



Smithsonian  
*National Museum of Natural History*

# Bridging the Digital Divide: Basic Guidelines & Best Practices for digitization projects

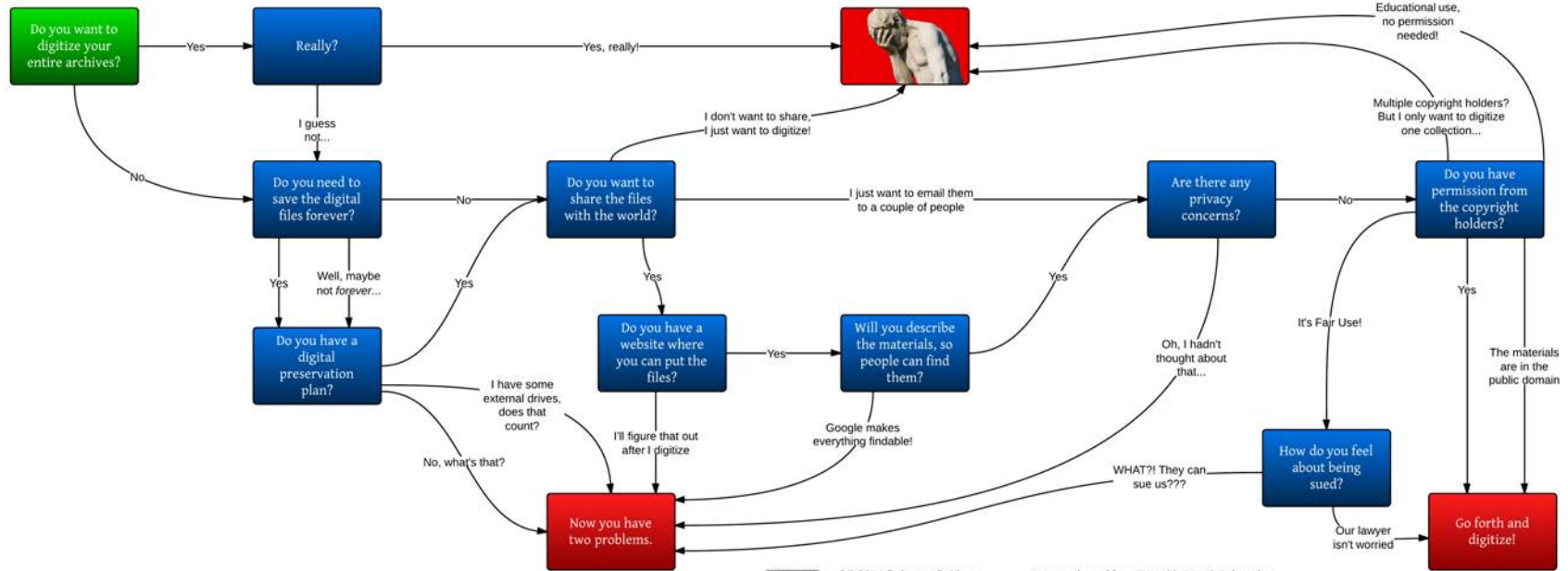
Jeanine Nault  
Digital Imaging Specialist  
National Anthropological Archives  
National Museum of Natural History  
Smithsonian Institution  
naultj@si.edu

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# Will digitization solve my problem?

