# DIGITIZATION ON A DIME

## Part 1 – Standards and Baseline Considerations

Isaiah Beard • Digital Data Curator, Rutgers University Libraries

NJLA Conference – June 1, 2018

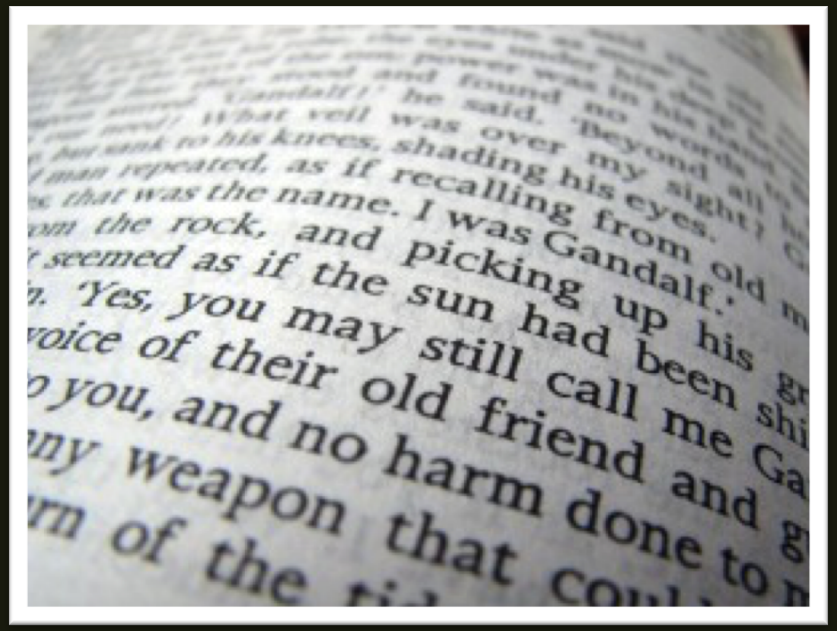
# Topics Covered

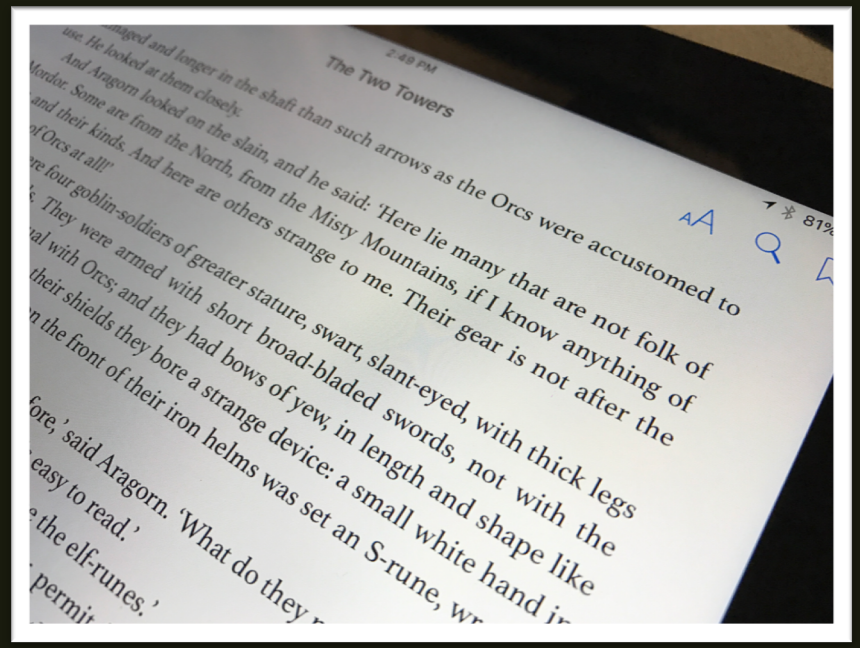
- Methodology: Terms, philosophy and purpose when digitizing
  - *Basic Elements*
  - *Considerations*
  - *Workflow*

- Standards
  - *Resolution*
  - *Color Spaces*
  - *Taking the source and purpose into account*

