

## SCAN TECH STUDIO WORKSHOP SERIES: HANDLING AND DIGITIZATION

### SCANNING FOR AN INSTITUTION VERSUS SCANNING FOR PERSONAL PROJECTS\*

	Institution	Personal
<i>Purpose?</i>	Long-term archival preservation and access for a wide range of collections materials	Safekeeping important and sentimental items like photos, documents, small books, home movies, and more
<i>Guidelines?</i>	Federal Agencies Digital Guidelines Initiative (FADGI) – set of strict standards used by libraries, archives, and other institutions that outline best practices; can be expensive and time consuming to execute	General guidelines for good lighting, image quality, and storage of digital files that can be adjusted depending upon the equipment and budget at your disposal
<i>Equipment?</i>	Specialized equipment to handle a wide variety of formats and heavy use for long-term digitization projects; outsourcing to vendors who are equipped to handle projects with very specific digitization specs	At home with flatbed scanner, cellphone camera, personal camera, etc. Projects may also be outsourced to companies like Legacybox, ScanMyPhotos, iMemories, Iron Mountain, etc.
<i>Processing?</i>	FADGI has best practices for image/audio editing and retouching; in general, does not permit adjustments to master files but can make certain adjustments for access/derivative files; save all non-access files as TIFF	You can do whatever you want – they're your files! Correct red eye, adjust brightness and color levels, apply various processing techniques, etc. Best to also keep an unedited TIFF or JPG copy of the image for future use
<i>Storage?</i>	Sophisticated organization and intentional redundancy to create several copies for safe keeping and retrieval of files for perpetuity	Best to keep at least two digital copies of files such as one on an external hard drive and a backup on a Cloud-based service

\*with a wide range of reasons for digitization and methods of scanning in between

## HANDLING AND STORAGE OF PHYSICAL MATERIAL

### IN GENERAL:

- Visually inspect the material for any obvious signs of deterioration or concern like detached pages, brittle and flaking paper, warping of material, mold growth, etc.
- Use clean, dry hands and exercise caution when moving material
- Handle fragile materials as little as possible and use proper supports to minimize stress on the material
- Stop immediately if you notice any damage occurring during scanning

---

### PHOTOGRAPHS, SLIDES, AND FILM

- Use nitrile gloves when handling archival photographic emulsion materials – handle from the edges and avoid touching the emulsion surface of an image
- Do not expose to light for longer than necessary to prevent fading and damage

---

### STORAGE

- Keep away from direct light sources (especially sunlight), extreme temperature changes, high humidity, pollutants, moisture, bugs and other critters
  - AKA: don't keep your family photos and great-grandma's Bible in the semi-finished basement, an outdoor shed, the garage, or the attic
- For items that are fragile or vulnerable to damage, best practice is to store them in acid-free paper mats/boxes/enclosures and chemically-stable polyethylene/polyester sleeves
- Books can be stored upright on clean shelves away from direct sources of light

## DIGITIZATION OF MATERIALS

See the Glossary at the end of this document for a detailed explanation of other terms.

### QUICK OVERVIEW OF IMAGE QUALITY

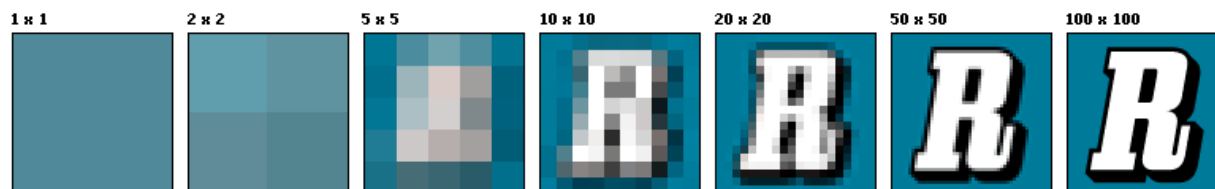
The most basic elements of image quality depend upon **resolution**, **color depth**, and **the format of the image**.