## Will digitization solve my problem?



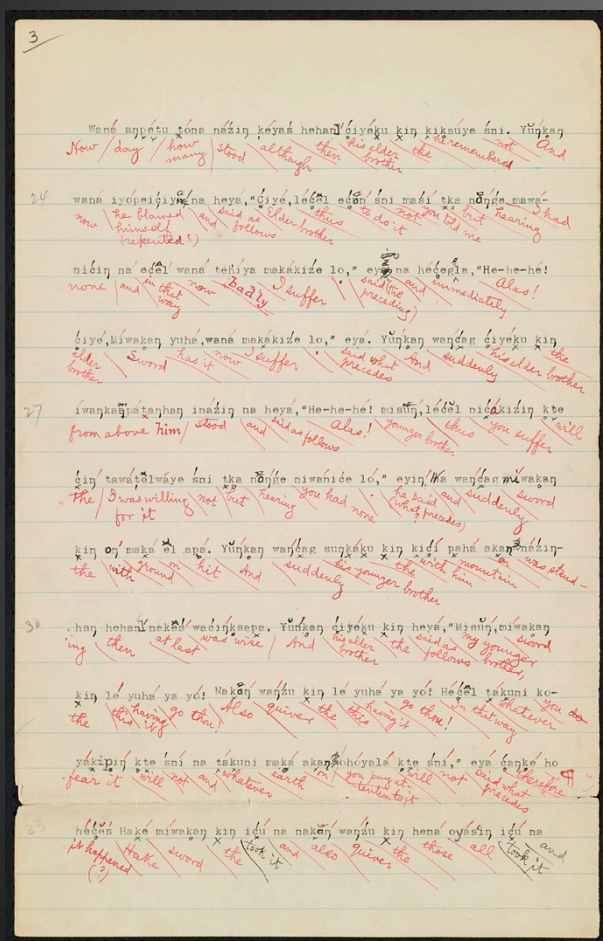
CC 2014 Rebecca Goldman  
Derangement and Description  
derangementanddescription.wordpress.com

Image adapted from "Head in Hands" photo by  
Alex E. Proimos. Photo at  
<http://www.flickr.com/photos/proimos/4199675334/>



Smithsonian  
National Museum of Natural History

# National Anthropological Archives Arcadia project



[NAA MS 4800 \[104\]](#), National Anthropological Archives, Smithsonian Institution



"Experts estimate that more than 50 percent of the world's living languages will cease to be spoken by the end of this century, and with them, the ability to confront worldwide environmental challenges will be diminished. The digitization project funded by Arcadia aims to conserve this knowledge, which is currently recorded on materials that are at high risk of degrading. The project team plans to create digital surrogates of voice recordings and paper documents in the NAA, and make them publicly available to support researchers and communities struggling to research, document and revitalize indigenous languages and cultures. Online access will make the material widely available for use without damage to the historic originals."

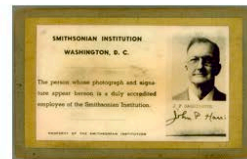


# John P. Harrington Collection web portal

<http://anthropology.si.edu/naa/harrington/index.html>

Harrington Home    Explore by Media    Explore by Map    Explore by Language

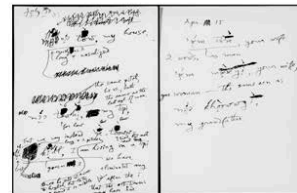
## Accessing the Collection



The Harrington collection is massive and contains several different types of media. Some material was deposited in the collection soon after it was created, but the majority did not arrive in the archives until after Harrington's death. Components have been processed at varying levels over a number of years. Some parts of the collection were duplicated to make them more widely accessible, utilizing the best technology available at different times. Digitization is the latest chapter in that ongoing effort.

The very size and depth of the collection can make it difficult to fully identify the range of materials on a given topic or to make connections across different media. This website is being developed as a means to reconnect parts of the collection with each other as well as to help scholars and communities connect with these archival materials. It is a work in progress.

This site provides two ways to discover what's in the Harrington collection - either through a map interface or through inventories of each media type. Both provide links to digital resources that are being added as they are produced. The entire collection is available onsite at the National Anthropological Archives.



## Explore the Collection by Media

- [Manuscripts and Microfilm](#)
- [Photographs](#)
- [Natural History Specimens](#)
- [Sound Recordings](#)
- [Material Culture Collection](#)



## Explore the Collection by Map

- [Native Languages of North America](#)
- [California Languages](#)



# Beyond slap & scan

## Before beginning a digitization project, consider the following:

- WHY are we digitizing? (The lifecycle of the digital file – Where did it come from? Where is going?)
- WHAT are we digitizing? (Are there different types of materials with different needs?)
- HOW are we digitizing? (Equipment, software, people, time, money, space)
- WHO are we digitizing for? (Who will use these files? What will they do with it?)

## Digitization projects are more than just slapping photos on a scanner

- Project planning and management
- Selection and Prioritization
- Technical Specifications
- Funding
- Preservation and Access
- Rights Issues
- To name just a few!