- Equipment
  - *Standard photo scanners*
  - *Negatives (Film, Plate Glass)*
  - *Large Format Items*

# PHILOSOPHY AND TERMINOLOGY



# The Goal: Digitize to Preserve

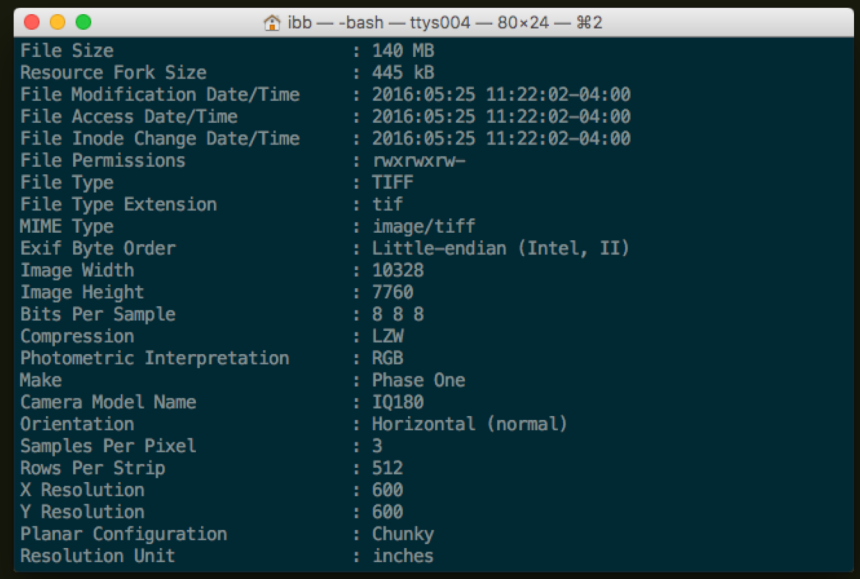
- Conserve brittle, old, or delicate documents
- Digitize only once: capture as much data as possible.
- Reproduce and share freely with the public when rights allow
- Increase object’s usability, accessibility
- Transcend the object’s physical limitations

# Philosophy

- Having minimum standards for preservation ensure:
  - *We can prepare for future technology advances*
  - *We can keep up with new displays and web standards*
  - *We can prepare for migrations if common formats evolve and change*

# STANDARDS

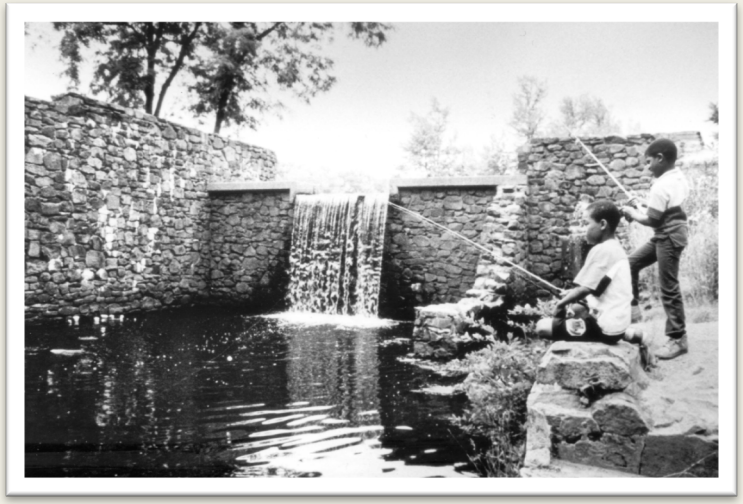

PLATE. 8.
Phila. & Trenton R.R.
EDDINGTON
Eddington Sta.
Est. of Dr. J. H. Schenk
150 A
J. C. Vandegrift
160 A
BENSALEM
George Jackson
115 A.
Mrs. Price
102 A
Vandegrift & Thomas
30 A
C. S. Vandegrift
19 A.
H. C. Fox
75 A
THE POINT FARM
Eddington Wharf
Dunks Ferry
DELAWARE RIVER


