



## Prioritizing Collections for Digitization

The Digitization 101 webinar series is a national education program of the Conservation Center for Art & Historic Artifacts provided with support from grants from the National Endowment for the Humanities and the William Penn Foundation.



# Why Prioritize?

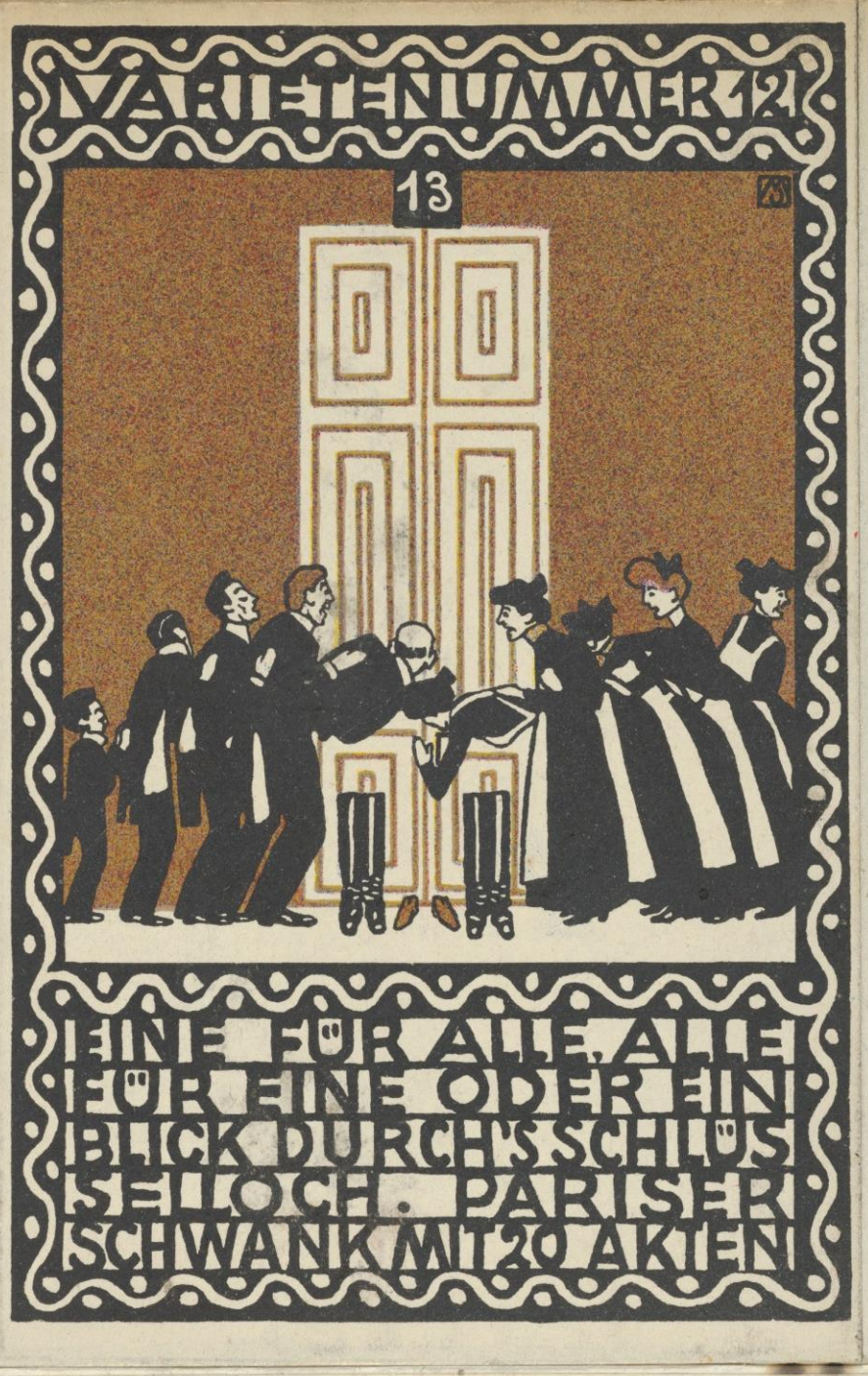
- Establish goals and scope for your project
- Make project more manageable
- Maintain momentum
- Help secure funding
- Help advocate for resources



# Why Prioritize?

Steps that follow will help you rank various collections depending on their:

- Value to institution and stakeholders
- Limiting factors such as copyright and intellectual control
- Physical condition



# Look at What You Have

- Ask questions about your goals and audience
  - What do you want to digitize?
  - Look to your mission for guidance.
  - How much material do you want to digitize?
  - Where will the images be made available?
  - Who is it for?

Moriz Jung, *Variety Act 12: One for All, All for One or a Glimpse through the Keyhole*, 1907. Metropolitan Museum of Art.



# Look at What You Have

- Review any materials that have already been digitized.
  - Why were these materials selected?
  - Who uses the digital files, and how?
  - How were these collections digitized? Did they do a good job?
  - How were prior digitization projects funded?
  - Where are digitized collections stored?

Moriz Jung, *Variety Act 12: One for All, All for One or a Glimpse through the Keyhole*, 1907. Metropolitan Museum of Art.

# Invite Input from Stakeholders

- Non-collection staff and upper administration
- Researchers or scholars
- General public
- Have an agenda and a plan



Moriz Jung, *Conversation between Mutes*, 1907. Metropolitan Museum of Art.

# Invite Input from Stakeholders

- What materials or collections:
  - Are considered the most valuable to the institution?
  - Are most popular with visitors or users?
  - Hold the highest scholarly value?
  - Have the highest monetary value?
  - Have grant or fundraising appeal?
  - Have visual appeal?



Moriz Jung, *Conversation between Mutes*, 1907. Metropolitan Museum of Art.



# Determine Limiting Factors

- Which parts of the collection have been accessioned / processed / inventoried?
- Do any materials have copyright issues?
- Do any materials contain personally identifying information, medical information, or culturally sensitive information?
- How unique are the materials? Are any materials in digital format elsewhere?
- How uniform are your collections?

Moriz Jung, *Variety Act 9: Also Mario Brasso, "Leap of Death" Artist*, 1907. Metropolitan Museum of Art.



# Ask Questions About Condition

- Where is your physical material stored? Are any collections difficult to access?
- Do materials require special handling, such as book cradles, weights, supports, or just extra care?
- Do materials require intervention by an archivist such as removing fasteners or encapsulation?
- Are any collections actively deteriorating? These could include film negatives, brittle newspapers, etc.
- Do any collections require conservation treatment before they're safe to handle?

Moriz Jung, *Greyhound*, 1912. Metropolitan Museum of Art.

# Weed Out Materials

Collection Name	Under Copyright	Contains Sensitive Information	Unprocessed	Digitize?
Smith Family Papers	No	No	Yes	Wait
Church Records	No	No	No	Proceed
Town Records	Maybe	Maybe	No	Evaluate further

# Prioritize Further

Collection Name	Scholarly Value	Grant Appeal	Collection Uniformity	Total
Early photography collection	5	5	2	12
Church Records	5	5	4	14
Correspondence	3	3	5	11

# Record Your Decisions

- Record decisions in writing
- Incorporate into Selection Policy or create a new policy
- Continue to update



Moriz Jung, *Bulldog*, 1912. Metropolitan Museum of Art.



# Examples

Dartmouth College Library - Selection Policy for Digitization

<https://www.dartmouth.edu/~library/digital/about/policies/selection.html>

Georgetown University Library - Digital Collections Development Policy

<https://www.library.georgetown.edu/digital-project-policy>

Amherst College – Digital Collection Development Policy

<https://www.amherst.edu/library/services/digital/digitalcollectiondev>



# Coming Up

- File Storage for Digitization – June 30
- File Specifications for Digitization – July 7
- Metadata for Digitization – July 14
- In-house vs Outsourcing and Quality Control – July 21
- Equipment Selection – July 28

Moriz Jung, *Act 8: Gottlieb Säuselin the Improviser*, 1907. Metropolitan Museum of Art.



# Questions?

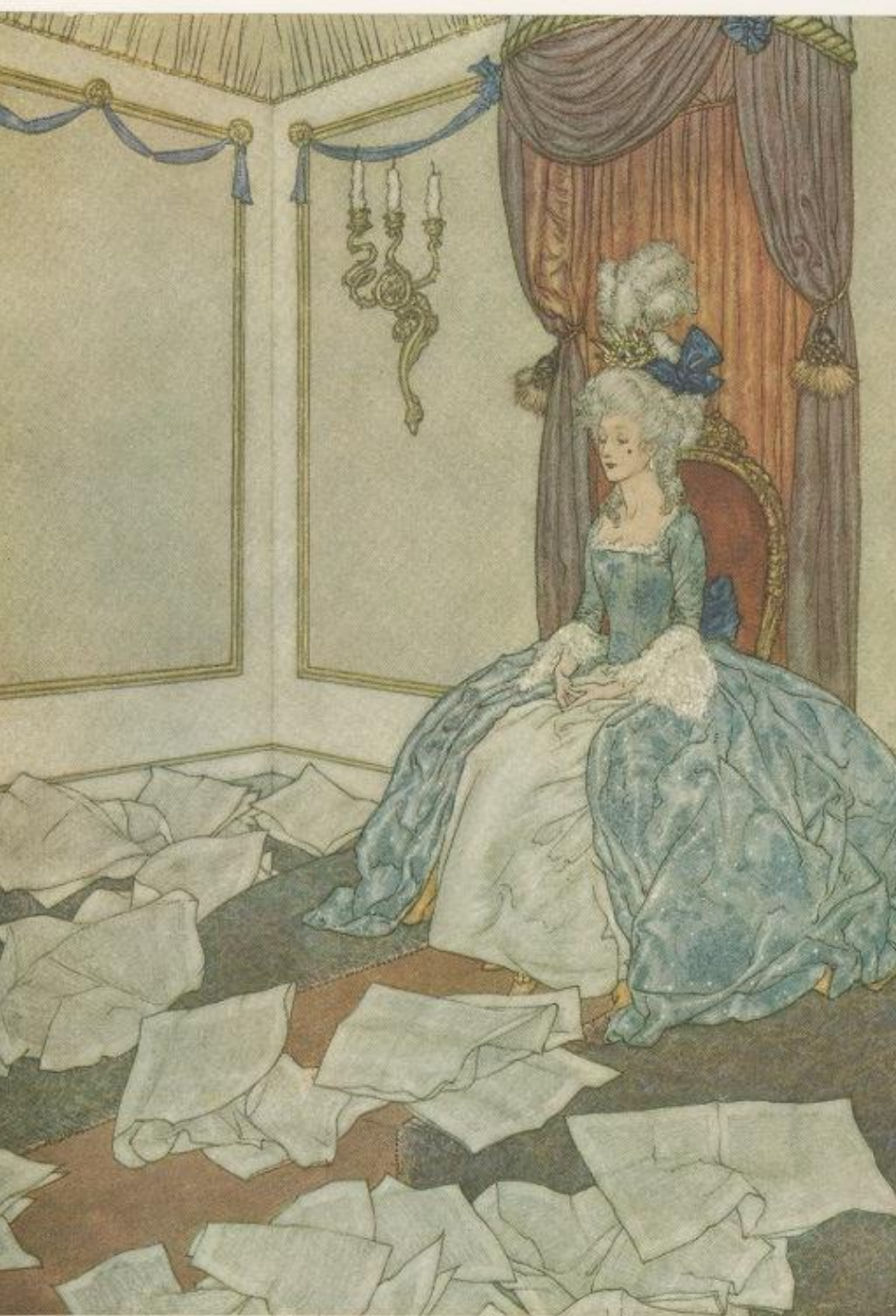
Email: [mdowning@ccaha.org](mailto:mdowning@ccaha.org)

Moriz Jung, *Variety Act 10: Man or Reptile or Razor?*, 1907. Metropolitan Museum of Art.



## File Storage for Digitization

The Digitization 101 webinar series is a national education program of the Conservation Center for Art & Historic Artifacts provided with support from grants from the National Endowment for the Humanities and the William Penn Foundation.



# File Storage Considerations

---

- Protect your investment!
- Management software programs
- Types of file storage
- Redundancy and backups
- Calculating storage space needs



# Collection Management System (CMS)

---

- Stores information about your collection items, and can include:
  - Donor information
  - Physical description, location, and condition
  - Loan information
  - Can store images
  - Can link to a web platform to make collections searchable online
- Locally stored or cloud-based
- Proprietary or open source
- Examples: PastPerfect, Embark, and Collective Access

E. Boyd Smith, illustration from *The Queen of Hearts*, 1919. Picture Collection, The New York Public Library.



# Digital Asset Management System (DAMS)

---

- Manages the organization and distribution of the files themselves
- Some programs are suitable as both a CMS and a DAMS
  - Such as DSpace, The Museum System (TMS), and ContentDM

E. Boyd Smith, illustration from *The Queen of Hearts*, 1919. Picture Collection, The New York Public Library.



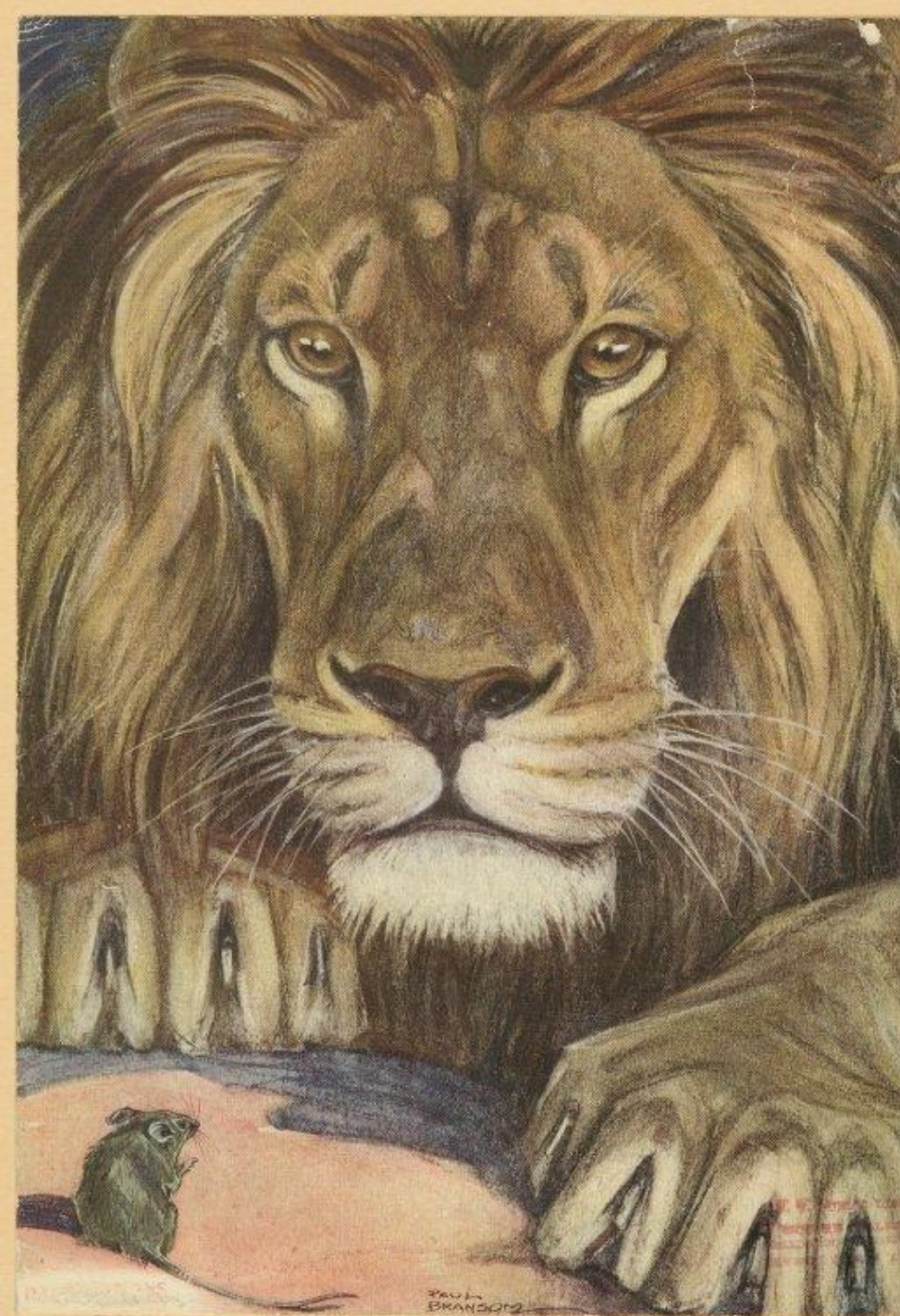
"HE SET OUT AT ONCE TO CLIMB THE BEANSTALK."  
#19,992 (©1927)

# Types of File Storage

---

- Network Attached Servers (NAS)
- Cloud Storage
  - Amazon Web Services (AWS)
  - Microsoft Azure
  - Costs vary based on amount of storage and level of access
- Hosted Storage
  - Preservica
  - DuraCloud
- Removable Media
  - Optical disks, hard drives, flash drives

Elizabeth Curtis, illustration from *Jack and the Beanstalk*, 1917. Picture Collection, The New York Public Library.