---

## RESOLUTION

The resolution of an image is the level of detail that an image holds; it is the quantity of pixels within an image. This is measured in PPI, or pixels per inch.

In the example below, the number of pixels within a given space determines the amount of detail that can be rendered. A low pixel count is blocky and poor quality whereas a higher pixel count leads to a crisp, higher quality image.




---

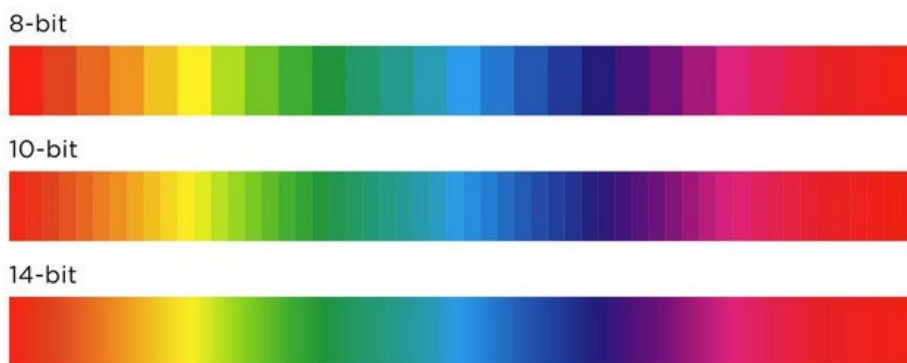
## COLOR DEPTH / BIT DEPTH

Color depth, also called bit depth, refers to the amount of color information that is stored in an image. Transitions between color and tone are much smoother in an image with a high bit-depth than a lower bit-depth.

The most common bit depths used in digitization are 8-bit and 16-bit.

- 8-bit can contain up to 16 million colors; ideal for web display, smaller prints
  - Suitable for most digitization projects with printed and handwritten documents
- 16-bit can contain up to 28 trillion colors; ideal for photo editing and printing
  - Use when digitizing photographs, film, and fine art prints

This graphic demonstrates the difference among the available color information in an 8-bit, 10-bit, and 14-bit color depth. The greater the bit depth, the smoother and more detailed color and tone gradations will be in an image.



A low bit depth may cause unsightly banding of colors, particularly in compressed and heavily edited images.

---

## SAVING YOUR IMAGE – RAW, JPG, OR TIFF?

Depending on the camera used, your images will either be shot in RAW or JPG format. The advantage of shooting in RAW format is that the highest level of detail is retained in the image; this image can then be converted into many other image types while still retaining the most amount of detail possible. Shooting in JPG format still results in a good quality image, but the very fine details will be lost due to compression of the image. For most non-professional digitization projects, this loss of detail should not be an issue.

It is not advantageous to keep your images in a RAW format. These files are ideal for editing and converting to another image type as they are extremely large and often require specific, proprietary software to view or edit them.

You should save your photos as either a TIFF or a JPG file. There are pros and cons to both and you should choose the format that best fits your specific needs.

	<b>JPG</b>	<b>TIFF</b>	<b>RAW</b>
<b>Digitization project</b>	Good for printed and handwritten documents/books; okay for photos/negatives/slides when other options are not available	Excellent for all projects; recommended for photos, slides, negatives, film, art prints, and other highly detailed items	Professional-level for all digitization projects
<b>Size</b>	Significantly smaller file size	Quite large due to high data content	Massive; not good for long-term storage
<b>Compression</b>	Lossy compression – image data is thrown out that cannot be recovered	Can use lossless compression – image data is retained	Uncompressed – retains all image data from time of capture
<b>Editing</b>	Not ideal; each additional edit will eliminate more data from the image and cause "compression artifacts"	Very good; can create further edits without introducing visual flaws in the image	Best; able to make highly specific edits without issue
<b>Color depth</b>	Limited to 8-bit	8-bit or 16-bit	Usually between 12 and 16 bit
<b>Sharing capabilities</b>	Small file size makes sharing and transfer of files quite easy; can be sent through e-mail	Can be difficult to share, especially if transferring over slower internet networks or in large batches	Not good – file size is massive and often requires proprietary software to even view the image