ibb — -bash — ttys004 — 80×24 — ⌘2
File Size : 140 MB
Resource Fork Size : 445 kB
File Modification Date/Time : 2016:05:25 11:22:02-04:00
File Access Date/Time : 2016:05:25 11:22:02-04:00
File Inode Change Date/Time : 2016:05:25 11:22:02-04:00
File Permissions : rwxrwxrw-
File Type : TIFF
File Type Extension : tif
MIME Type : image/tiff
Exif Byte Order : Little-endian (Intel, II)
Image Width : 10328
Image Height : 7760
Bits Per Sample : 8 8 8
Compression : LZW
Photometric Interpretation : RGB
Make : Phase One
Camera Model Name : IQ180
Orientation : Horizontal (normal)
Samples Per Pixel : 3
Rows Per Strip : 512
X Resolution : 600
Y Resolution : 600
Planar Configuration : Chunky
Resolution Unit : inches

# Resource/object considerations

- What type of item are you digitizing?

Color photograph or document



Grayscale photograph or document


Black and White document with halftones

# Resource/object considerations

- Some collection owners might want to show off the patina, yellowing or sepia of a photograph or document.
- To do this, they will scan in full color, even if the original material is black and white or grayscale.
- Less image information is available as you go from color, to grayscale, to pure black and white. Choose wisely.

# Digitization Standards: File Format

- Start with a Preservation Master
  - *A high-resolution image, that is uncompressed or uses lossless compression*
  - *Commonly used formats:*

  - *TIFF (6.0, Uncompressed, or with LZW compression)*
  - *JPEG2000-lossless*

- Presentation copy to serve the public (e.g. website)
  - *JPG, PNG, PTIF… any web-friendly, user-friendly image format, can be lossy-compressed.*

# Digitization Standards: Color Mode Selection

- Any item that contains color, or, any item for which there is any doubt about what color space to use: **Full color.**
  - *sRGB or if supported by your software: DCI-P3*

- 24-bit (8 bits per channel) vs 48-bit (16 bits per channel) color
  - *This describes how much bandwidth – or amount of data - is used to encode the color information in your image. More bits means a broader depth of information to describe each color, but takes up more space.*
  - *24 bit color is the most common. 48-bit color is also used, but not always supported well by software packages.*

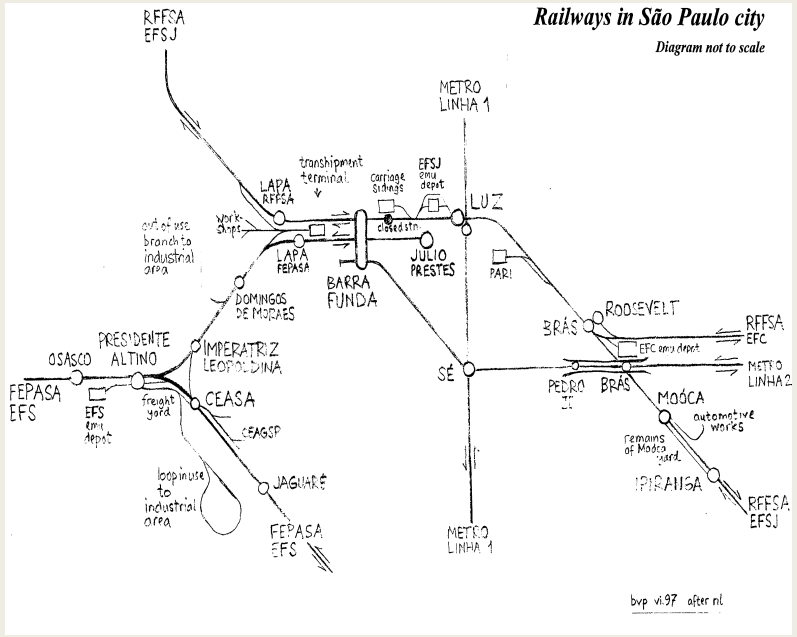
# Digitization Standards: Color Mode Selection

- Use **Grayscale** for most black and white photographs, where there is no color beyond shades of gray.