# Digitization Then and Now

## Then

- Microfilm as nationally accepted standard for decades – cost effective way to preserve text materials
  - But what about graphic materials like paintings and photos?
- Copy negatives – expensive and two generations removed from original
  - Original object → New negative → Copy print

## Now

- Digital imaging and digitization allows for preservation of original object (less handling, high resolution file for detailed viewing) and increases access to collections



# Then and Now

<p>Syl. Hin. madse.'i'tsi, chief.  = Crow. madse.'its'e. carefully compared. From Crow 'its'e. <del>crow</del> <del>its'e</del>, good, this is the old word, but the Hin. have is Hin. dsaki, good; dsaki-ts, it is good. Hib. madse', man. madse', man, has the same sound that you got in the</p>	<p>Syl. We still keep the Hib. word for chief is madse.'i'tsi, chief, lit. good-man = Crow madse.'its'e, chief, lit. good-man. Ep. Moj. kwaloita. The Hib. now say mats'e-tzaki, = an ordinary good-looking man (not a chief, &amp; we say tsaki' for good, but the Crows still say 'its'e', good.</p>
---	--

John Peabody Harrington papers: Mandan/  
Hidatsa/Crow, 1950-1951  
Image name: Accession #1976-95 [NMNH-  
Harrington\_mf5\_r14\_0002]

23

u m m w d m o w t t -  
m m e t t e m m m k o -  
m m m w d m o w e  
e t t e - o m m m k o - w i k o  
m m e t t e - k o - k y e e t t e - t e  
d a m m - a s w e e t t e - e m a  
m a - k e b e w e e t t e - m  
k a b a w e e t t e - e t t e w i m k o -  
i g e w e

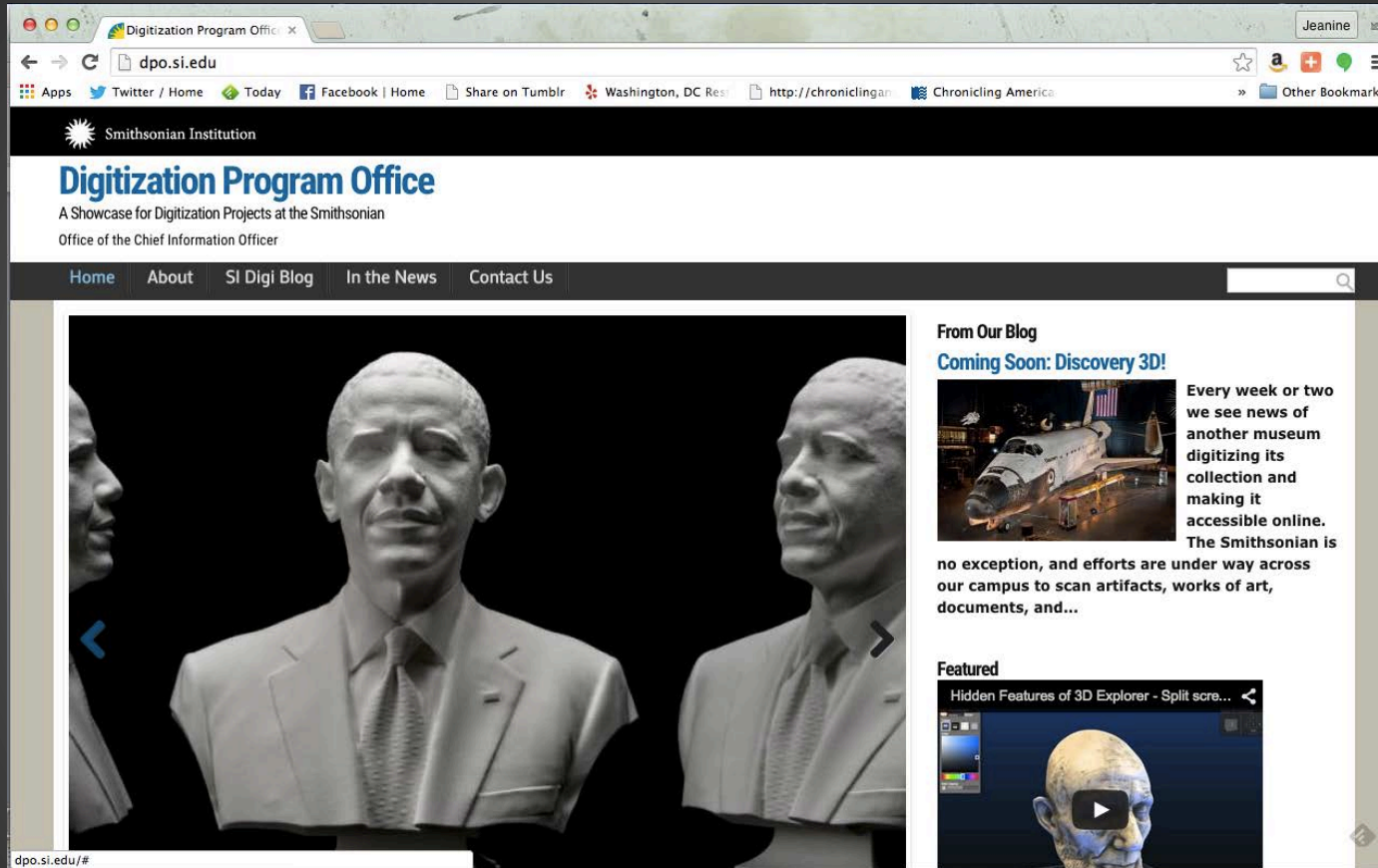
ya w o - m m m e k o - e t t e  
k a b e w a e t t e - m m m e t t e - k o -  
i m m m k o - k e m e t t e - e m e  
k e w e e t t e - m m m e t t e - k o -  
k e w a w e m m m o - e t t e - k y e  
e t t e - i m m m k o - e m m m a  
w a e t t e - k e w a k o - m m k o  
i m m m k o - i m m m k o - e  
m m m e t t e - o m m - k  
k o w i - m m m m k o m m  
i - e k y e t t e - k y a k o t e  
k a w a e t t e - i g e w a k o -  
m m m m k o m m m m m m m  
m o m m k o -

