file, how should I archive it?

Archival files should be stored in a facility separate from the one containing the original. Some repositories write their archival files to CDs. If you use CDs, they should be stored offsite in archival conditions (controlled access, light,

temperature, and humidity). You can also contact the [Alabama Digital Preservation Network](#) (ADPNet) to inquire about using its archival storage facilities.

What formats are commonly derived from master TIFF files?

The most common practice for AlabamaMosaic contributors is to save the presentation file as a JPEG at 72 dpi resolution. Other commonly derived formats include:

GIF (Graphics Interchange Format) is generally only used for creating thumbnails and 1 bit bitonal (black & white) images.

JPEG (Joint Photographic Experts Group) is widely used for creating medium and high resolution images for Web delivery.

PDF (Portable Document Format) is commonly used for printed text and requires the repository to use [Adobe Acrobat](#)[™] software to apply the format. Acrobat.com provides an online PDF converter so you can easily create PDF documents from your existing files in Microsoft Word, Excel, PowerPoint, and many other applications. Users must download and install the free [Adobe Reader](#) to view the document.

What is the difference between a TIFF and a JPEG/JPG?

A TIFF is an uncompressed image and is the best format for your master archival file. A TIFF image preserves more details from the original item. To create smaller files that will load faster, AlabamaMosaic stores a compressed file for presentation online. The most common compressed file format is JPEG/JPG. This smaller file requires less server space to store and less time to download for a user to view. This smaller file is fine for online viewing.

Can I use JPEG 2000 instead of JPEG/JPG?

Yes. JPEG 2000 is a published [International Organization for Standards \(ISO\) standard](#) (ISO 15444). It provides better image quality at smaller file sizes than JPEG and its compressed images retain an image without much distortion or loss. For more information about JPEG and JPEG 2000, visit the [Joint Photographic Experts Group and Joint Bi-Level Image Experts Group](#) website. You may also want to read the Technology Watch Report published by the Digital Preservation Coalition, [JPEG 2000 – a Practical Digital Preservation Standard?](#)

Can I scan and save the originals as JPEGs without saving TIFFs?

AlabamaMosaic **STRONGLY** recommends that you save the initial scan as an unaltered TIFF. However, this is your own digital file, and you determine what file format you will use. Alabama's rich collections will take years to scan, and everyone wants to avoid re-scanning items. You should organize your work to handle and scan an item only once. Saving a TIFF will give you the greatest capability to use the image for other purposes without subsequent handling of the original for additional scanning. .

What does “unaltered” mean?

The look of an original historical document is an important part of its history. The initial master TIFF should retain, as far as possible, the appearance of the original document. After the document has been scanned, you should not alter it (change the appearance) by disguising imperfections such as holes, tears, or streaks or by digitally removing notes or other content added to the document. You should not crop an image or eliminate parts such as background scenery. Once you save your master TIFF, you make a copy to create your presentation copy. Your presentation copy should retain the appearance of the original, but acceptable alterations include:

- cropping an image if it is surrounded by a lot of blank white space
- inserting the name of the contributing library or a copyright notice at the bottom of the image
- adding a “watermark” overlay that names the owning repository
- removing evidence of a tear line

Can I use a digital camera to make images for AlabamaMosaic?

Most consumer digital camera images do not offer sufficient image quality for digital collection building and most do not create a TIFF image. Only high quality professional cameras, such as PhaseOne or BetterLight, are adequate. This is your digital collection, and you may decide to use a digital camera that does not create TIFF formats if you have no other option.

What if my document is too large for my scanner?

AlabamaMosaic can advise you about possible sources for scanning oversized materials such as posters. These may be fee-based, and charges vary by the amount of handling that will be required.

What do bitonal, grayscale, and color mean for scanning?

A **bitonal scan** can represent only two tones, black and white. It is best used only for high contrast black type font on white paper because only these two colors are represented. There are no shades of gray or other colors.

A **grayscale scan** can represent black, white, and a range of gray shades. As the number of shades in the range of gray increases, the quality of the scanned image increases. This also increases the size of the file. Even black and white photographs have shades of gray, so you should use grayscale to scan black and white photographs.

A **color scan** can represent a wide range of colors. These are very large TIFF files because they preserve color resolution up to 48 bits (a 16-bit field each for Red, Green, and Blue), either as full RGB color or in a 64k-color palette.

Some items we want to scan are fragile and rare. Should we scan them anyway?

Many repositories are scanning fragile and rare historical materials so the originals can be stored in closed archives. This makes the intellectual content available without handling and endangering the physical item. Your first priority is security for the item. Your workflow should insure that the every valuable item

is maintained in a secure environment and cannot “disappear” or be harmed. A second and equal priority is to do nothing that will endanger the material during the digitization workflow and actual scanning.

If I scan a lot of items, how should I name the individual digital files?

The scanner will give each file a generic name but each repository should determine how best to name its digital files when saving them. Some repositories use an already assigned “identifier” if the item is part of collection that names or numbers each item. Some use a short identifier for the collection and then the number. Other AlabamaMosaic contributors can advise you about file naming procedures that work for their collections. Above all: Be consistent.

Are there other resources for information about scanning?

The AlabamaMosaic website includes a [Directory of AlabamaMosaic Trainers](#) on this website. These are from AlabamaMosaic who volunteer to advise new contributors. Cornell University maintains an online digital imaging tutorial <http://www.library.cornell.edu/preservation/tutorial/> derived from the out-of print book *Moving Theory into Practice: Digital Imaging for Libraries and Archives* (Anne R. Kenney and Oyua Y Rieger. Research Libraries Group, 2000)

Many digital programs post their scanning guidelines online:

Library of Congress. *Technical Standards for Digital Conversion Of Text and Graphic Materials*. 28 pages.

<http://rs6.loc.gov/ammem/about/techStandards.pdf>

Digital Library Production Guide, Kentuckiana Digital Library

<http://kdl.kyvl.org/guidelines.html>

North Carolina ECHO (Exploring Cultural Heritage Online). Digitization Guidelines <http://www.ncecho.org/dig/digguidelines.shtml>

BCR’s CDP Digital Imaging Best Practices, Version 2.0, June 2008. 71 pages. <http://www.bcr.org/cdp/best/digital-imaging-bp.pdf>

These publications can be read online:

Sitts, Maxine K., ed. *Handbook for Digital Projects: A Management Tool for Preservation and Access*, Northeast Document Conservation Center, 2000 (out of print) Available online (182 pages) at:

<http://www.nedcc.org/resources/digitalhandbook/dman.pdf>

Humanities Advanced Technology and Information Institute (HATII), University of Glasgow, and the National Initiative for a Networked Cultural Heritage (NINCH). *The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Material* (see chapter six: Capture and Management of Images) <http://www.nyu.edu/its/humanities/ninchguide/>