– *Make sure you are not concerned about depicting patina or other physical aging artifacts of the item.*

# Digitization Standards: Color Mode Selection

- Use **Black and white** for pure text documents and line diagrams, with absolutely no color or gray information.
  - *This is often reserved for typed documents and textbooks with no photographs.*

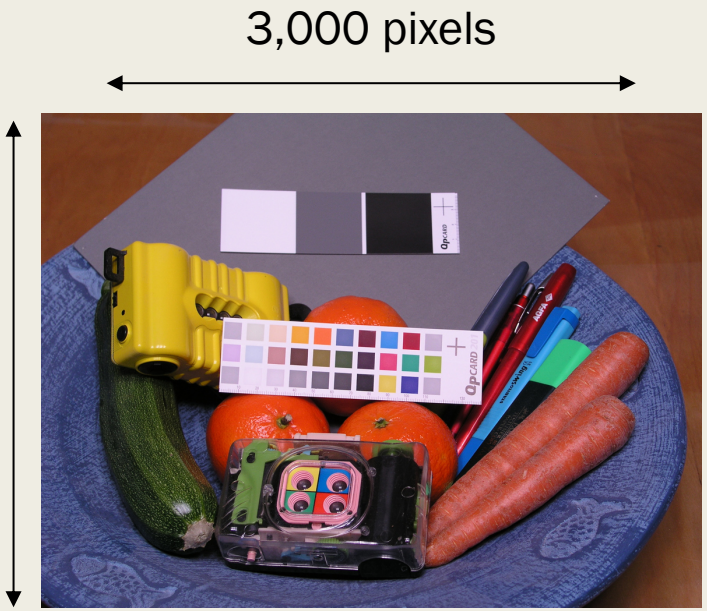


# Digitization Standards: Resolution

- Minimum for most images: 600 dpi
- The 3,000 pixel rule

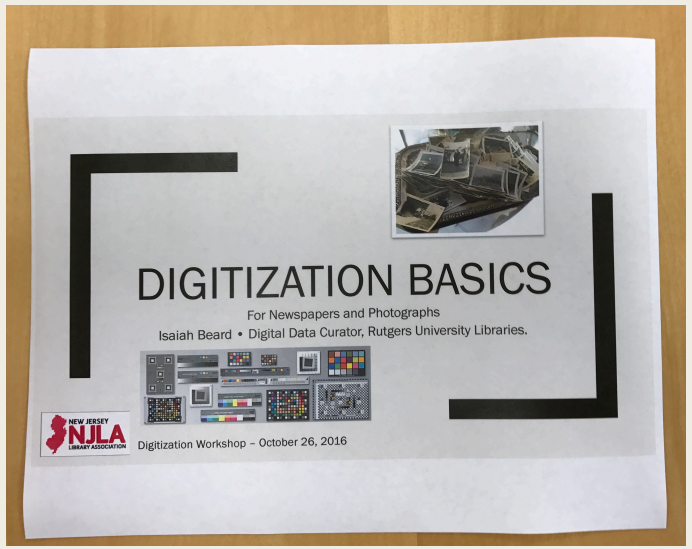
  – *Every image scanned must be at least 3,000 pixels in length or width.*

  – *If, even at the minimum dpi, the image is not at least 3,000 pixels on one axis, the resolution must be increased until this minimum is met.*