NAA MS 2046, Fox text by Alfred  
Kiyana on Buffalo dance,  
undated



Smithsonian  
National Museum of Natural History

# Digitization Now



- Smithsonian Digitization Program Office Rapid Capture Digitization Process

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*





# FADGI: Federal Agencies Digitization Guidelines Initiative

[Glossary](#) | [Participants](#) | [Contact Us](#)

Search this site



Federal Agencies  
Digitization Guidelines Initiative

Your Comments  
Submit comments on guidelines or on our digitization efforts in general.

[Home](#)

[About](#)

[Guidelines](#)

[Resources](#)

[Review Process](#)

[News & Events](#)

## About This Initiative

Started in 2007, this is a collaborative effort by federal agencies to define common guidelines, methods, and practices for digitizing historical content. As part of this, two working groups are studying issues specific to two major areas, Still Image and Audio Visual.

[Learn more about the initiative »](#)

### Still Image Working Group

This group is involved in a cooperative effort to develop common digitization guidelines for still image materials.

### Audio-Visual Working Group

The goal for this working group is to identify, establish, and disseminate information about standards and practices for the digital reformatting of audio-visual materials.

## Digitization Guidelines

### File Format Comparison Projects

Posted by both Working Groups | December 2, 2014

Comparisons of file formats for *still-image reformatting*, *video reformatting*, and *preservation-oriented treatment of born digital video*. For the reformatting examples, matrixes are used to compare formats in terms of about forty factors. For born digital video, current practices in six federal agencies are described in case histories. The offerings include narrative explanatory documents and supplementary reports.

### MXF Application Specification for Moving Image and other Audio-Visual Content

Review draft for comment | Comments requested by November 10, 2014

The MXF Application Specification for Archiving and Preservation is a detailed specification for a file "wrapper" intended to serve reformatting programs in which an archive is making a digital-file copy of a videotape or motion picture film, as well as to serve as a container for born-digital content when the native encoded essence is judged suitable for retention for at least the medium term.

