# Photo Digitization

DECISIONS FOR PLANNING YOUR WORKFLOW

This presentation is on standards and best practices related to the digitization of photographs.

# General Outline

## Selection
◦ Which version of the photograph to digitize?

## File Specifications
◦ Color

## Metadata
◦ Defining metadata fields

This will cover three basic topics:
Selection issues specific to photographs
File specifications for digital images
Defining metadata fields.

## Selection: Background

Format designed for mass production

One image -> Potentially many copies

Raises questions for selection: which copy to chose?

Can also raise question for metadata: relating copies of same image

One issue that you may run into when digitizing photographs is that you have different versions of the image. You may have the negative, several prints, and objects containing the image, like flyers and cards. When digitizing, which copy should you choose?

# Selection

First question: What can I digitize?
◦ Prints? Negatives?
◦ Oversized formats?

Second question: What should I digitize?
◦ Authenticity
◦ Versatility

Third question: What is the context of the photograph?
◦ Inscriptions on or around
◦ Social context

The first question to ask is what format can you digitize? Do you have a scanner that can scan negatives? What size formats can you scan?

The second question is to ask which copy is the most authentic and which will provide the best quality? Remember that a high quality scan from a high quality image is versatile and can be applied to many different uses. Generally speaking, the negative is the most authentic copy and the copy from which you can extract the most information. A preservation print or negative are also great choices. If you choose a print, an uncropped, untouched print best meets the qualities we want in the archival master.

Other questions to ask is there any information one of the copies that's useful for contextualizing the image? Also, is there a social context for the image that is worth capturing?

# Selection: Authenticity and Versatility

Lean toward digitizing negatives
- It was actually in the camera when the photograph was taken
- Provides the most information
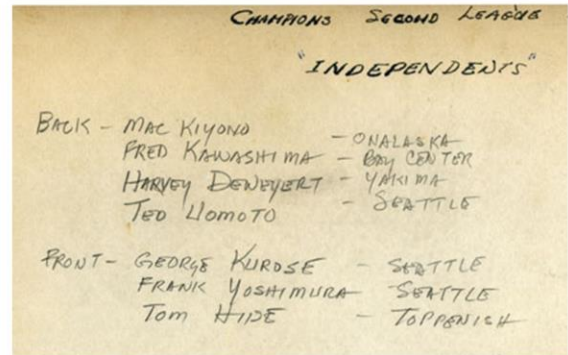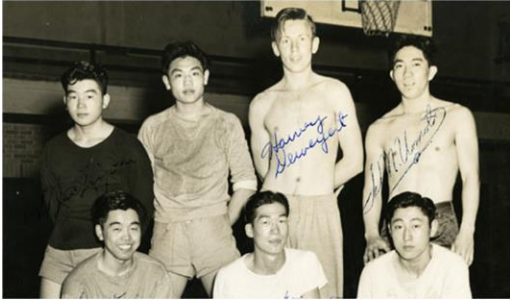- Is the highest quality image
- Preserves better

For the purposes of authenticity and versatility, negatives are often our best choice.

An example of a print with added information. Even if I had the negative, I would be compelled to digitize this print in addition to or maybe even instead of the negative. Sometimes, the meaning of an image is impacted by other objects around it. One obvious example is a photograph in a scrapbook. In this case, we may feel compelled to digitize a less than ideal copy of an image because its meaning is enhanced by other objects on the scrapbook page.
Some versions of an image have special social contexts (see next page).

# Social Context



The original image is on the left and an air brushed image is on the right. One of the men removed from the image fell out of favor with the ruling party and his presence was erased from the event documented in the photograph. If all we had was the airbrushed negative and a print or some other document with the original image, we may feel compelled to digitize both because of the unique social context surrounding the image.

# Other Notes about Format

Glass Formats
◦ Do not set emulsion on scanner glass

Direct Positives
◦ Do not remove direct positives from frame

A preservation note:
Do not place the emulsion of glass negatives on scanner glass for fear of scratching or otherwise damaging the image. Also placing glass directly on glass (e.g., glass slides) can produce Newton rings, which is a distracting visual effect that impacts the digitized image. You can purchase anti-Newton ring glass to mediate this effect or elevate the glass image slightly with the use of thin objects (e.g. cardboard tabs) that won't damage the scanner glass or original image.
Do not remove direct positives like tintypes and daguerreotypes from their frames. These provide crucial physical support for the items. We also want to avoid disrupting the environment of the original if at all possible.

# File Specifications

Resolution

Color Bit-Depth

Tonal Range and Color Space

These are a few specifications we'll want to set specific measurements for before digitizing.