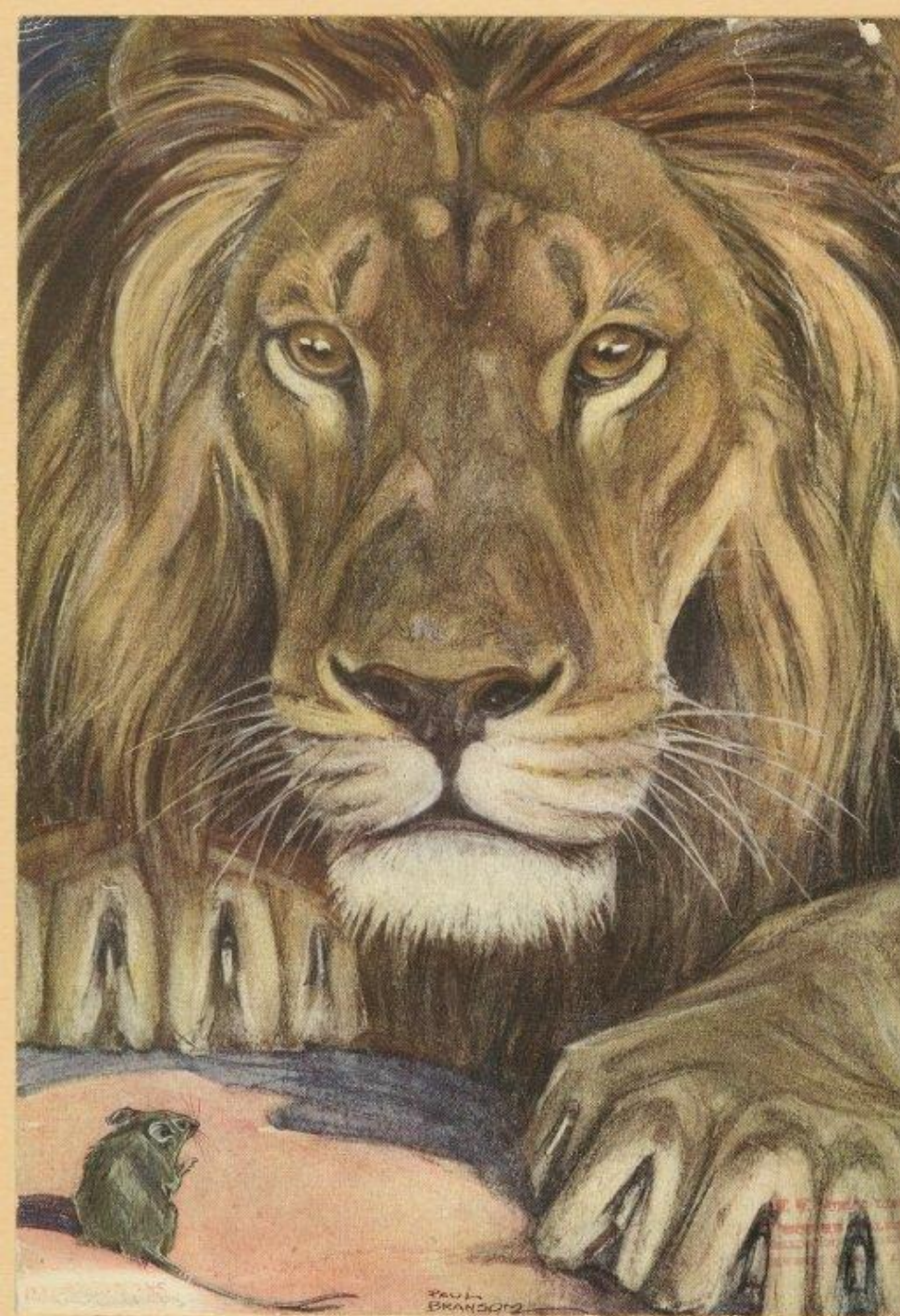


# Keep Your Files Safe

---

- Threats to digital files
  - Obsolescence, accidental deletion, viruses, hacking
- LOCKSS – Lots Of Copies Keeps Stuff Safe
- 3-2-1 Principle for backups
  - 3 copies of your files
  - 2 media types
  - 1 in another geographic location

Paul Bransom, illustration from *The Lion and the Mouse*, 1921. Picture Collection, The New York Public Library.



# Keep Your Files Safe

---

- Generate a checksum for your files
  - Digital fingerprint
  - Changes in the bitstream of the file changes
  - Digital Preservation Coalition Handbook section on Fixity and Checksums:

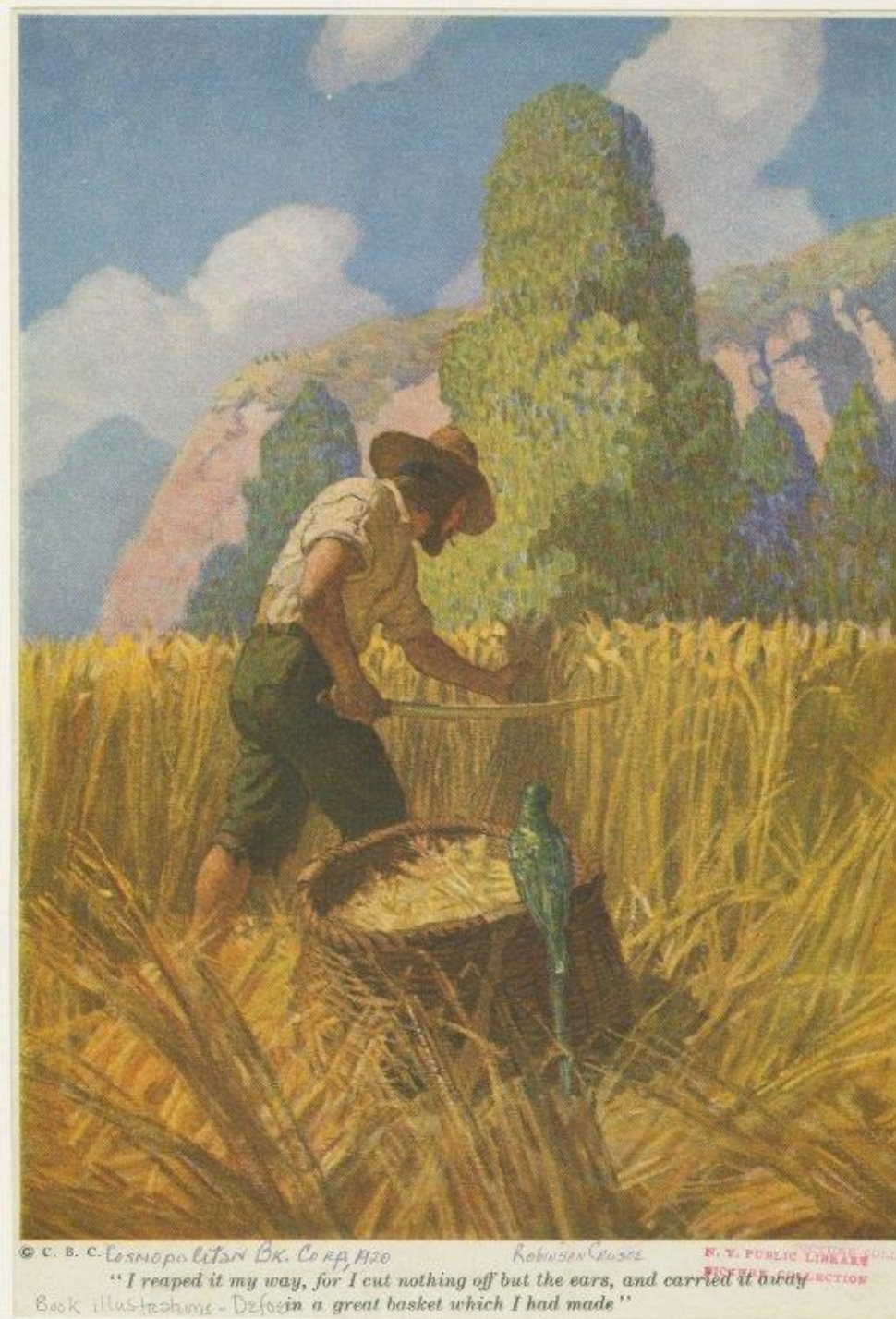
<https://www.dpconline.org/handbook/technical-solutions-and-tools/fixity-and-checksums>

Paul Bransom, illustration from *The Lion and the Mouse*, 1921. Picture Collection, The New York Public Library.

# How Much Storage Do You Need?

---

- How many items you want to digitize?
- What file size will you get from scanning at the resolution you need?
- Will you be scanning negatives and then making additional positive files?

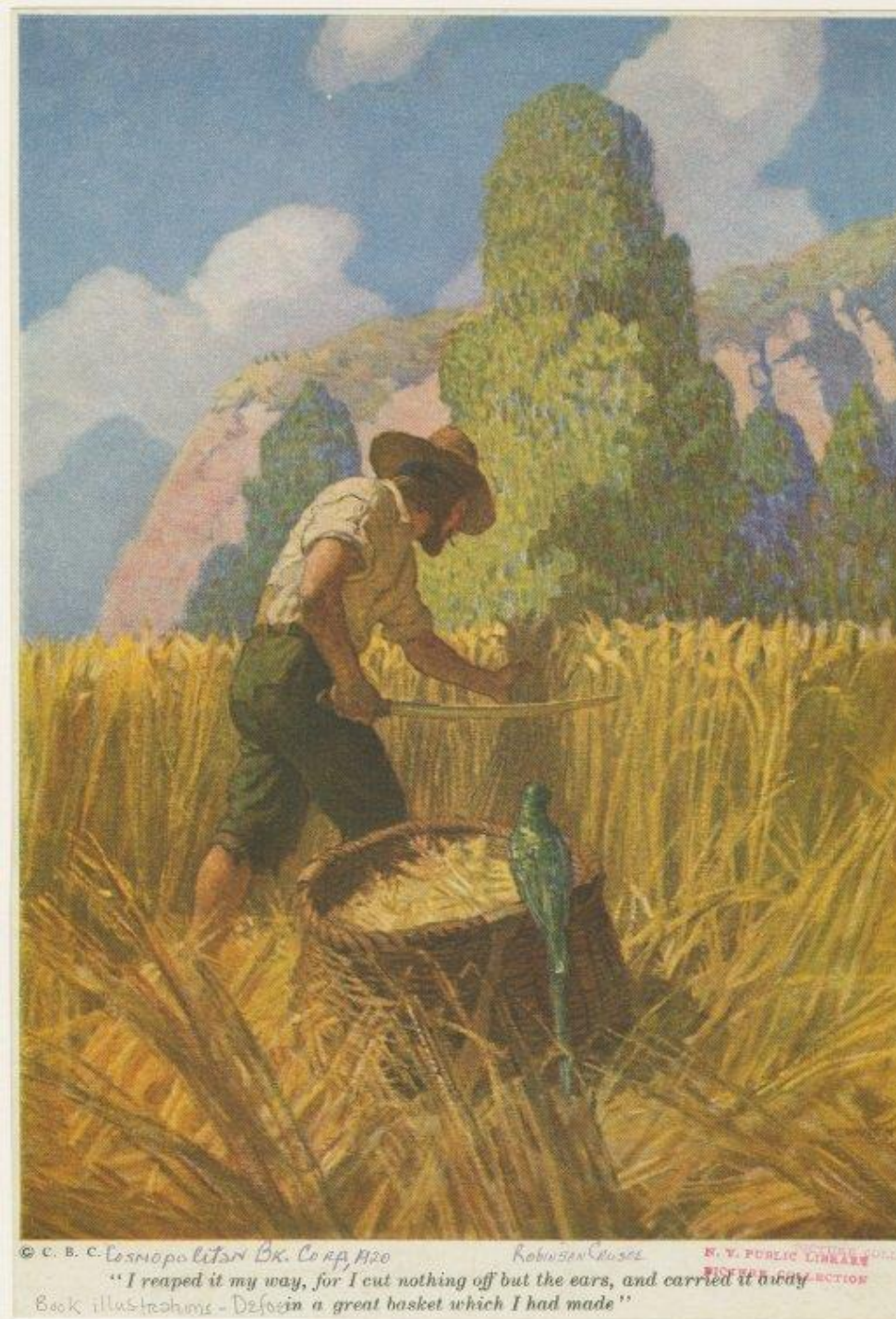


N.C. Wyeth, illustration from *Robinson Crusoe*, 1920. Picture Collection, The New York Public Library.

# How Much Storage Do You Need?

---

- Don't forget the derivative files you want to add to master copies.
- Are you incorporating video or sound collections that you want to digitize?
- Plan for the future! Select a storage option that you can add onto incrementally.



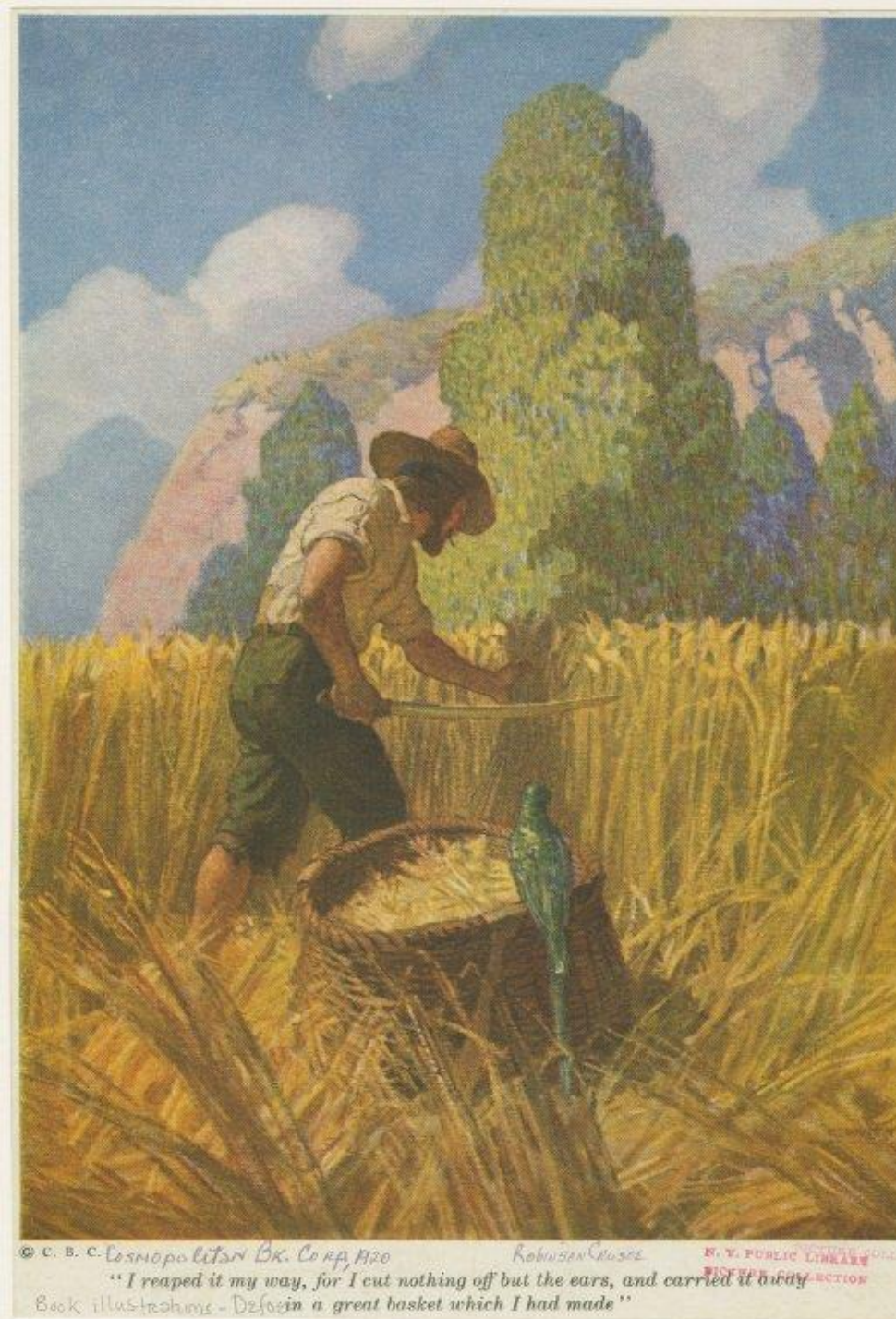
N.C. Wyeth, illustration from *Robinson Crusoe*, 1920. Picture Collection, The New York Public Library.

# How Much Storage Do You Need?

---

Example:

- 8x10 print – 16-bit, color, 400 dpi TIFF
  - Produces a file size of about 75 MB
  - A JPEG derivative copy will add about 7MB
  - 1,000 prints = at least 82 GB for one set of files



N.C. Wyeth, illustration from *Robinson Crusoe*, 1920. Picture Collection, The New York Public Library.