<b>Equipment and software</b>	Extremely common shooting option for digital cameras and cell phones; compatible with most devices and software	Common format, especially for photography; may not be compatible with the most basic of image viewers; most cameras do not shoot in TIFF but it is a common conversion from RAW	Most common in traditional digital cameras; iPhone 12 Pro and later Pro models can shoot in RAW format; typically requires proprietary software
-------------------------------	---	---	---

### BEFORE BEGINNING A DIGITIZATION PROJECT, CONSIDER THE FOLLOWING:

- Technical capabilities of the available scanning equipment
  - Color depth (bit depth)
  - Optical resolution and sensor
  - Capture format (RAW vs. JPG files)
- Specifications required for the digitization of an item
  - Processing/storage format (TIFF vs JPG)
  - End goal? Digital reproduction, print reproduction? OCR?
    - Image resolution (PPI)
    - Color depth (bit depth)
- Physical condition of item and format of scanner
  - Flatbed scanners are good for loose documents, photographs, slides, very thin bound materials
  - Cradle scanners to support bound materials

Specifications for general, all-purpose scanning with a compromise between quality and size of files:

- 8-bit color
- 250-300 ppi
- TIFF or JPG

If you are planning to run optical character recognition (OCR) software on a document with clearly printed text, a minimum of 300 ppi is recommended.

For digitizing highly detailed materials or when you intend to produce a print copy, scanning at a minimum of 600 ppi is recommended.

## GLOSSARY

### **Access file**

Also called *service, delivery, viewing, output, or derivative* files. A version of the original archival master file that has been altered and/or optimized for a specific use (e.g., to provide end-user access; high quality reproduction; processing optical character recognition)

### **Adobe RGB (1998)**

A red-green-blue color space developed to display on computer monitors most of the colors of CMYK color printers; it is significantly larger than the sRGB color space, particularly for cyan and green regions

### **Archival master file**

A file that represents the best copy produced through digitization, with best defined as meeting the objectives of a particular project or program; archival master files are digital content intended for long-term archival use without loss of essential features and are meant to be maintained in a preservation-oriented data management system

### **Artifact**

A visual distortion in an image often caused by loss of image data; can appear as fuzziness, rings, banding of colors, discoloration, pixilation, and more

### **Bit**

The most basic unit of information in computers/digital communications; has one of two possible values usually represented as either "1" or "0" but may also use true/false, yes/no, on/off, +/-

### **Bit depth**

Also referred to as *color depth*. Number of bits used for each color component of a single pixel. Archival digitization uses 8-bit and 16-bit color. Higher bit depth equates to increased capacity of color variants.

### **Color space**

Specific organization of colors that supports reproducible representations of color in combination with color profiling supported by various devices; can be helpful for describing or understanding color capabilities of a specific device or file. Examples include Adobe RGB (1998), sRGB, and ProPhoto RGB (also see *gamut*)

### **Derivative**

See *access file*. Derivative files are accessed frequently by end-users

**DPI**

Dots per inch. A measurement for printer resolution that is independent of image pixel resolution (see *PPI*). The higher the number means the more dots per inch that a printed image will have, the more detailed a printed image will be. Scanners use DPI to indicate the number of points of data that the scanning sensor will pick up from the image

**Gamut**

The specific range of colors that can be displayed by a given *color space*. A wider gamut range means more colors can be displayed than a narrower gamut.

**JPG/JPEG**

An image file named for the group that defined its standard format (Joint Photographic Experts Group). Uses lossy compression which results in a smaller file size than TIFF and RAW files but contains significantly less data within the image, sacrificing image quality for more manageable file sizes

**Pixel**

The smallest unit of information in an image structure; an abbreviation of *picture element*.