Image on the left is a higher resolution than the image on the right.
For archival masters, resolution standards are 4000-8000 pixels along the long edge
(e.g., an 8x10 photo should be digitized at a resolution of 400-800 ppi. 400 ppi x 10 in
= 4000 pixels).
For access copies, a resolution of 150 ppi generally makes sense for prints. For film,
150-200 ppi or 600 pixels along the long edge.
High resolution images take up more storage space. Digitize at a resolution that
makes sense. If the image clarity doesn't improve beyond 4000 pixels, there is no
reason to apply a higher resolution.
High resolution images can be expanded to accommodate larger formats and they
can be more effectively cropped if you want to use only a portion of the photograph
for a specific purpose. Low resolution images can severely impact your options for
magnifying and cropping.

The image on the left is 48-bit color. The image on the right is 12-bit. You can see how much detail is lost when there isn't enough color represented.

Color is generally represented along three channels: Red, Green, and Blue (RGB). The number of bits you dedicate to each channel impacts the number of shades you can represent in each channel and the number of colors you can create by mixing these different shades.

The way bit-depth is measured is by multiplying the number of bits per channel by the total number of channels. If you are digitizing in RGB at 8 bits per channel you are digitizing at a color bit-depth of 24 (8 bits x 3 channels = 24).

At a color bit-depth of 24, you can represent over 16 million colors, which is consider true color in that the human can differentiate approximately this many colors.

For archival masters, it is recommended that you digitize at a color bit depth of 48 (16 bits per channel) in color and 16 in grayscale.

For access copies, 24 and 8 work fine.

# General Standards (CDP Digital Imaging BPs)

Standards employed depend upon format (e.g., print, 35mm, etc.)

In general, for archival masters
- File format: TIFF
- Bit-depth: 48-bit for color, 16-bit for grayscale
- Resolution: 4000-8000 pixels along long edge

Specific uses may require higher resolutions
- There are limits to resolution

Limitations
- Storage
- Software and scanner limitations

General rule of thumb is to digitize at a high level of quality so that you only have to digitize once, but that means….

Your main limitations when it comes to scanning a high-quality image are storage because high resolution, high bit-depth photos take up more space than low-res, low bit-depth images.

Generally speaking, most scanners and software will allow you to digitize according to best practices.

I will also mention that high-resolution images take more time to scan, but any gains based on time will be lost if you have to rescan because higher-res images are needed.

# Tonal Range and Color Space

Color Space = Range of colors that can be represented

Tonal Range = The range of colors represented in an image

Both can be set upon scanning



Synthesized Macbeth ColorChecker
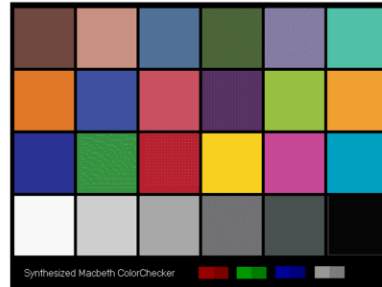
Color space is a measurement of the range of colors that can be captured by a scanner, camera, or other input device, or the range of colors that can be represented within an editing program like Photoshop. For our purposes, it's important for calibrating scanners. You want to calibrate in as broad a space as possible. Adobe RGB is an example of a broad space.

When you capture an image from a scanner, you can adjust the tonal range through the histogram tool. According to the FADGI Technical Guidelines for Digitizing Cultural Heritage Materials, you may want to adjust the extremes of your tonal range (pure white and pure black), as well as the middle neutral point. For archival masters, this adjustment should only be done within strictly specified standards. In addition, this process is time consuming and may not be appropriate for every project.

# Metadata

Which fields contain fixed vocabularies?

What standards do you want to place upon fields that are more variable?

Existing standards, vocabularies, and thesauri
◦ CDP Dublin Core Best Practices
◦ Art and Architecture Thesaurus
◦ Thesaurus of Graphical Materials
◦ DCMI Type (http://dublincore.org/documents/dcmi-type-vocabulary/#H7)

When creating a metadata plan, it is useful to identify which fields should have a fixed, limited vocabulary and which need to be more flexible.
If you're using an existing metadata schema like Dublin Core, you will want to at least consult the published best practices, even if you choose to apply the schema differently.
For describing visual images, there are pre-defined vocabularies (like the Art and Architecture Thesaurus) that can be useful for fields like keyword, subject, type, and format. While these vocabularies may not be completely relevant to your needs, they may be worth consulting so that you don't have to invent your vocabularies from scratch.
The CDP Dublin Core Best Practices have set some useful guidelines for forming titles and descriptions and other categories in a standardize way that may be applicable to your situation.