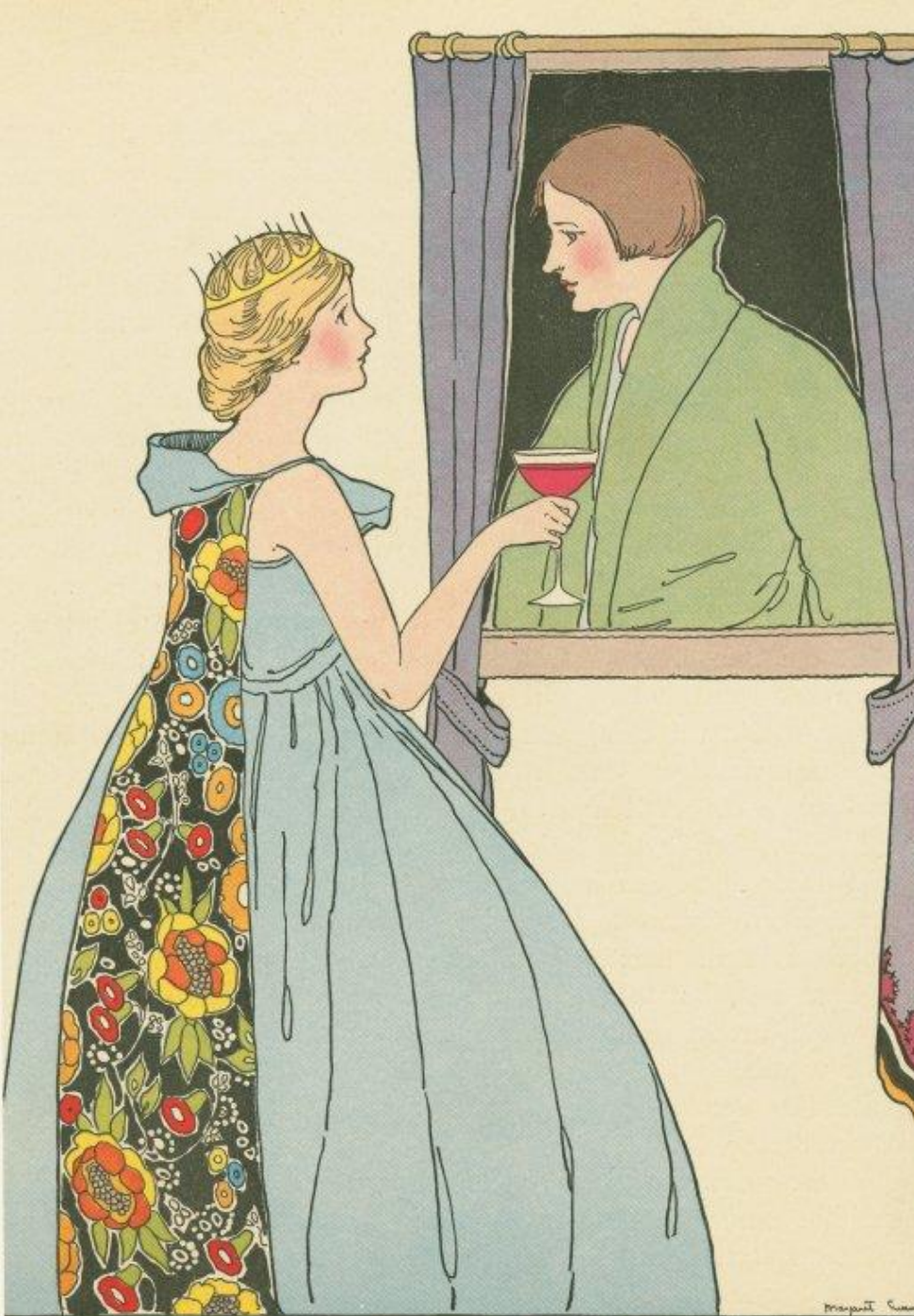


[I2I] #14,778 (n.d.)  
Illus. by Edmund Dulac

# Manage Your Files

---

- Monitor files
  - Fixity checks
- Maintain technical infrastructure:
  - Maintain hardware, software, facilities, supplies
- Migrate files
  - Convert data to latest file formats or relocate to new storage media
  - Consider a migration schedule every 5-10 years



# Resources

---

- DHPSNY webinar "Managing Your Digital Collections": <https://www.youtube.com/watch?v=PHD92p2cUjo&feature=youtu.be>
- Additional DHPSNY & CCAHA webinars: <https://dhpsny.org/webinars> / [ccaha.org/events](https://ccaha.org/events)
- Digital Preservation Coalition: <https://www.dpconline.org/>
- Digital POWRR: [digitalpowrr.niu.edu/](https://digitalpowrr.niu.edu/)

Margaret Evans Price, illustration from *The Twelve Dancing Princesses*, 1921. Picture Collection, The New York Public Library.

THE PRINCESS OFFERED THE SOLDIER A GOBLET OF WINE

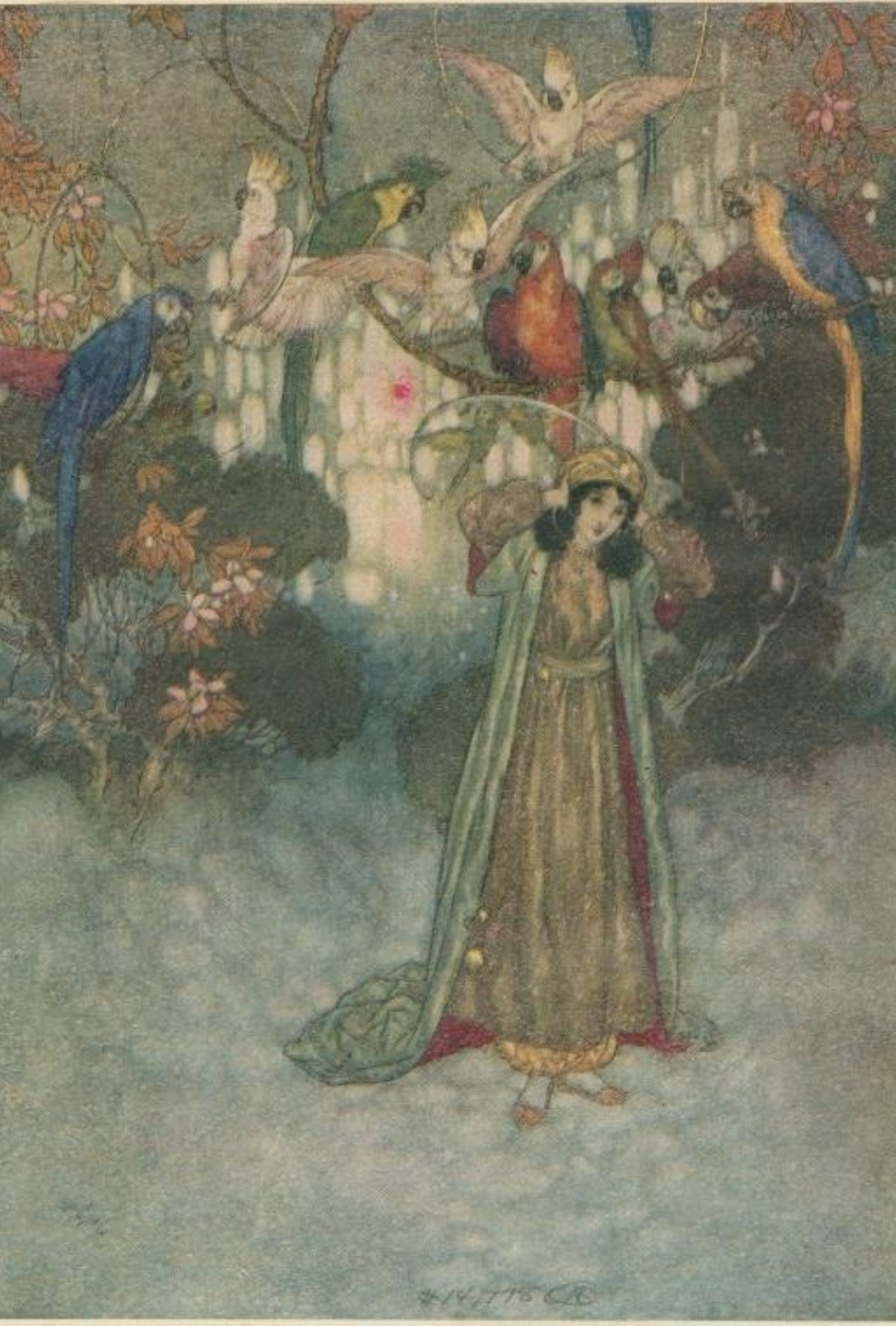


# Coming Up

---

- File Specifications for Digitization – July 7
- Metadata for Digitization – July 14
- In-house vs Outsourcing & Quality Control – July 21
- Equipment Selection – July 28

Charles Folkard, illustration from *Little Bo Peep*, 1922. Picture Collection, The New York Public Library.



# Questions?

---

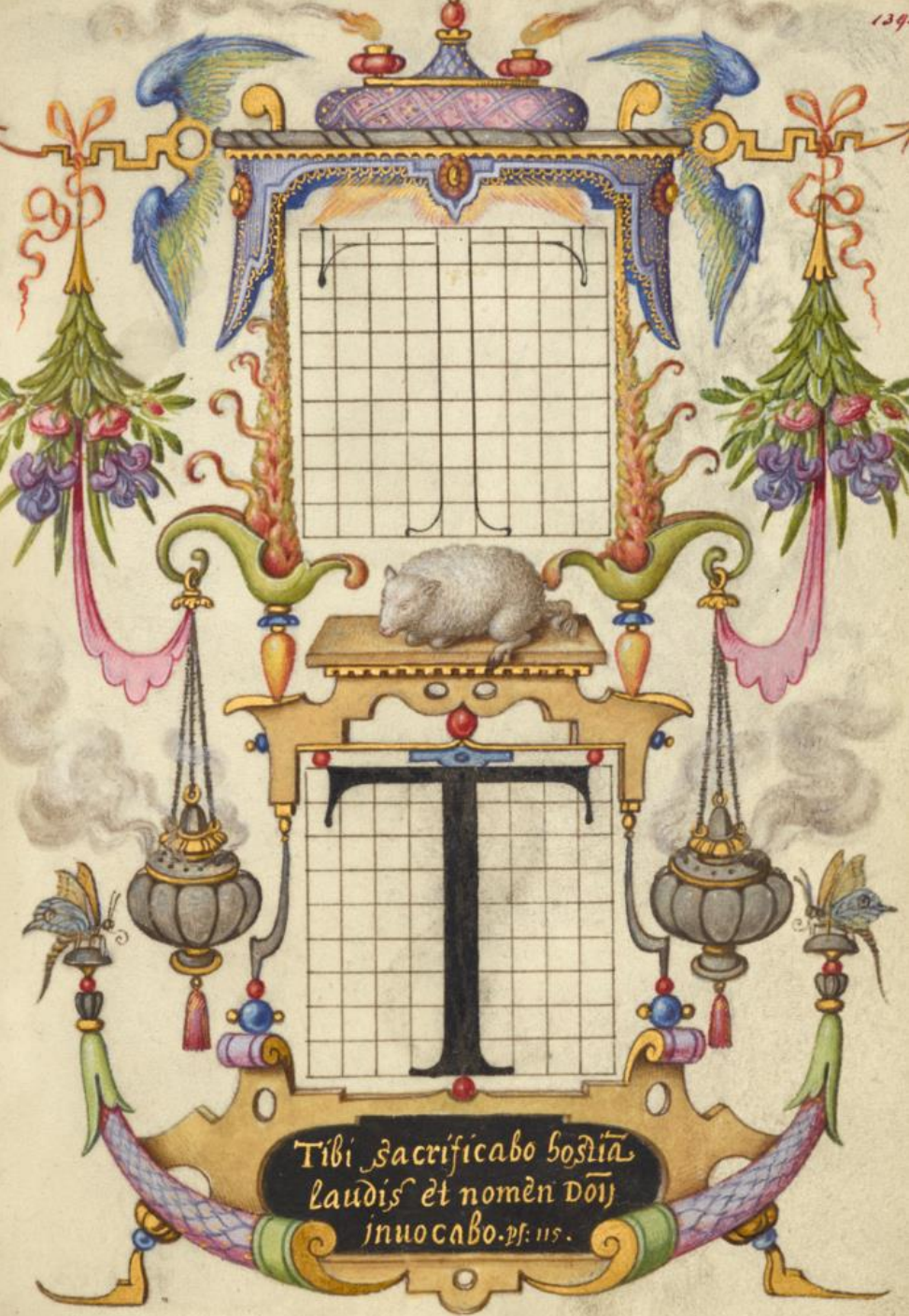
Email: [mdowning@ccaha.org](mailto:mdowning@ccaha.org)

Edmund Dulac, illustration from *Beauty and the Beast*. Picture Collection, The New York Public Library.



## File Specifications for Digitization

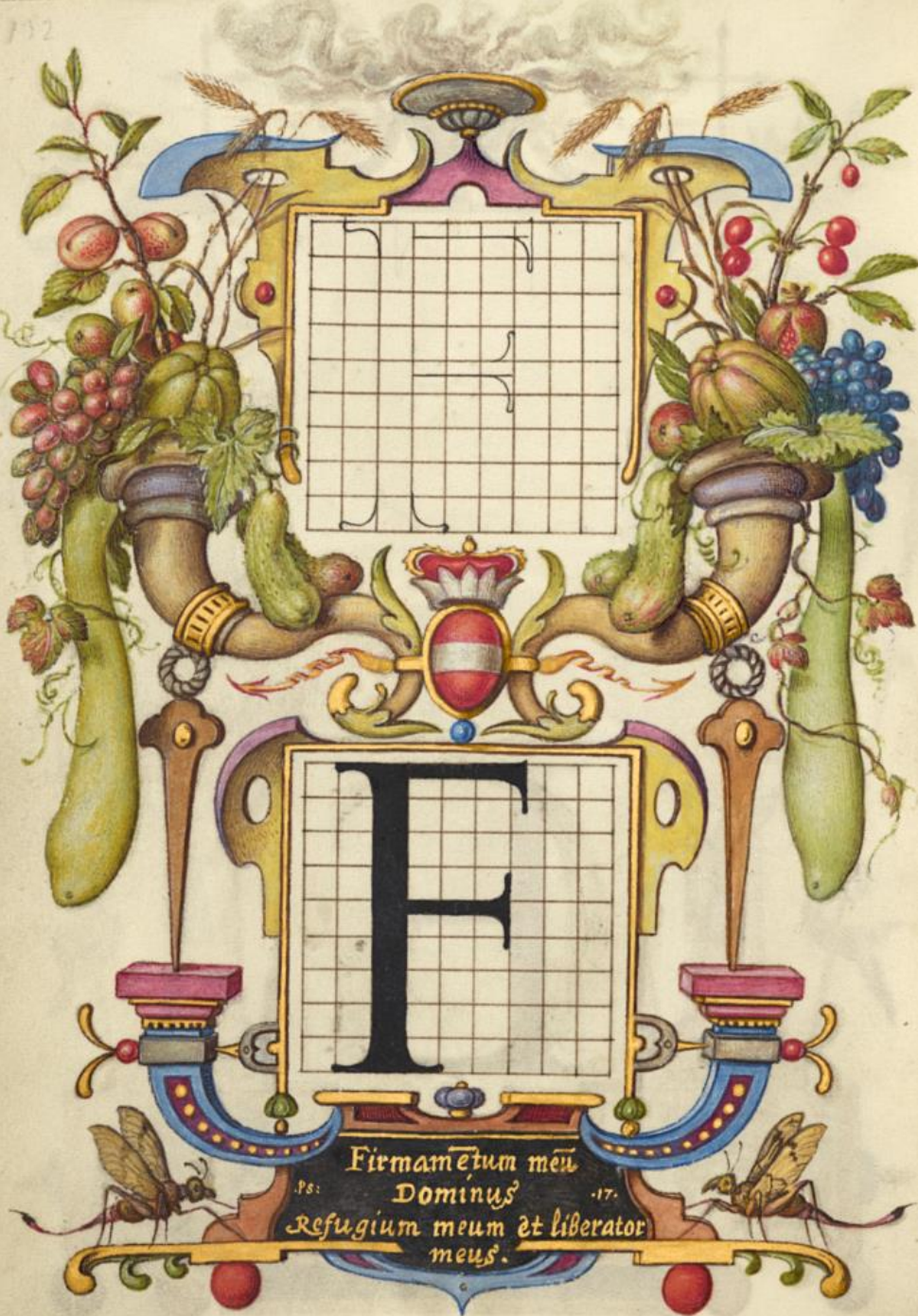
The Digitization 101 webinar series is a national education program of the Conservation Center for Art & Historic Artifacts provided with support from grants from the National Endowment for the Humanities and the William Penn Foundation.