# Digitization Standards: Resolution Examples

11” x 8 ½” sheet of paper
Scanned at 600 dpi




5,100 pixels

6,600 pixels

4” x 6” photographic print
Scanned at 600 dpi

3,600 pixels

2,400 pixels

35mm film slide
Scanned at 600 dpi

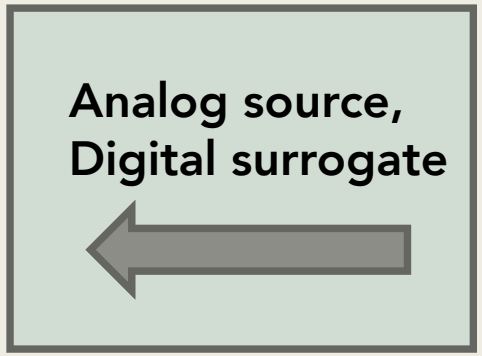

564 pixels

852 pixels

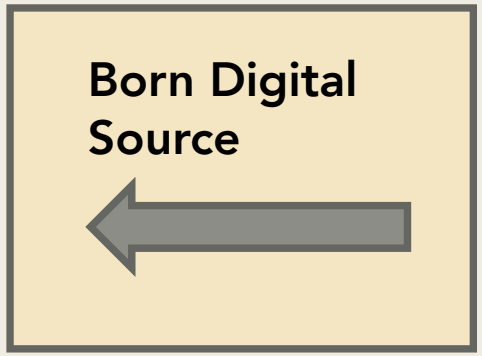
# EQUIPMENT AND PROJECT CONSIDERATIONS

# Consider the scope of the project

- **What are we trying to preserve?**
  - *Photographs, Maps, Slides, Manuscripts and printed documents?*
  - *Audio tapes, transcription discs, wire recordings, film, video tapes?*



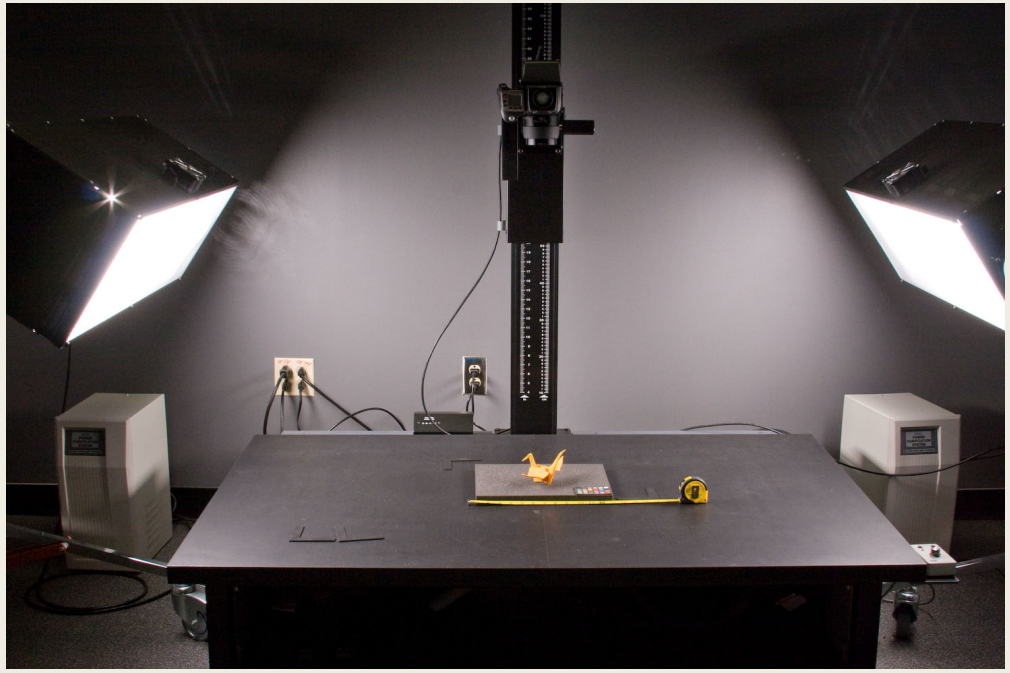

  - *Digital Video/Audio, Computer Documents, Web Sites?*
  - *Creating new content, like oral histories?*
  - *Documenting current events?*



# Consider all uses for the assets you want to use

- Is equipment and/or software already available internally, or with a partner organization?  Can we share it?

- Can resources we must purchase be used by other projects, existing or future? Can those costs be shared?



# Hardware Recommendations

- **“What equipment should I use?”**
- Flatbed scanner

Typical price for letter-sized scanners:
$99 - $200


Tabloid sized models, edge scanners
with more features: ~$2,000

# Hardware Recommendations

- **“What equipment should I use?”**
- Sheet Feed or Document Scanners
  - *Excellent choice if scanning very large volumes of loose letter or legal-sized documents, especially multi-page. Makes fast work of multipage documents.*
  - *Drawbacks: Not useful for photographs or “non-standard” sized objects.*
  - ***NOT for brittle pages!***
  - *Very pricey, but may be worth it in labor savings.*
  - *600-1200 dpi resolution is acceptable.*