**PPI**

Pixels per inch. A measurement of image resolution or a descriptor of the resolution capabilities of a scanner/camera. More pixels in an image means more detail and variation in color to create depth, clarity, and sharpness.

**Production master file**

Files that are produced using one or more archival master files

**ProPhoto RGB**

Highly detailed color space developed by Kodak for photography purposes

**RAW file**

An uncompressed, unprocessed image file created by the camera at the time of capture. It contains a high amount of detail and can be edited, compressed, and/or converted into another file format without compromising image quality. RAW files are often proprietary and require specific software to view and edit

**Resolution**

The level of detail that an image holds; also, an imaging system's ability to disseminate finely spaced detail. Resolution refers to the quantity of pixels in an image. Images with high resolution are high quality but are also large in file size; low resolution images have lower quality but smaller file sizes

**sRGB**

Standard RGB color space created by HP and Microsoft for use on monitors, printers, and the Internet. Can be directly displayed on typical computer monitors

**TIFF**

Tag Image File Format. A widely supported file format for images with lossless compression meaning that image quality is retained while reducing file size; however, the size of a TIFF file is still often quite large. Ideal for high-resolution scans and photographs



## CITATIONS AND FURTHER READING

---

Apple. "About Apple ProRAW." Published 28 September 2022. <https://support.apple.com/en-us/HT211965>

Adobe. "JPEG vs. TIFF." Accessed 19 February 2024. <https://www.adobe.com/in/creativecloud/file-types/image/comparison/jpeg-vs-tiff.html>

Adobe. "Our Guide to PPI (Pixels per Inch) and Pixel Density." Accessed 19 February 2024. <https://www.adobe.com/uk/creativecloud/photography/discover/pixels-per-inch-ppi-resolution.html>

Adobe. "RAW files." Accessed 19 February 2024. <https://www.adobe.com/creativecloud/file-types/image/raw.html>

Brittany Centeno's super nifty *Staff Guidelines for the Identification, Handling, and Care of Negatives and Slides* handout from 2023.

Federal Agencies Digital Guidelines Initiative. *Technical Guidelines for the Still Image Digitization of Cultural Heritage Materials*. 3<sup>rd</sup> ed., May 2023. [https://www.digitizationguidelines.gov/guidelines/FADGI%20Technical%20Guidelines%20for%20Digitizing%20Cultural%20Heritage%20Materials\\_3rd%20Edition\\_05092023.pdf](https://www.digitizationguidelines.gov/guidelines/FADGI%20Technical%20Guidelines%20for%20Digitizing%20Cultural%20Heritage%20Materials_3rd%20Edition_05092023.pdf)

Mansurov, Nasim. "Raw vs JPEG: Which Image Format is Better and Why." *Photography Life*. Last modified 18 September 2023. <https://photographylife.com/raw-vs-jpeg>

Nicholson, Angela. "Image formats (RAW, JPEG, HEIF) and compression." *Canon*. Accessed 19 February 2024. <https://en.canon-cna.com/pro/infobank/image-file-types/>

US National Archives. "How to Preserve Family Archives (papers and photographs)." Accessed 19 February 2024. <https://www.archives.gov/preservation/family-archives>

US National Archives. "When We Say We Want 'Resolution': DPI and PPI Explained." Accessed 19 February 2024. <https://records-express.blogs.archives.gov/2023/08/22/when-we-say-we-want-resolution-dpi-and-ppi-explained/>

Wikipedia. "Comparison of graphics file formats." Last modified 19 December 2023. [https://en.wikipedia.org/wiki/Comparison\\_of\\_graphics\\_file\\_formats](https://en.wikipedia.org/wiki/Comparison_of_graphics_file_formats)

Wikipedia. "RAW Image Format." Last modified 19 February 2024. [https://en.wikipedia.org/wiki/Raw\\_image\\_format](https://en.wikipedia.org/wiki/Raw_image_format)