# Terminology

---

- File Format
- Resolution
- Color
- Bit Depth
- Color Space
- Cropping
- File Naming



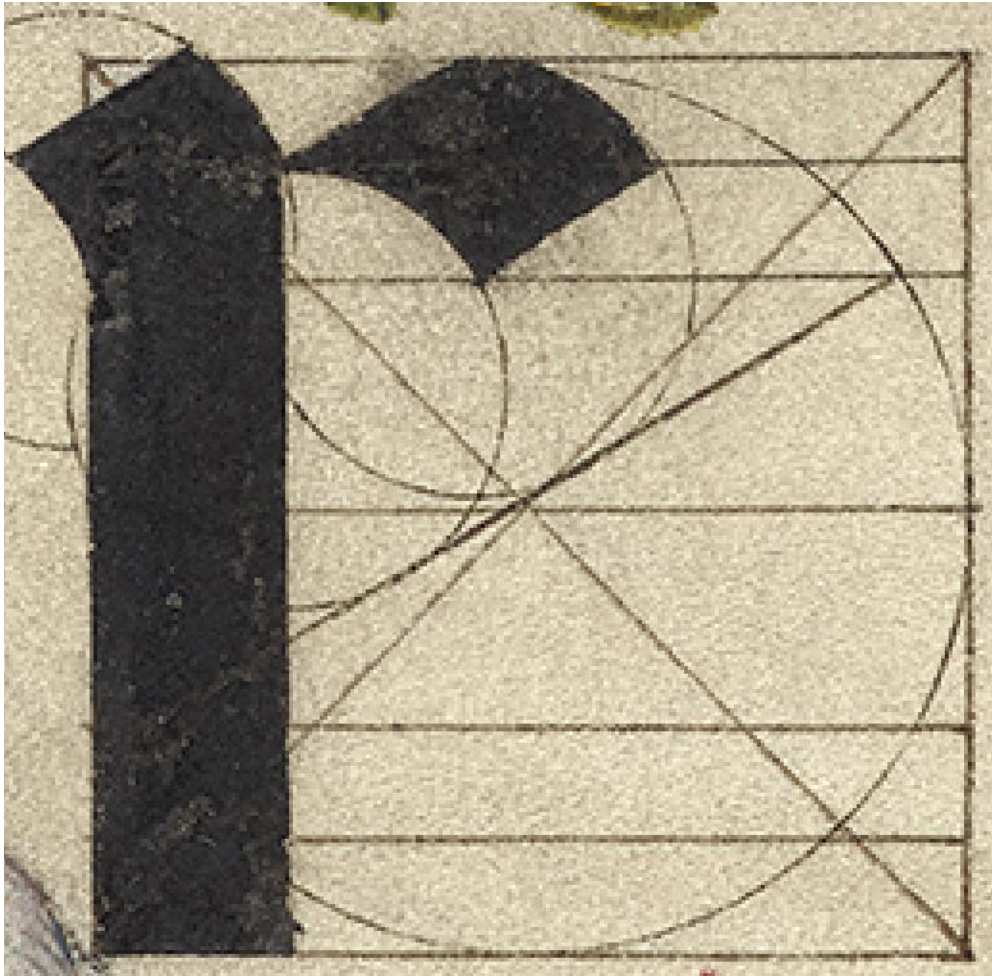
# File Format

- Convention for encoding data into human-readable form

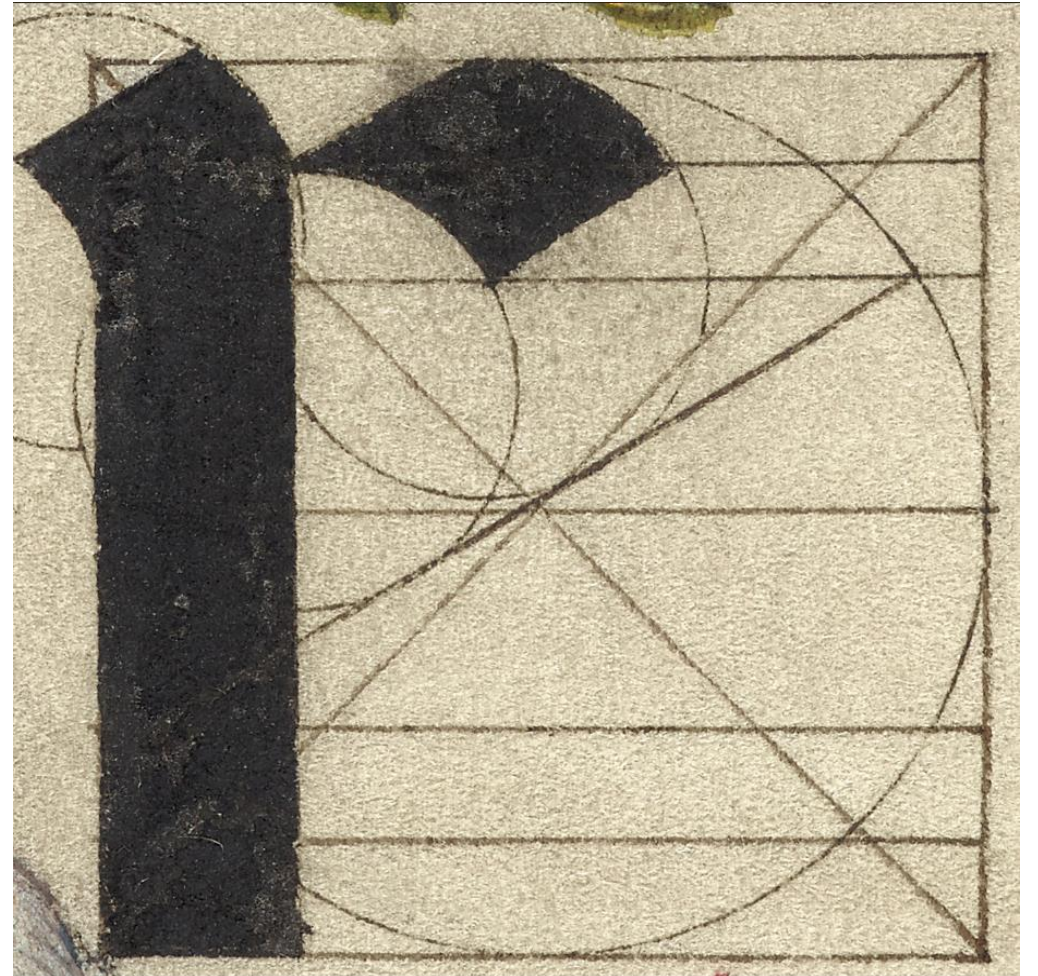


# Resolution

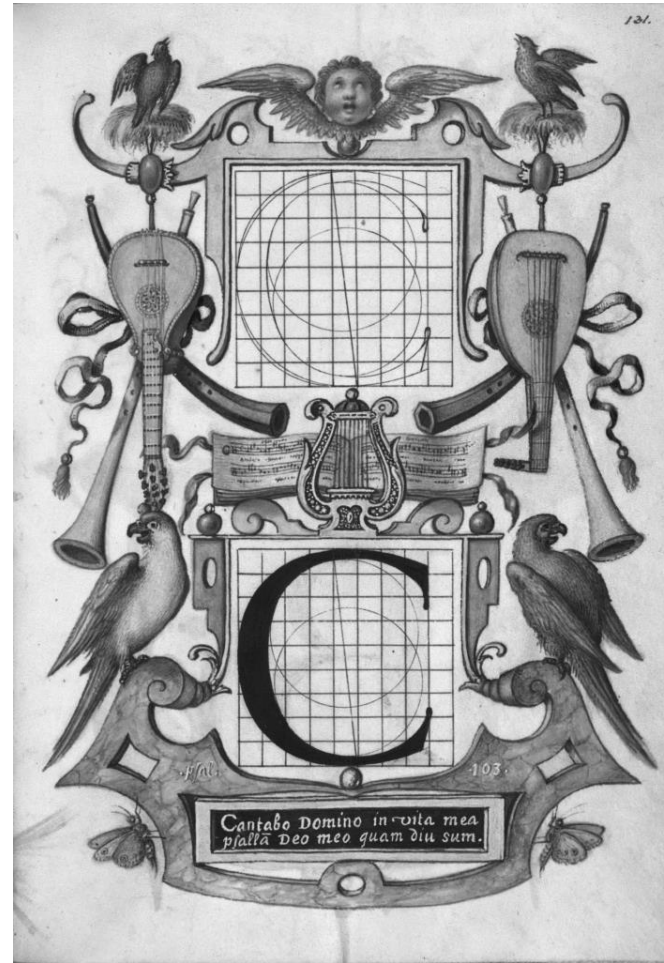
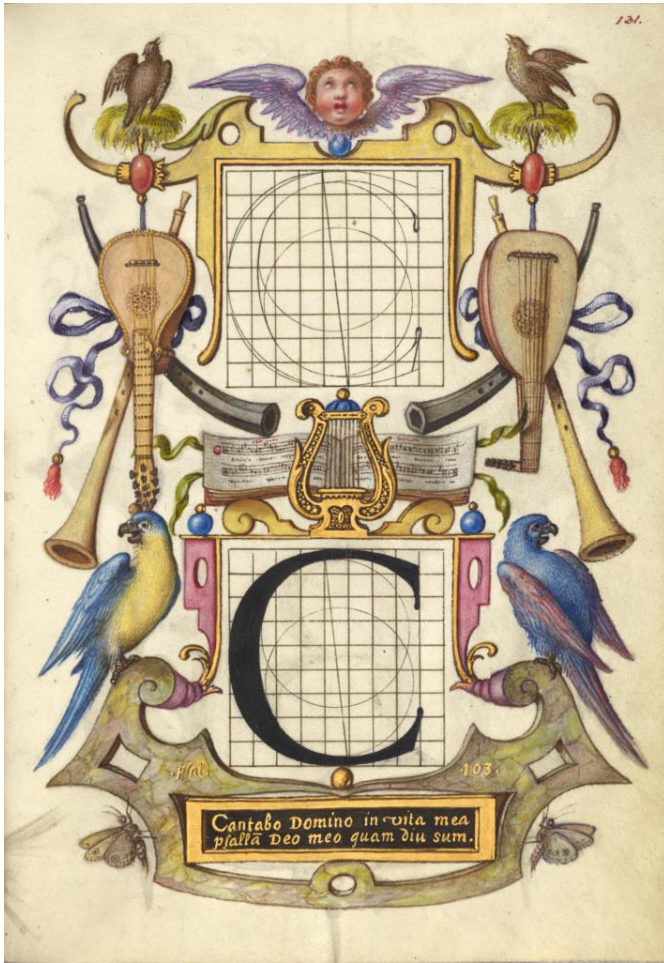
150 ppi



800 ppi



# Color



Joris Hoefnagel, *Guide for Constructing the Letter C*, 1591–1596. J. Paul Getty Museum.

# Bit Depth

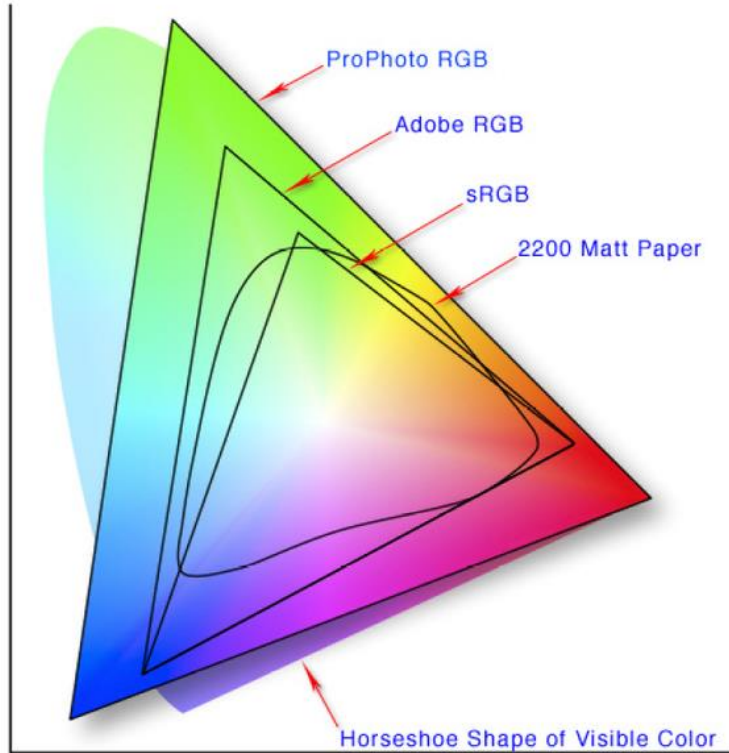
1-bit (2 tones available per pixel)



8-bits (256 tones available per pixel)



# Color Space / Color Profile



Defines the range and tones available in the file.

Examples of common color spaces:

- **sRGB** – low range of color, common in web publication
- **Adobe RGB 1998** – medium range of color, widely adopted
- **ProPhoto RGB** – huge range of color, newer

Color tutorial: Understanding Color Spaces

<https://www.youtube.com/watch?v=KKX08oOTMkk>

# Cropping



Albrecht Dürer, *Stag Beetle*, 1505. J. Paul Getty Museum.



# File Naming

- Structure for naming files

Name

StOlaf\_RML\_Felland\_023\_P  
StOlaf\_RML\_Felland\_053\_P  
StOlaf\_RML\_Felland\_062\_P  
StOlaf\_RML\_Felland\_117\_P  
StOlaf\_RML\_Felland\_228\_P  
StOlaf\_RML\_Felland\_234\_P  
StOlaf\_RML\_Felland\_271\_P  
StOlaf\_RML\_Felland\_335\_P  
StOlaf\_RML\_Felland\_348\_P  
StOlaf\_RML\_Felland\_360\_P  
StOlaf\_RML\_Felland\_381\_P  
StOlaf\_RML\_Felland\_410\_P  
StOlaf\_RML\_Felland\_420\_P



# Standards and Recommendations

---

American Library Association (ALA): Minimum Digitization Capture Recommendations

<http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations>

National Archives and Records Administration (NARA): Technical Guidelines for Digitizing Archival Materials

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

Federal Agencies Digitization Guidelines Initiative (FADGI): Technical Guidelines for Digitizing Cultural Heritage Materials

[http://www.digitizationguidelines.gov/guidelines/FADGI%20Federal%20%20Agencies%20Digital%20Guidelines%20Initiative-2016%20Final\\_rev1.pdf](http://www.digitizationguidelines.gov/guidelines/FADGI%20Federal%20%20Agencies%20Digital%20Guidelines%20Initiative-2016%20Final_rev1.pdf)



# FADGI Guidelines

---

- Bound Volumes
- Documents
- Oversize Items: Maps, Posters and Other Materials
- Newspapers
- Prints and Photographs
- Photographic Transparencies of various sizes
- Paintings
- X-Ray Film
- Microfilm

Example of Performance Levels Table

Prints and Photographs

Performance Level:

	1 Star	2 Star	3 Star	4 Star
Master File Format	TIFF	TIFF	TIFF	TIFF
Access File Formats	All	All	All	All
Resolution	100 ppi	200 ppi	400 ppi	600 ppi <sup>1</sup>
Bit Depth	8	8	8 or 16	16
Color Space	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2	Adobe 1998 ProPhoto, ECIRGBv2	Adobe 1998 ProPhoto, ECIRGBv2
Color	Grayscale or Color	Grayscale or Color	Color	Color

# FADGI Guidelines

---

- Quality-Level Ratings – 1-star through 4-star
  - Focus on 3- or 4-star levels for digital surrogates
- Performance Levels
  - File format
  - Resolution
  - Bit Depth
  - Color Space
  - Color



# File Format

## Prints and Photographs

Performance Level:

	1 Star	2 Star	3 Star	4 Star
Master File Format	TIFF	TIFF	TIFF	TIFF
Access File Formats	All	All	All	All

### Master files – TIFF

- Widely adopted
- Accommodates variety of color spaces
- High-bit compatible, suitable for large file sizes

### Derivative files

- PDF – Great for multi-page objects
- JPEG—Smaller files, good for email and websites



# Resolution

## Prints and Photographs

Performance Level:

	1 Star	2 Star	3 Star	4 Star
Master File Format	TIFF	TIFF	TIFF	TIFF
Access File Formats	All	All	All	All
Resolution	100 ppi	200 ppi	400 ppi	600 ppi <sup>1</sup>

## Examples:

- Newspapers – 300 ppi or 400 ppi
- Prints and Photographs – 400 ppi or 600 ppi
- 4x5" photographic transparency – 1,500 ppi or 2,000 ppi
- 35mm photographic transparency – 3,000 ppi or 4,000 ppi



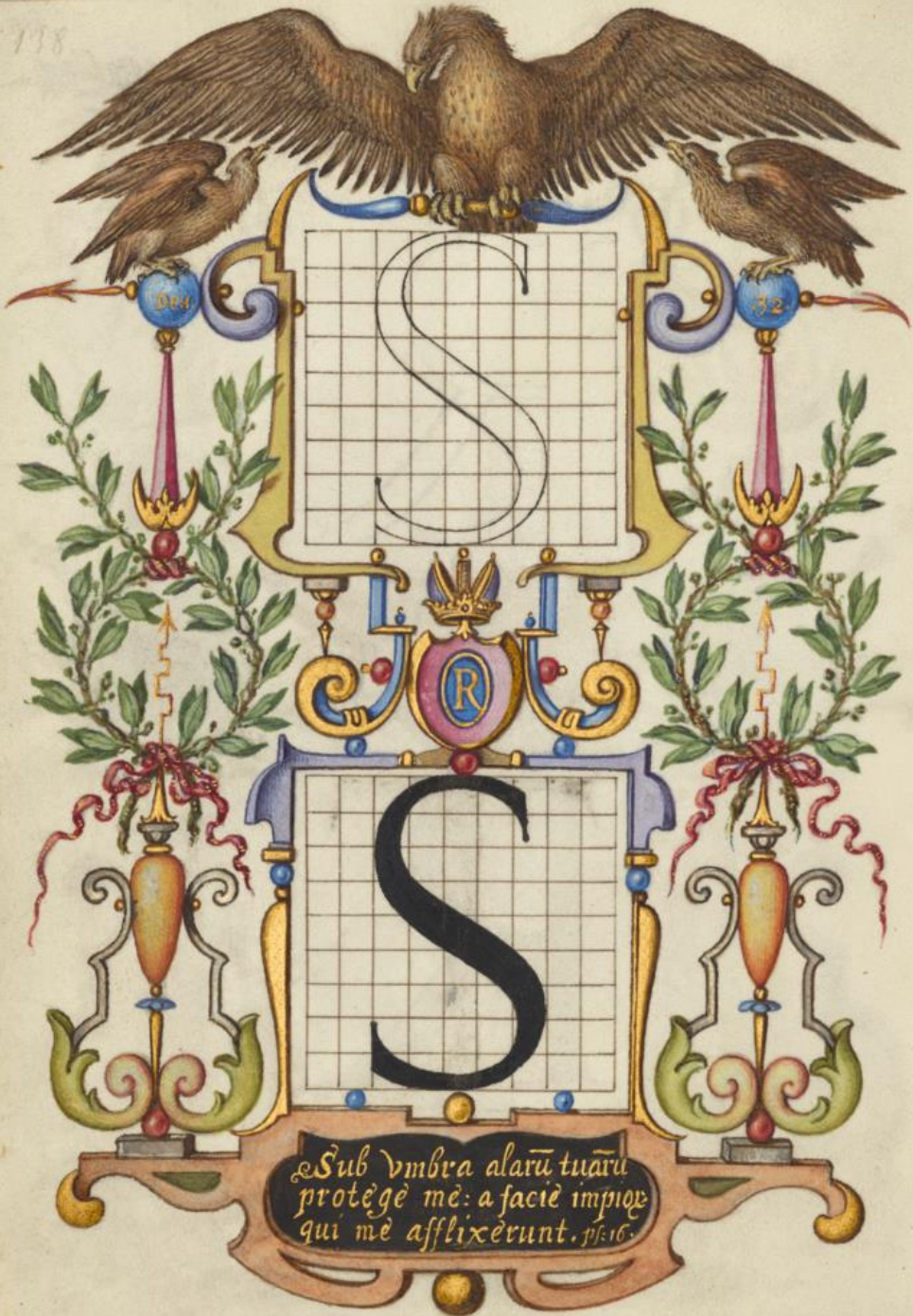
# Bit Depth

## Prints and Photographs

Performance Level:

	1 Star	2 Star	3 Star	4 Star
Master File Format	TIFF	TIFF	TIFF	TIFF
Access File Formats	All	All	All	All
Resolution	100 ppi	200 ppi	400 ppi	600 ppi <sup>1</sup>
Bit Depth	8	8	8 or 16	16

- Lower bit depth – 8-bit
  - Documents, newspapers, microfilm, e.g.
- Higher bit depth – 16-bit
  - Fine art, prints, and photographic negatives, e.g.