Typical price : $600 - $1500

# Hardware Recommendations

- **“What equipment should I use?”**
- Slide/Film Scanners
  - *Virtually required if you have a significant number of slides or negatives that need to be digitized.*
  - *Capable of up to 3600 ppi resolution or higher*
  - *Some large flatbed scanners are capable of also scanning slides, film, and plate glass negatives*


Typical price : $600 - $2500
(standalone)

# Hardware Recommendations

- **“What equipment should I use?”**
- Medium/Large format imaging stations / Planetary scanners
  - *Highly flexible, for a wide variety of large objects (maps, posters, large photo prints, blueprints).*
  - *Can scan 3D objects*
  - *Can be very efficient for large volumes of reprographic work, and for brittle books*

Typical price : $30,000 and up

# Outsourcing to a vendor

- Hiring a third-party company to digitize objects that require special handling, or where the right hardware doesn’t exist in-house
- Often a requirement for large numbers of bound volumes, and large format items (maps, posters, items greater than letter size).
- Per-item cost can be very inexpensive if done in significant volume
- Make sure the vendor is aware of imaging standards and can adhere to them. Ask for test images, and perform stringent quality control.

# Software considerations

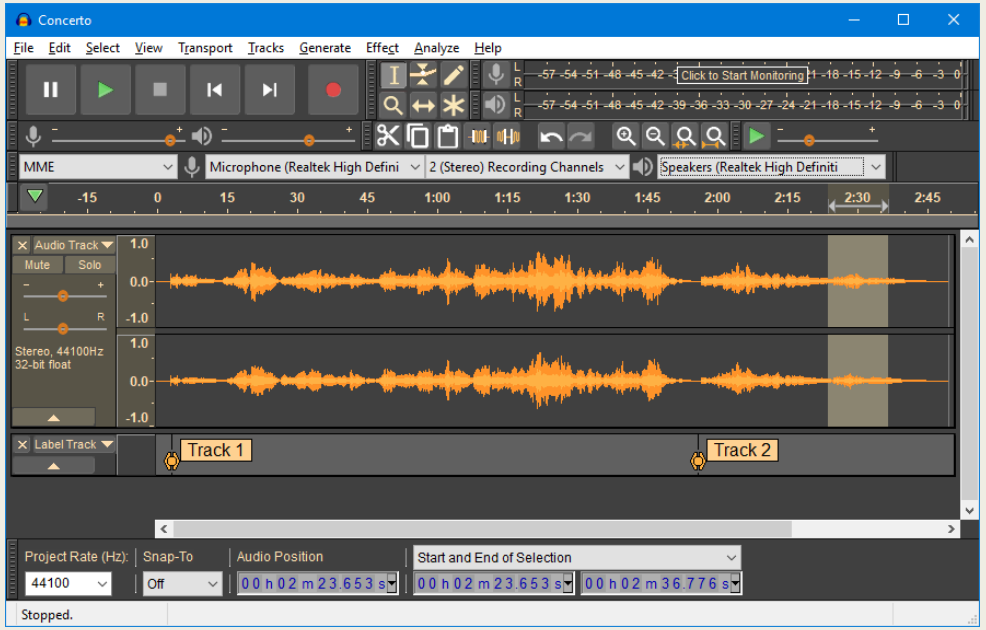

- **Adobe Creative Cloud**
  - *$200+ per year*
  - *($99 for year for photos only)*
  - *Recurring cost, per station*
  - *Full version provides “all in one” functionality, cross-integration*

# Software considerations

- **GiMP (GNU Image Manipulation Program) gimp.org**
  - *Open Source, Multi platform,* ***free***
  - *Can do most photo manipulation tasks*
  - *Has a bit of a learning curve*

# Software considerations

- **Audactity (Sound editing) audacityteam.org**
  - *Open Source, Multi platform, free*
  - *Can do most sound editing tasks*
  - *Has a bit of a learning curve*



# Software considerations

- **OpenShot Video Editor (Sound editing) openshot.org**
  - *Open Source, Multi platform, free*
  - *Can do most video editing tasks*
  - *Has a bit of a learning curve*