# Color Space

Color Space	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2	Adobe 1998 ProPhoto, ECIRGBv2	Adobe 1998 ProPhoto, ECIRGBv2

- FADGI Recommendations
  - Grey Gamma 2.2 - black and white
  - SRGB – low color range, common in web publication
  - Adobe 1998 – medium color range, widely adopted
  - ProPhoto – wide color range, newer



# Color

## Prints and Photographs

Performance Level:

	1 Star	2 Star	3 Star	4 Star
Master File Format	TIFF	TIFF	TIFF	TIFF
Access File Formats	All	All	All	All
Resolution	100 ppi	200 ppi	400 ppi	600 ppi <sup>1</sup>
Bit Depth	8	8	8 or 16	16
Color Space	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2	Adobe 1998 ProPhoto, ECIRGBv2	Adobe 1998 ProPhoto, ECIRGBv2
Color	Grayscale or Color	Grayscale or Color	Color	Color

- Will you scan in color or in black and white?
- Black and white files are about 1/3 the size of color files.



# Cropping

- Master file – Show all edges of object
- Derivative file – Can be cropped



O.G. Felland, *Valborg Feeding Doll*, 1899. St. Olaf College Rolvaag Memorial Library.

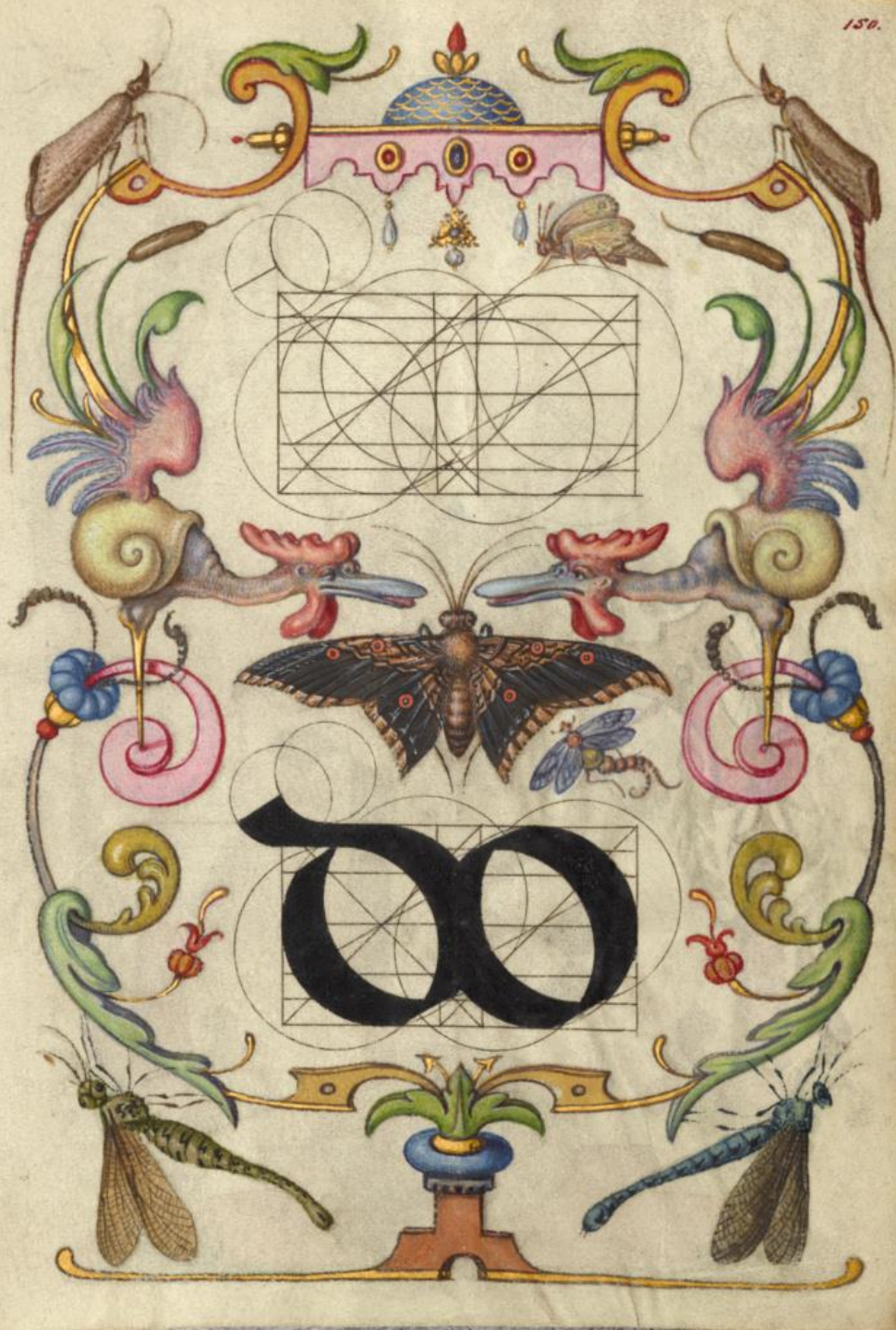


# File Naming

---

- Unique and consistently structured
- Should have some level of description
  - Example: **StOlaf\_RML\_Felland\_0001.tif** vs. **0001.tif**
- Use leading zeros when you are numbering items or pages
  - 0001, 0002, ... 1023, etc.
- Do not include shelf numbers or locations in file names
- ALA Minimum Digitization Capture Recommendations

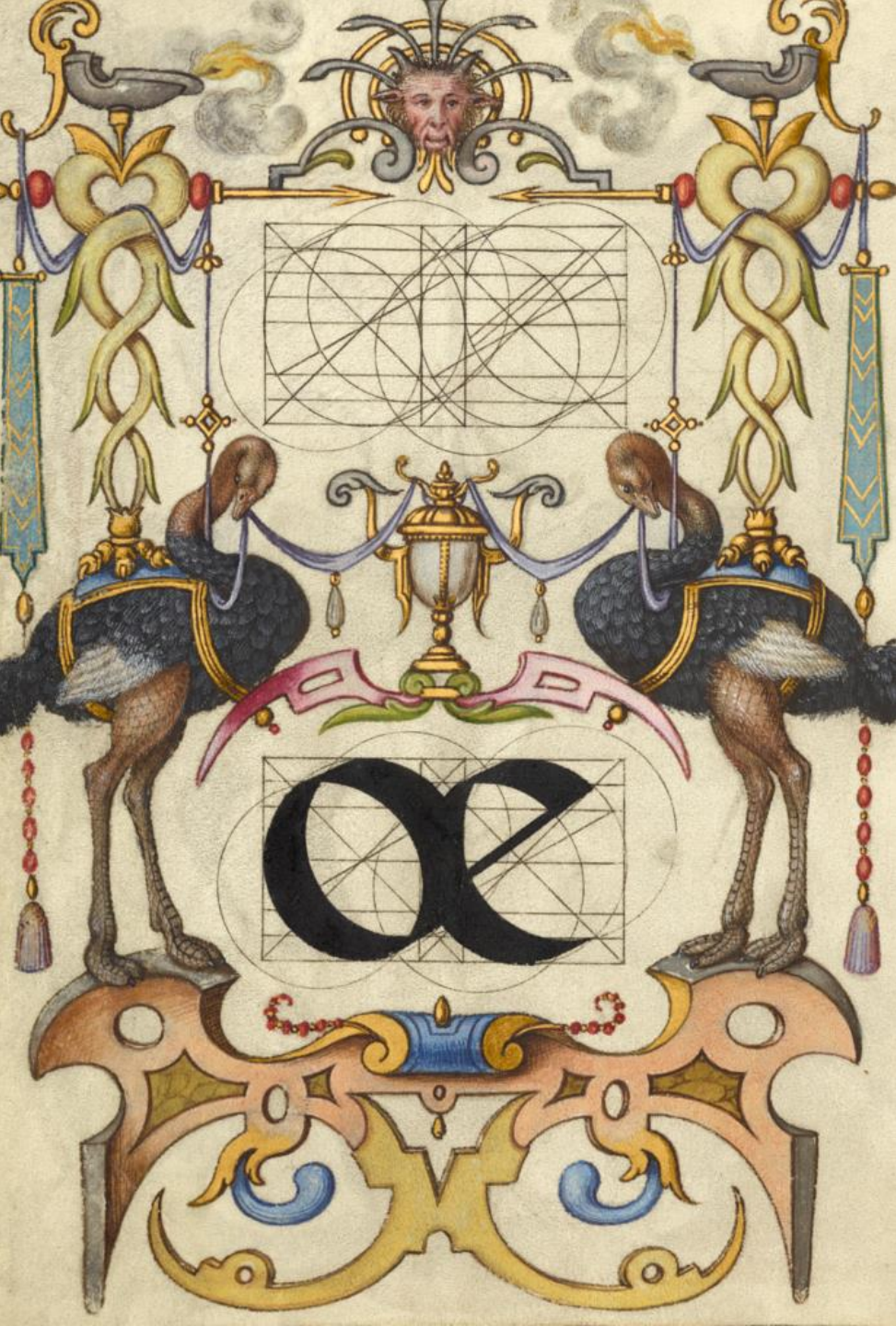
<http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations>



# Record Your Decisions

---

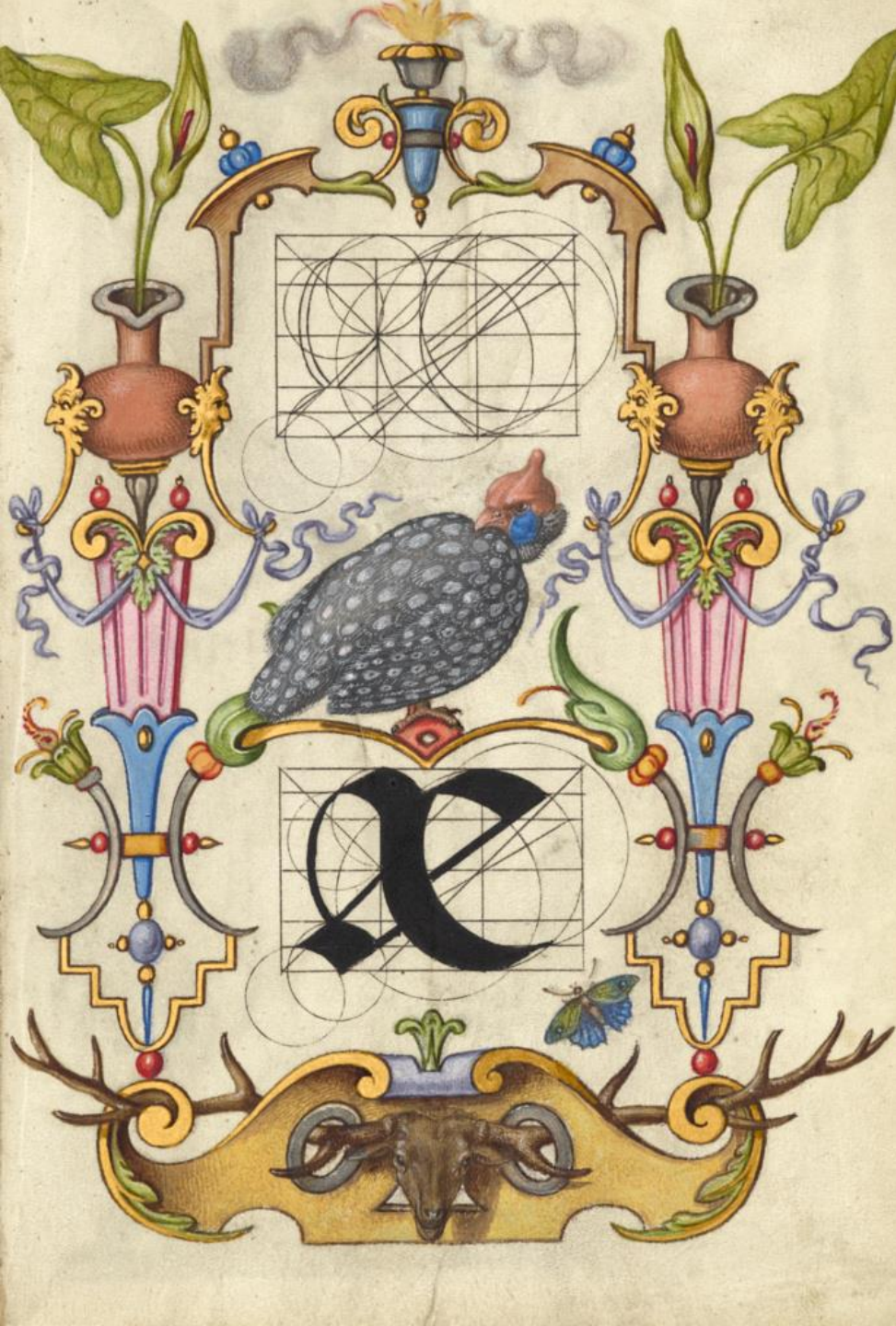
- Consult guidelines to select your file specifications
  - FADGI, NARA, ALA
  - Colleagues at other institutions
- Record in a digitization plan for your institution
- Will help you to digitize consistently over the life of the project
- Can be referred to by people picking up the project later



# Resources

---

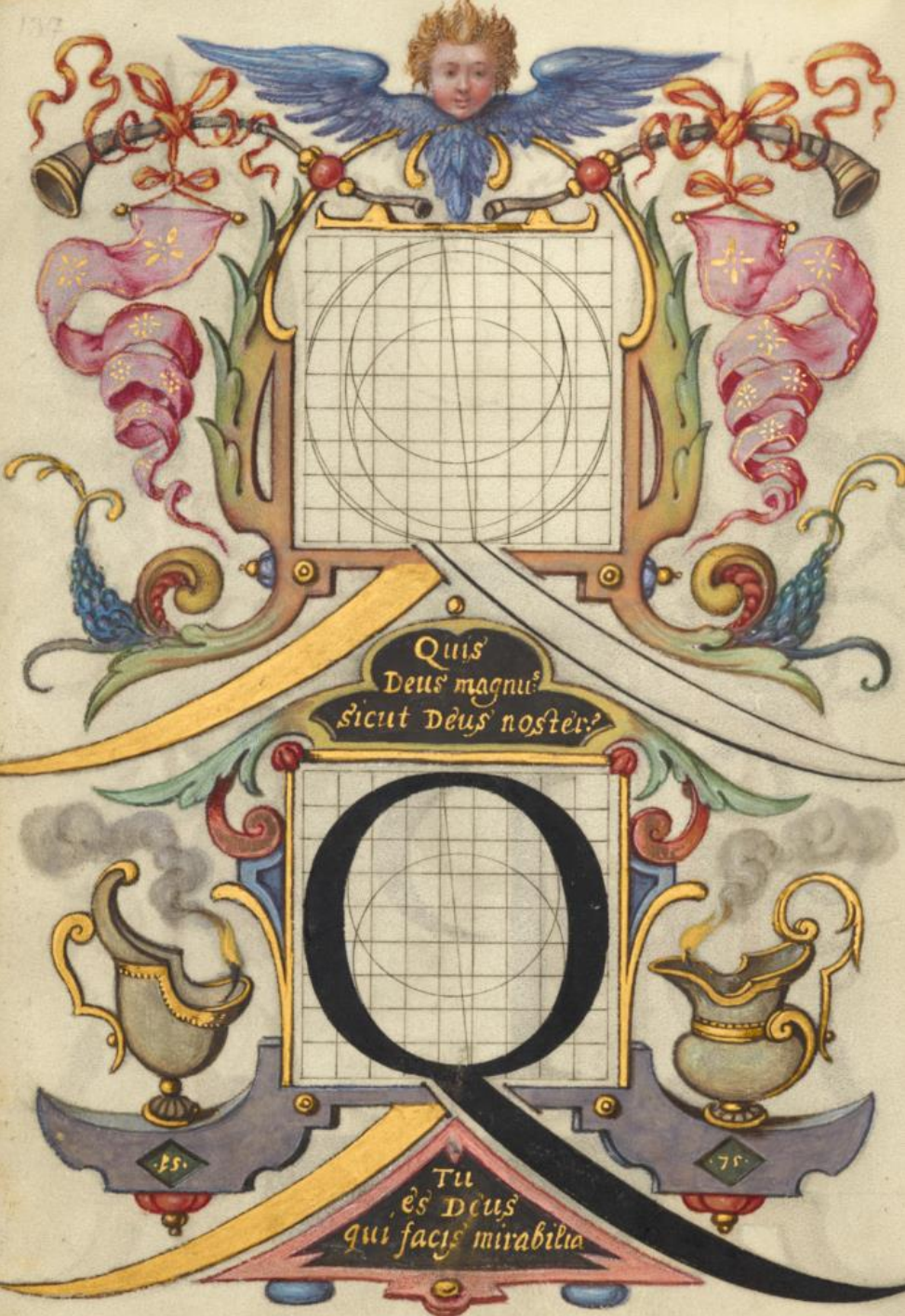
- Examples of institution digitization guides
  - Digital North Carolina: <http://www.digitalnc.org/policies/digitization-guidelines/>
  - Columbia University Libraries Imaging Standards & Procedures: <https://library.columbia.edu/bts/imaging.html>
- FADGI Audio Visual Working Group resources:
  - <http://www.digitizationguidelines.gov/audio-visual/>



# Coming Up

---

- Metadata for Digitization – July 14
- In-house vs Outsourcing & Quality Control – July 21
- Equipment Selection – July 28



# Questions?

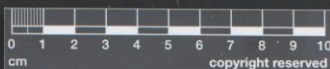
---

Email: [mdowning@ccaha.org](mailto:mdowning@ccaha.org)



## Metadata for Digitization

The Digitization 101 webinar series is a national education program of the Conservation Center for Art & Historic Artifacts provided with support from grants from the National Endowment for the Humanities and the William Penn Foundation.



CARNEGIE MUSEUM  
HERBARIUM (CM)  
PITTSBURGH, PA.  
copyright reserved

CARNEGIE MUSEUM  
HERBARIUM (CM)  
PITTSBURGH, PA.

Nº 483816



BARRENS HABITAT SURVEY

BALSAMINACEAE

*Impatiens pallida* Nutt.  
det: B.L. Isaac 1998

U.S.A.: PENNSYLVANIA: Huntingdon County: 2.9 km ESE of  
Alexandria, Neff Barrens. 40-32-39N, 78-04-04W. Alt.  
275 m. Roadside.

Flowers yellow.

Bonnie L. Isaac 9827  
and Joseph A. Isaac

7 August 1997

HERBARIUM, CARNEGIE MUSEUM (CM), PITTSBURGH, PA.

# Metadata

This session will discuss:

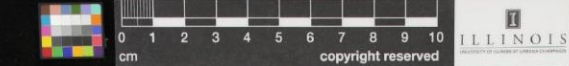
- Definition and importance of metadata & controlled vocabularies
- Creating internal procedures for metadata creation
- Associating metadata with your files
- Collection Management System selection

## Definition of Metadata:

Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource.

- from the National Information Standards Organization (NISO)

*Impatiens pallida*. Carnegie Museum of Natural History Herbarium.



# Metadata Types

---

Some metadata mirrors what is created for physical records

- Descriptive – Describes the item, including title, date, subject, etc.
- Administrative – Copyright and access restrictions

Some metadata is just for digital records

- Technical – Scanner/camera, date created, pixel dimensions, etc.
- Structural – File's relation to other files
- Preservation – Checksums, history of data corruption or recovery

*Impatiens pallida*. Illinois Natural History Survey.

# Importance of Metadata

---

Metadata supports:

- Discovery of the resources through the ability to search
- Provides insight into your collections
- Facilitates sharing resources across institutions
- Helps maintain control over access and restrictions
- Helps manage digital records for the future



Margaret Armstrong, *Lilac Clematis (Atragene occidentalis)*, 1909. Metropolitan Museum of Art.



# Metadata Schema

---

Metadata standards help to ensure that metadata is applied consistently within an institution and across institutions.

Examples:

- Dublin Core Metadata Initiative (DCMI)
- Machine-Readable Encoding (MARC)
- Encoded Archival Description (EAD)

University of Texas Libraries: Metadata Basics

<https://guides.lib.utexas.edu/metadata-basics/intro>



Names from Getty ULAN

**Muybridge, Eadweard (preferred)**

Eadweard Muybridge

Helios

Muggeridge, Edward James

Muybridge Eadweard J.

Muybridge, E. J.

Muybridge, Eadweard J.

Eadweard J. Muybridge

# Controlled Vocabularies

---

Metadata schema are supported by controlled vocabularies.

- Standardized and organized arrangements of words and phrases that provide a consistent way to describe data.

Examples:

- Library of Congress Name Authority File (LCNAF)
- Library of Congress Subject Headings (LCSH)
- Getty Union List of Artist Names (ULAN)
- Getty Art & Architecture Thesaurus (AAT)



# Create a Data Dictionary

---

- Data dictionaries document:
  - Element names and definitions
  - Obligation (required or optional)
  - Guidance on use
- When drafting metadata procedures, consider:
  - What metadata do you currently use?
  - Who is your audience, and what terminologies will they understand?
  - What resources are available to create the metadata?
    - What tools do you have?
    - What skills do your staff and/or volunteers have?
  - Who will you share your metadata with?

### Metadata Field Element Chart

Field Name	Map to Dublin Core Schema	Mandatory for Upload	Best Practice	Controlled Vocabulary
1. <a href="#">Title</a>	Title	Yes	Yes	None
2. <a href="#">Description</a>	Description		Yes	None
3. <a href="#">Creator</a>	Creator		Yes	LCNAF; ULAN; VIAF
4. <a href="#">Subject</a>	Subject		Yes	LCSH; LCNAF; AAT; TGM; MeSH; VIAF
5. <a href="#">Location</a>	Coverage-Spatial		No	LCSH; TGN
6. <a href="#">Contributors</a>	Contributor		No	LCNAF; ULAN; VIAF
7. <a href="#">Publisher of Original</a>	Publisher		No	LCNAF
8. <a href="#">Date of Original</a>	None		Yes	None
9. <a href="#">Hidden Date</a>	Date		Yes	ISO 8601 W3CDT
10. <a href="#">Physical Format</a>	Format		Yes	NY Heritage List
11. <a href="#">Physical Description</a>	Source		Yes	None
12. <a href="#">Local Location</a>	None		No	None
13. <a href="#">Relation</a>	Relation		No	None
14. <a href="#">Type</a>	Type		Yes	NY Heritage List (DCMI)
15. <a href="#">Language</a>	Language		No	ISO 639-2

# Data Dictionary

## Digital Public Library of America (DPLA)

- Review metadata guidelines for your state or regional hub

Example: New York Heritage Data Dictionary, 2019

<https://nyheritage.org/sites/default/files/pages/NYH-MetadataDictionary-V5.pdf>

“How to Create a Descriptive Metadata Plan.” Sustainable Heritage Network. March 16, 2018.

[http://sustainableheritagenetwork.org/system/files/atoms/file/How\\_to\\_Create\\_a\\_Descriptive\\_Metadata\\_Plan.pdf](http://sustainableheritagenetwork.org/system/files/atoms/file/How_to_Create_a_Descriptive_Metadata_Plan.pdf)

# Associating Metadata

**Modify Object**

Accession Number: SHIN.213.65 Temp ID: 123 InActive: ☐ Confidential: ☐ Web Access?: ☒

Display Title: Rooster Hen & Hydrangea Display Artist: Ito Jakuchu Artist Code: I-894

Object Status: Cataloging:Needs Cataloging, Conservation:Excellent Condition, Imaging:Resize for the Ir

Keywords: Eccentric, Hydrangea, Middle Edo, Rooster

Location: Green St. Warehouse : 2nd Floor : Storeroom 1 (moved 7/6/1998)

Page 1 | Page 2 | Page 3 | Notes & Histories | Edited | Surrogate | User Defined Fields | Surrogate List

Accession Number: SHIN.213.65 Temp ID: 123

Display Title: Rooster Hen & Hydrangea

Display Artist: Ito Jakuchu Artist Sort Name: Jakuchu, Ito

No. of Components: 1 Whole? ☒ Yes ☐ No Multiple Artists? ☐ Yes ☒ No

State: original

Creation Date: 1756

Update Range: 1756 - 1756

Alt. Date Type: n/a

Technique: Painting

Medium: Color

Support: Silk

Media & Support: Color on Silk

Artist Code: I-894 Attribution:

Editioned? ☐ Yes ☒ No

Century: 18th Century

Period: Edo Middle

Alt. Date: n/a

Color:

Medium 2:

Support 2:

- Collection management system
  - The Museum System (TMS), Embark, PastPerfect, or ContentDM
- Excel or Access table for smaller-scale projects



# Associating Metadata

---

- Metadata can be embedded into the file itself
  - Technical metadata is embedded automatically
  - Additional metadata can be embedded within the file
    - At the time of capture using scanner or photo processing software
    - Or on the final files using tools such as Adobe Bridge
- Not a requirement if metadata is recorded elsewhere and corresponds with file names
- Added layer of security

# Selecting a CMS

---

- Lots on the market
- Vary in cost, functionality, and level of required staff expertise
- Important things to consider:
  - Open-source or proprietary software
  - Hosting files locally or online
  - Cost and ease of use
  - Ease of integration with current data management system
  - Ease of importing and exporting metadata



# Selecting a CMS

---

- American Association for State and Local History (AASLH)  
Technical Leaflet on Choosing a Collection Management Software  
[https://d221a1e908576484595f-1f424f9e28cc684c8a6264aa2ad33a9d.ssl.cf2.rackcdn.com/aaslh\\_a28f1bb1d5ab7592af01b9bbac64dc95.pdf](https://d221a1e908576484595f-1f424f9e28cc684c8a6264aa2ad33a9d.ssl.cf2.rackcdn.com/aaslh_a28f1bb1d5ab7592af01b9bbac64dc95.pdf)
- Ashley Blewer's "Collection Management System Collection"  
<https://bits.ashleyblewer.com/blog/2017/08/09/collection-management-system-collection>



# Document Metadata Decisions

---

- Data dictionary and metadata plan for your institution
- Procedures for metadata entry
- Include or reference in your digitization plan



# Resources

---

- Sustainable Heritage Network Resources  
<http://sustainableheritagenetwork.org/digital-heritage>
- Riley, Jenn. "Understanding Metadata: What is Metadata, and What is it For?" National Information Standards Organization, 2017.  
[https://groups.niso.org/apps/group\\_public/download.php/17446/Understanding%20Metadata.pdf](https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf)



Margaret Armstrong, *Lilac Clematis (Atragene occidentalis)*, 1909. Metropolitan Museum of Art.



# Coming Up

---

- In-house vs Outsourcing & Quality Control – July 21
- Equipment Selection – July 28



# Questions?

---

Email: [mdowning@ccaha.org](mailto:mdowning@ccaha.org)



## In-house vs. Outsourcing + Quality Control

The Digitization 101 webinar series is a national education program of the Conservation Center for Art & Historic Artifacts provided with support from grants from the National Endowment for the Humanities and the William Penn Foundation.



# What We'll Discuss

---

- Where to do the actual digitization of your material:
  - In-house with purchased equipment
  - Outsourcing to a digitization vendor
- What to look for in a vendor
- Quality Control Procedures
  - Image Quality
  - Metadata



# Decisions, decisions...

---

- Scanning is only about 1/3 of project time in a digitization project.
- Cataloging description, indexing, preparation, and project management make up the other two thirds.
- If you're applying for grants, this staff time can often count toward in-kind fundraising requirements.

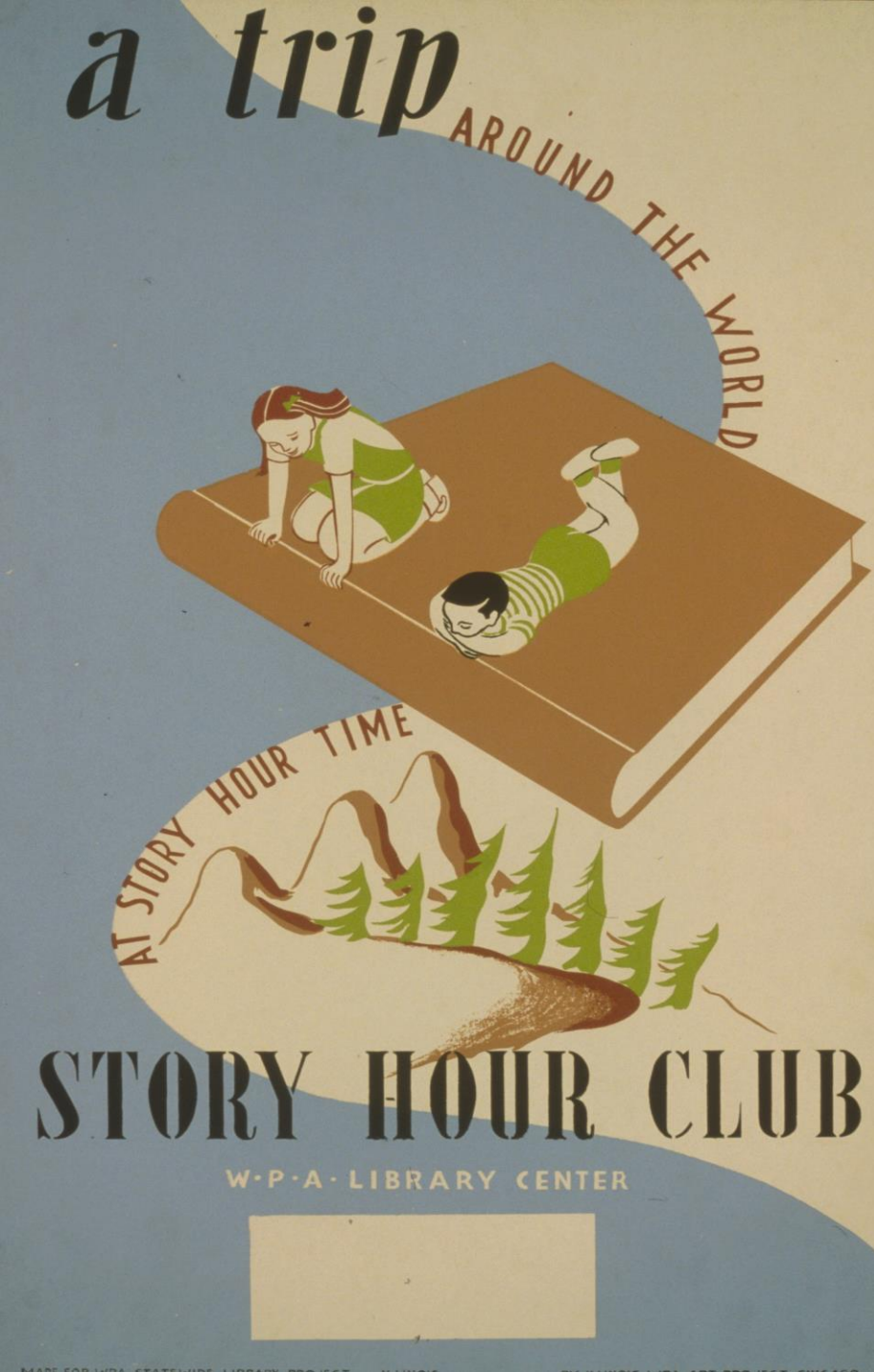
# Books for everyone



## Digitizing In-House

---

Pros	Cons
Have more control over process	Requires lots of resources: staff training and equipment purchase
In-house expertise on hand	Need to budget for equipment repairs and upgrades
Security	Need for dedicated space
Learn new skills	Potentially limited staff expertise
Once studio is set up, can establish an ongoing digitization operation	Responsible for all quality control



# Outsourcing Digitization

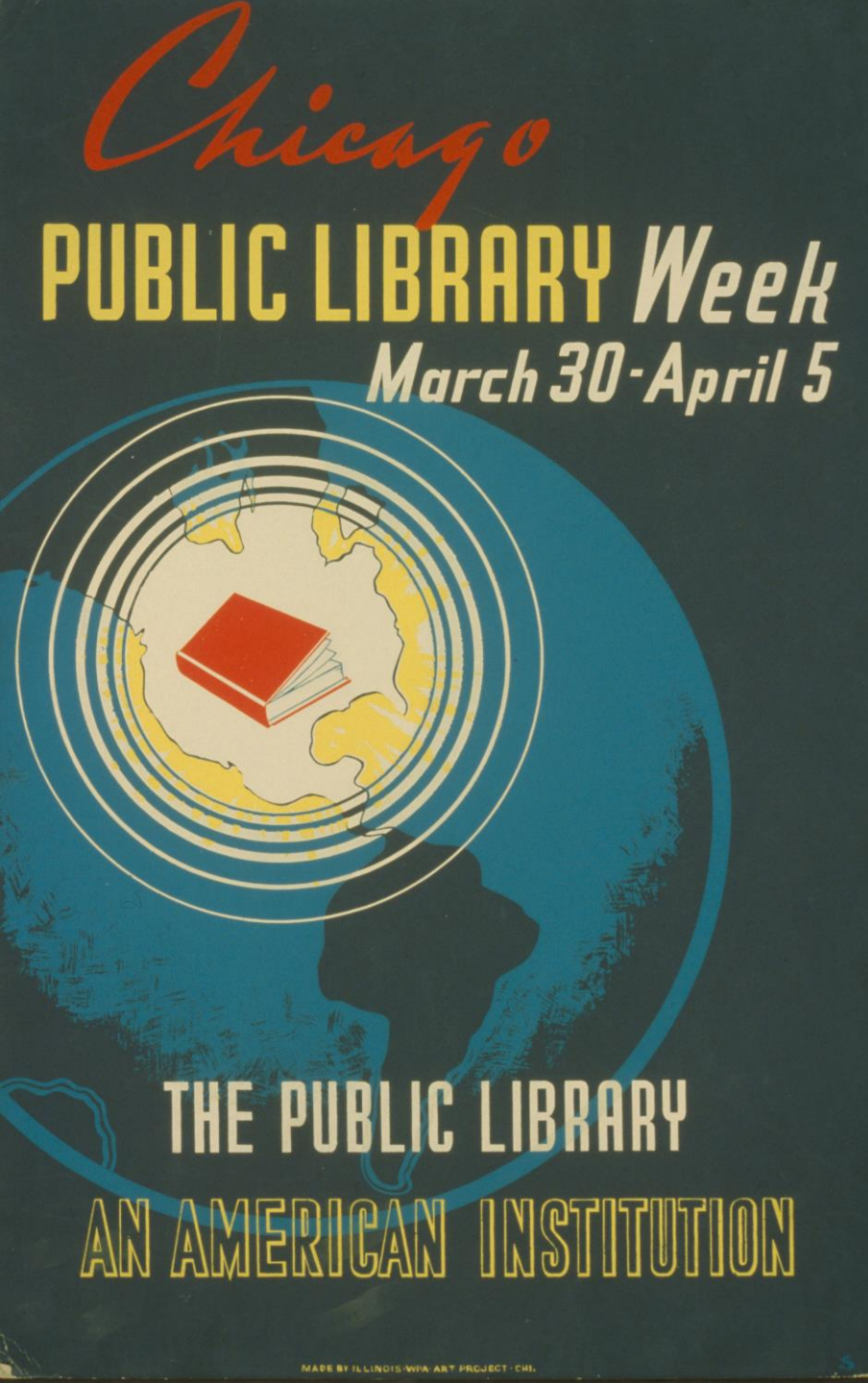
Pros	Cons
Faster	Less flexibility, project scope must be clearly established, and all decisions made ahead of time
No in-house technical expertise required	Must have very clear communication
High production levels	Risk to originals in transportation
Vendor provides quality control	
Don't have to pay for equipment or training	



# What To Look For in a Vendor

---

- Written quality control procedures
- Staff expertise in handling fragile materials
  - Conservators on staff if necessary
- References and sample projects
- Security and safety measures to protect objects in their car
  - Secure, fire-resistant, climate-controlled storage
  - Climate-controlled lab and digitization studio spaces
  - No smoking, eating, or drinking in areas where documents will be handled



# What to Include in an RFP

---

- Project timeline
- Can work be done at vendor facility or on-site?
- Number and size of originals
  - Approximate number of pages or images to capture
- Type of materials
  - Are materials bound, flat, or a mix?
- Condition issues or special handling notes
  - Are materials in mats or encapsulation?
  - Are there any fasteners like paper clips or staples?
- File specifications
  - File naming, cropping, resolution, bit depth, color, etc.
- Metadata requirements
- Do you to include packing and transportation costs?



# What is Quality Control?

---

- ☐ Quality Control (QC) or Quality Assurance (QA)
  - ☐ Processes used to ensure that digitization and metadata creation are done properly.
- ☐ Address specific issues, identify accuracy requirements, and acceptable error rates for all aspects evaluated
- ☐ A quality control program should address :
  - What elements you're looking at when reviewing materials
  - What percentage of files you'll review
  - What causes a file to fail QC, and procedures for fixing
- ☐ NARA Technical Guidelines for Digitizing Archival Materials  
<https://www.archives.gov/files/preservation/technical/guidelines.pdf>

# CCAHA IMAGING QUALITY CONTROL CHECKLIST

ACCESSION #:  
OWNER:

## 100% Check of RAW Files

Inspection	Project Requirements	Initial / Date	Notes
<b>Document Related</b>			
Orientation			
Proportion / Distortion			
Skew / Rotation			
Cropping			
Imaging of blank pages			
Scale reference (if present)			
Background color			
<b>Metadata Related</b>			
Named properly			
<b>Image Quality Related</b>			
Tone - Brightness, Contrast, Clipping			
Color accuracy			
Noise / Artifacts			
Sharpness			
Even tonal values			
<b>General QC Notes</b>			

# QC: Image Quality

- ☐ CCAHA procedures for large-scale projects include an initial inspection and a final inspection
- ☐ Initial inspection
  - ☐ Correct file names
  - ☐ Orientation
  - ☐ Distortion
  - ☐ Scale reference (if required)
  - ☐ Cropping
  - ☐ Background color
  - ☐ Even tonal values and contrast
  - ☐ Completeness

#### 10% Check of Final Files

Inspection	Project Requirements	Initial / Date	Notes
<b>File Related</b>			
Files Open and Display			
Proper Format			
Compression (if required)			
Color Mode			
Bit Depth			
Color profile			
Layers (present if desired)			
<b>Document Related</b>			
Correct Dimensions			
Spatial Resolution			
Orientation			
Proportion / Distortion			
<b>Metadata Related</b>			
Named properly			
Data in header tags			
Descriptive metadata			
Technical metadata			
Administrative metadata			
<b>Image Quality Related</b>			
Tone			
Color accuracy			
Noise / Artifacts			
Sharpness			
Even tonal values			
<b>File Delivery</b>			
All files on drive			
Files Open and Display			
Drive is labeled properly			

# QC: Image Quality

---

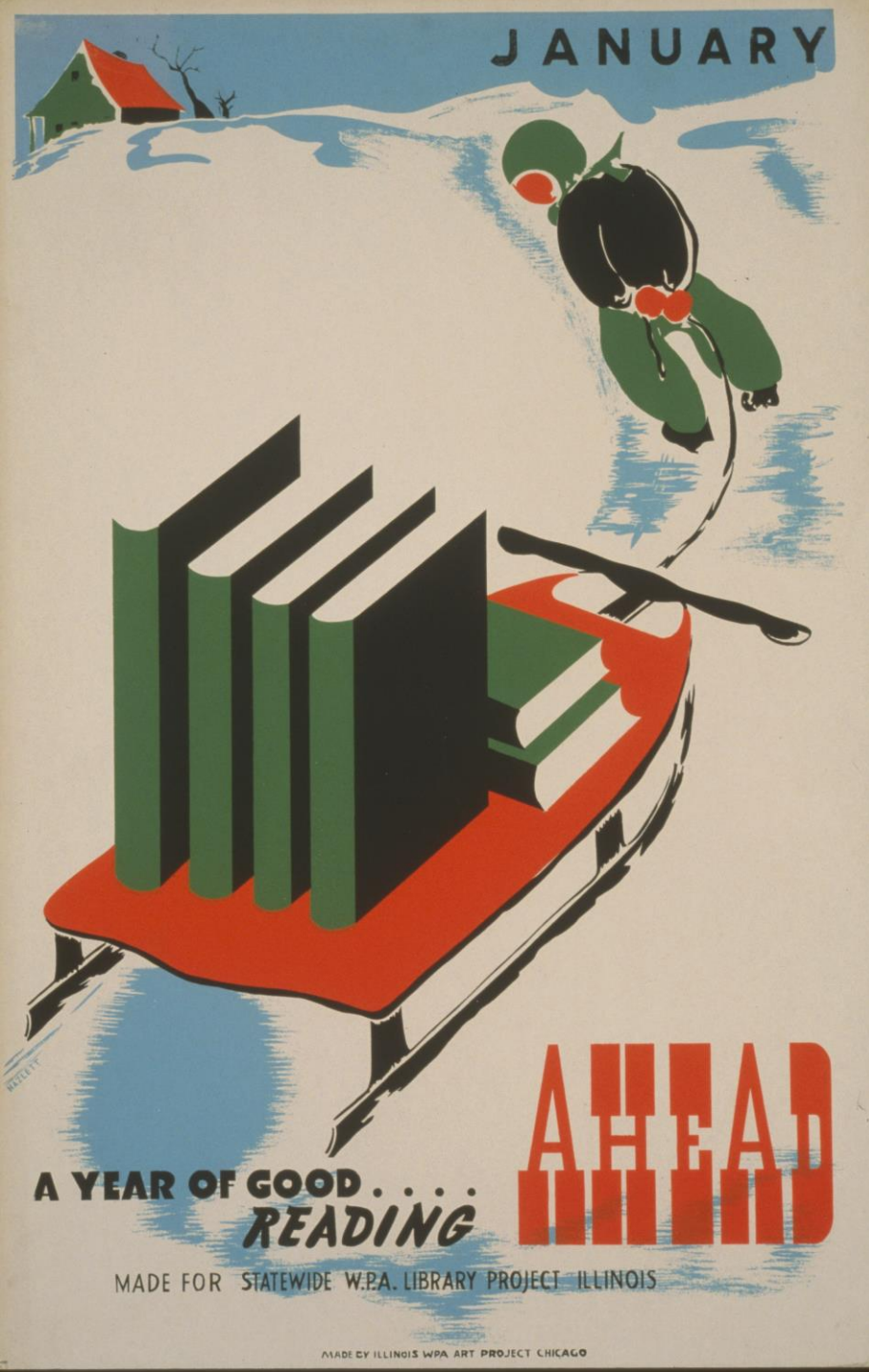
- 10% inspection of final files
  - File format
  - Compression
  - Color mode
  - Bit Depth
  - Dimensions
  - Resolution
  - Metadata (as needed)
  - Artifacts (physical or digital)
  - Focus and overall image quality
- If a file fails, others around it are inspected.
  - If more than 1% of inspected files fail, then the entire batch is inspected, failed files are corrected.



# QC: Metadata

---

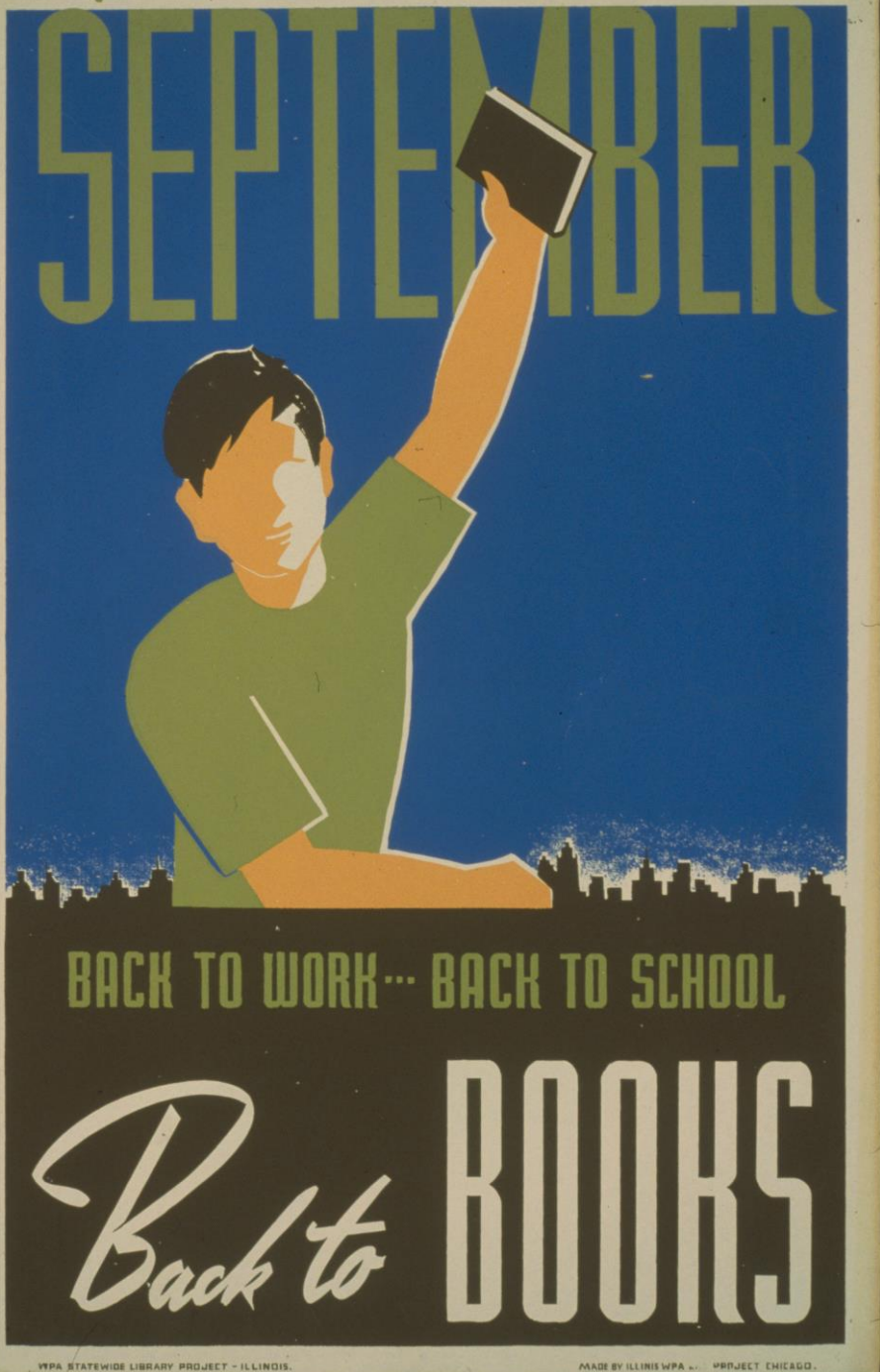
- ☐ Less defined in evaluation metrics
  - ☐ Adherence to standards set by institutional policy
    - ☐ Guided by data dictionary
  - ☐ Relevancy and accuracy of metadata
  - ☐ Consistency in the creation of metadata
  - ☐ Consistency and completeness in the level at which metadata is applied
  - ☐ Representation of different types of metadata.
    - ☐ Has sufficient descriptive, technical, and administrative metadata been provided?
- ☐ NARA Technical Guidelines for Digitizing Archival Materials  
<https://www.archives.gov/files/preservation/technical/guidelines.pdf>



# Next Steps

---

- ☐ Decide whether you want to digitize in-house or outsource to a vendor
- ☐ Draft written Quality Control procedures
- ☐ Start requesting estimates from vendors or planning for equipment purchase



# Questions?

---

Coming Up

- Last webinar in series: Equipment Selection – July 28

Email: [mdowning@ccaha.org](mailto:mdowning@ccaha.org)



## Equipment Selection

The Digitization 101 webinar series is a national education program of the Conservation Center for Art & Historic Artifacts provided with support from grants from the National Endowment for the Humanities and the William Penn Foundation.



# What We'll Discuss

---

- Types of Equipment
  - Flatbed scanners
  - Camera and copystand
  - Specialty scanners
- Software
- Targets and calibration



# Flatbed Scanners

Pros	Cons
Small, can fit on a desk in a small space	Scan time can be slow
Enclosed environment	Not ideal for bound material
Affordable	Not ideal for friable media
Easy to learn	Limited to size of scan bed
Good for collections of small, unbound, flat material, with stable surfaces	
Good for photographic negatives and transparencies	



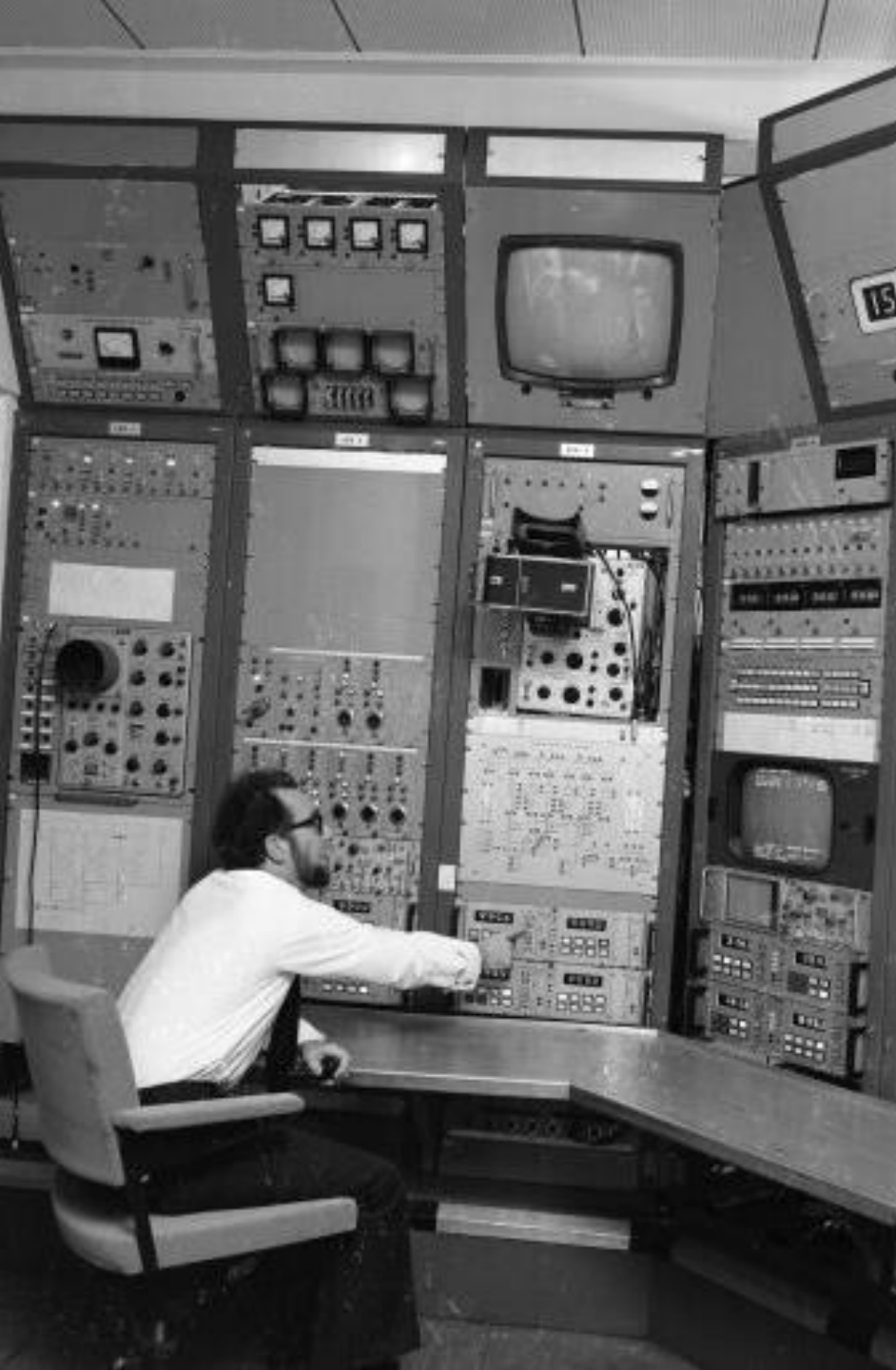
# Scanning Software

---

- Third-party scanning software such as Silverfast or Vuescan
  - Provides customization in scanning options
  - Comes with built-in calibration software
- Image editing and processing software
  - Creative Cloud applications – Lightroom, Bridge, Photoshop
- Optical Character Recognition (OCR)
  - Adobe Acrobat Pro or Abbyy FineReader

Discounted rates for non-profits through TechSoup:

<https://www.techsoup.org/>



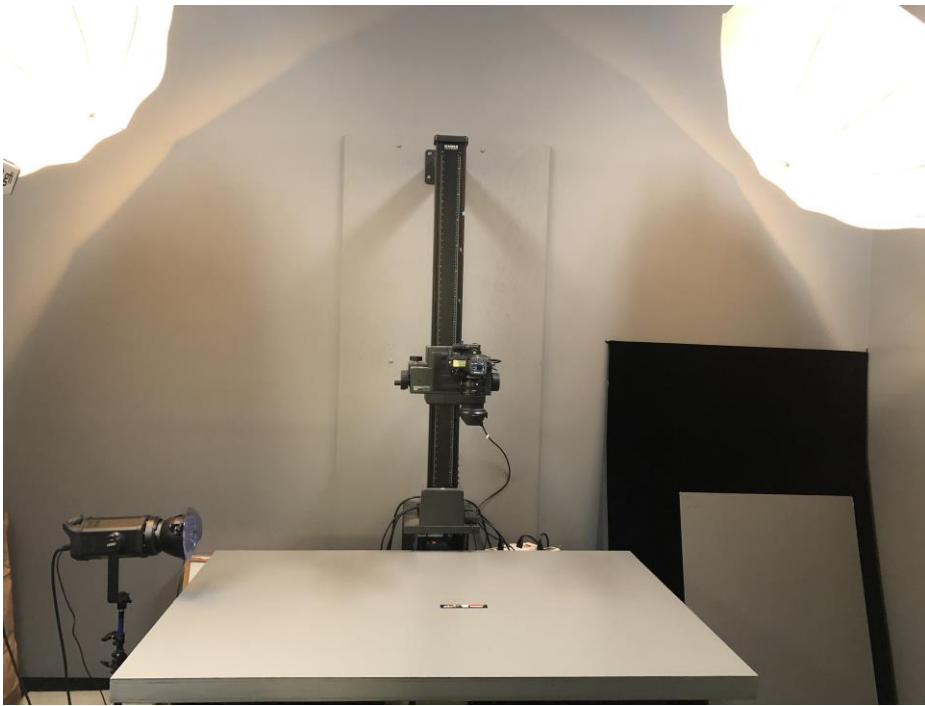
# What to Look For in a Scanner

---

- Size of scan bed
- Resolution
- Bit depth and color space options
- Ability to scan transparencies
- Compatible with third-party scanning software
- Return policy
  - Test out scanner with a range of materials

Library of Congress Tip Sheet: What to Look for in a Scanner

- <https://www.loc.gov/rr/print/tp/LookForAScanner.pdf>



CCAHA Studio 1

# Camera / Copystand

---

- Components
  - 2 lights
  - Camera wired to a computer
  - Copystand
  - Table surface
  - Book cradle or easel



Linhof Book Copying Easel



DT V Cradle



# Camera / Copystand

Pros	Cons
Very versatile	Expensive
Speedy capture	Higher learning curve
Modular setup	More variables that affect image quality
No contact with artifact, good for fragile items	
Good for bound materials and objects	
Can be adapted easily to capture negatives or objects	

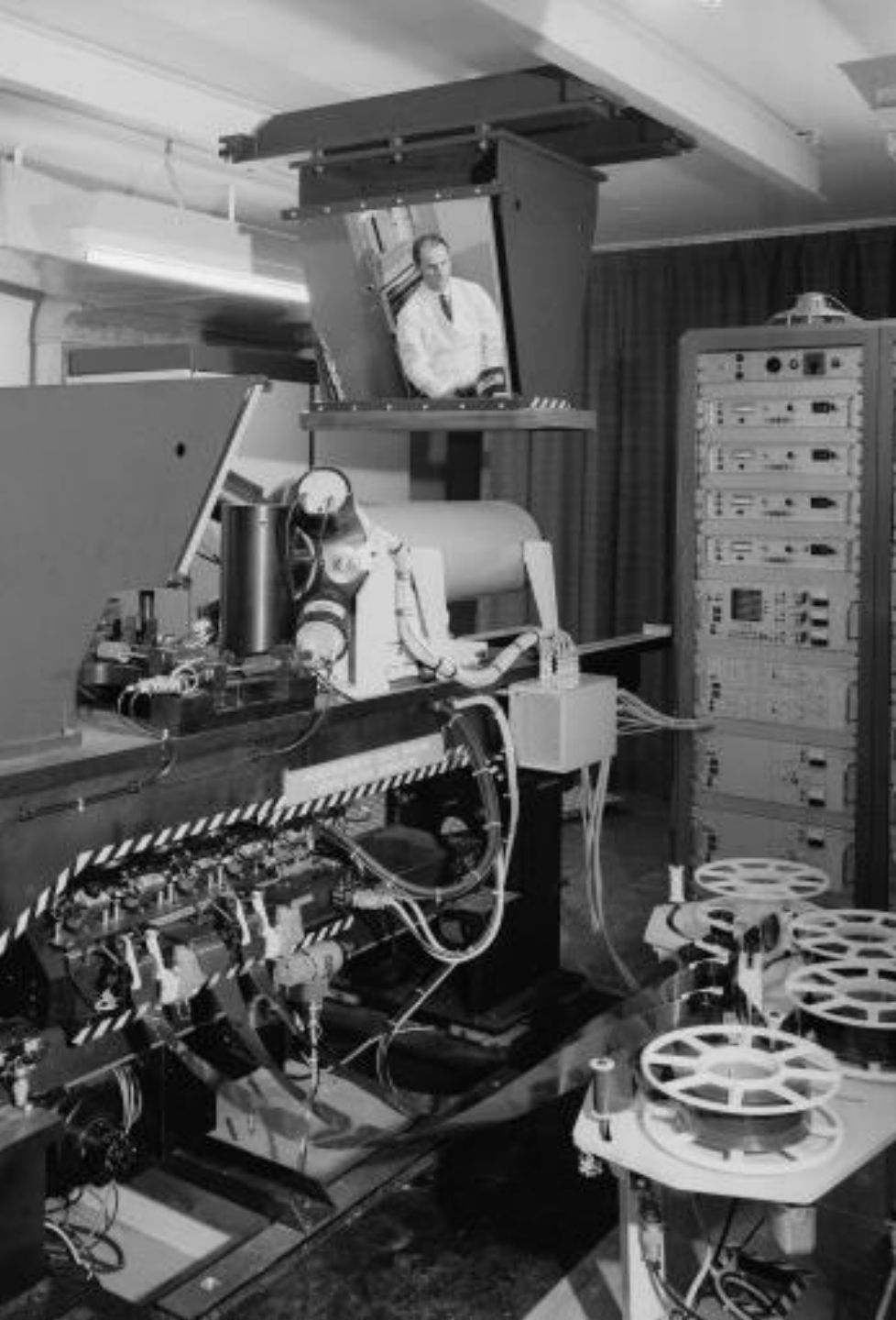
CCAHA Photographer imaging a parchment document



# Supporting Software

---

- RAW processing software
  - Can apply adjustments to images in batches, crop, add metadata
- Adobe Creative Cloud
  - Photoshop, Lightroom, Bridge, and Acrobat Reader and more



# What to Look For in Equipment

---

- Camera
  - Camera body – High resolution – Look at pixel dimensions
  - Lens –sharp and even in exposure
  - Ability to test camera and lens, in your own studio if possible
    - Check for even focus and exposure
    - Check for digital artifacts
- Copystand
  - Secure and sturdy, motorized if possible
- Lighting
  - Can either be continuous illumination or strobe
  - Well-balanced and even



Cruse table scanner



BookEye book scanner

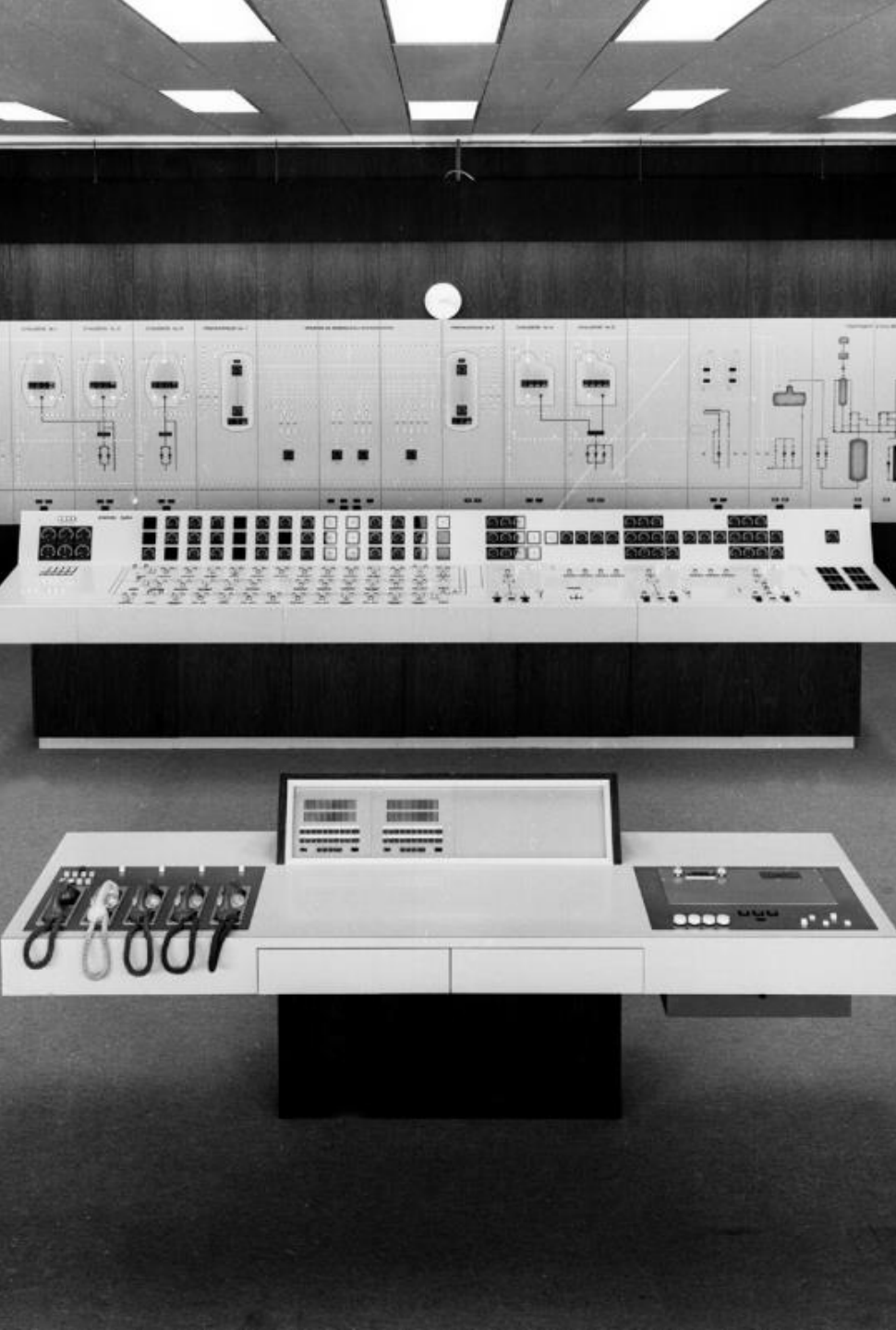


Plustek film scanner

# Specialty Scanners

---

- ❑ Oversize scanners
  - Materials pass under scanning element on large table
- ❑ Book scanners
  - Overhead scanners
  - Good for large book collections or libraries
- Film Scanners
  - High quality scans, can be faster than flatbed scanners



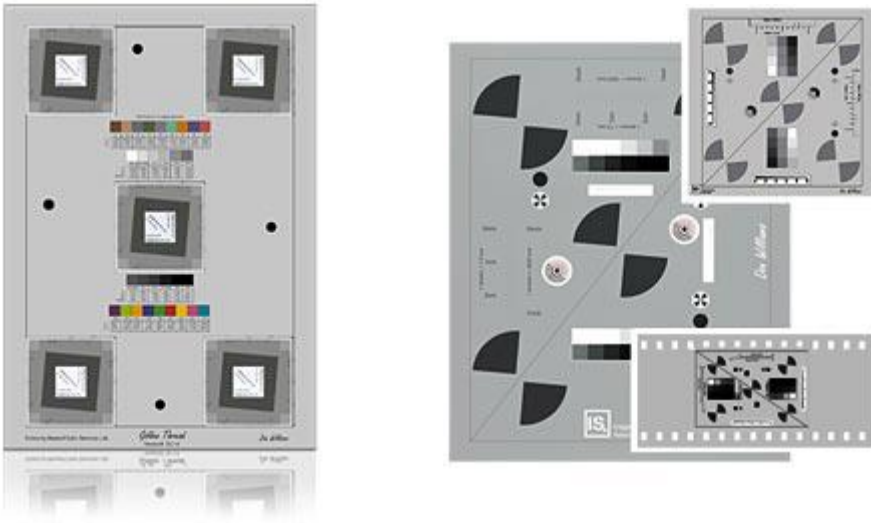
# FADGI Metrics

---

- Federal Agencies Digitization Guidelines Initiative (FADGI) Guidelines have specific metrics for equipment calibration
- Metrics designed to ensure that digital surrogates are scanned at a certain proven level, important for government agencies
  - Tone Response
  - White Balance
  - Illuminance Uniformity
  - Color Accuracy
  - Contrast
  - Scale Accuracy
  - Sharpening

# DICE Targets and Software

---



Targets from Image Science Associates

- Specially designed targets for assessment
- Software tools analyze the targets
  - Open DICE – free and open source, developed by Library of Congress <http://www.digitizationguidelines.gov/guidelines/digitize-OpenDice.html>
  - Golden Thread – proprietary, sold by Image Science Associates <http://www.imagescienceassociates.com/>

Image Access resource on Quality Controlled Scanning

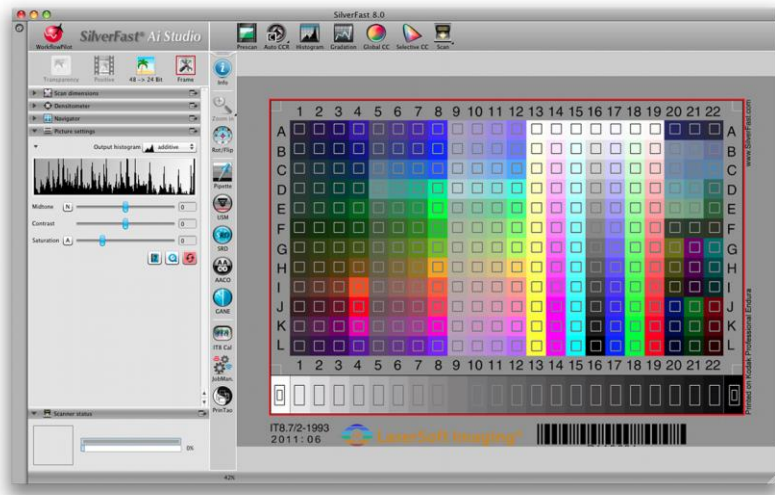
[https://support.imageaccess.de/downloads/product\\_manuals/FAQ/FAQ-Quality-Controlled-Scanning.pdf](https://support.imageaccess.de/downloads/product_manuals/FAQ/FAQ-Quality-Controlled-Scanning.pdf)



X-Rite ColorChecker

# Color Calibration

- Color calibration target
  - Assigned values for each color
- Can create custom profiles automatically
- Adjust in image editing programs by adjusting white balance and exposure
- Best to calibrate equipment as much as possible before making digital adjustments



IT8 color calibration in Silverfast



# Successful Digitization

---

- FADGI compliance not easy to achieve
  - Specialized targets and software
  - Top of the line equipment
- Don't be discouraged from digitization
- Things you can do to ensure a successful digitization program:
  - Color calibration
  - Visual assessment
  - Checking for correct file specifications
  - Metadata and cataloging
  - Quality control program
  - Ongoing digital collection management



# Wrapup and Questions

---

- What topics would you like to see in future sessions?
- Email: [mdowning@ccaha.org](mailto:mdowning@ccaha.org)