### Audio Analog-to-Digital Converter Performance

Approved by Working Group | August 20, 2012

This guideline concerns metrics and measurement methods for analog-to-digital converters. Future performance-related documents will discuss the problem of interstitial errors, i.e., accidental loss or transformations of audio samples within the digitizing system before the data stream is written to file.

### Embedding Metadata in Broadcast WAVE Files

Approved by Working Group | April 23, 2012

Print Subscribe Share/Save

## News & Events

- [Review drafts](#) for MXF Application Specification AS-07 including Baseband Shim posted at AMWA website; comments requested by November 10, 2014.
- FADGI welcomes Tom Rieger as the new coordinator for the Still Image Working Group.
- [File format comparisons: still image formats, formats for reformatted video and born digital video exploration](#)  
Documents posted September 9, 2014

## Recent blogs:

- [Let's Start at the Very Beginning: Guiding Principles for Creating Born Digital Video](#) (LC, February 2014)
- [AV Artifact Atlas: By the People, For the People](#) (LC, January 2014)
- [Can I Get a Sample of That? Digital File Format Samples and Test Sets](#) (LC, December 2013)
- [On the Road with FADGI: Recent Conference Presentations Highlight Current Audio and Video Projects](#) (LC, November 2013)
- [Connecting Communities: FADGI Still Image Working Group's Impact on the Library of Congress and Beyond](#) (LC, November 2013)
- [One Format Does Not Fit All: FADGI Audio-Visual Working Group's Diverse Approaches to Format Guidance](#) (LC, October 2013)

[See all News & Events »](#)

## Resources



Smithsonian  
National Museum of Natural History

# Materials: What are we working with?



Paradigm of Verb Forms (Infinitive construct)				Root sé, separate Dakota (road), etc.			
Simple	Instrumental	Instrumental	Abdote	Abdote	Instrumental	Instrumental	Instrumental
ba-wat	i-ba-wat	i-ba-wa-wa	i-ba-wa-wa	wa-bat	hi-pwat	i-ki-pwat	gi-pwat
bi-wat	i-bi-wat	i-bi-wa-wa	i-bi-wa-wa	wa-wat	hi-pwat	i-ki-pwat	gi-pwat
ba-wat	i-ba-wat	ba-wa-wa	i-ba-wa-wa	wa-wat	hi-gtat	i-ki-gtat	gi-gtat
ti-wat	i-ti-wat	ti-wa-wa	i-ti-wa-wa	wa-wat	hi-gtat	i-ki-gtat	gi-gtat
ga-wat	i-ga-wat	ga-wa-wa	i-ga-wa-wa	wa-wat	hi-gtat	i-ki-gtat	gi-gtat
ma-wat (1)	i-ma-wat	ma-wa-wa	i-ma-wa-wa	wa-wat	ma-wat	ma-i-ki-wat	ma-gi-wat
na-wat (2)	i-na-wat	na-wa-wa	i-na-wa-wa	wa-wat	ma-wat	ma-i-ki-wat	ma-gi-wat
na-wat	i-na-wat	na-wa-wa	i-na-wa-wa	wa-wat	hi-nat	i-ki-nat	gi-nat
wa-fa (to make)	ig-wa-fa	ig-wa-fa	ig-wa-fa	wa-danbe	hi-tan-be	i-ki-tan-be	gi-tan-be
danbe (to use)	ig-danbe	ig-danbe	ig-danbe	u-tin	u-ki-tin	i-ki-tin	u-gi-tin
u-tin (to strike)	i-tin	u-tin-tin	i-tin-tin	ja-ke	ja-ki-ke	ja-ki-ke	ja-gi-ke
ja-ke (to strike)	i-ja-ke	ja-ka-ka	i-ja-ka-ka	wa-ja-ke	ja-ki-ke	ja-ki-ke	ja-gi-ke
aga-ka-de (to walk)	ig-aga-ka-de	aga-ka-ka-de	i-aga-ka-ka-de	—	a-ki-gta-hade	i-ti-ki-gta-hade	ag-gi-gta-hade
a-tin (to have)	—	—	—	—	a-ki-gta-tin	i-ti-ki-gta-tin	ag-gi-gta-tin
—	—	—	—	—	—	—	—
ka-ti-še (to love)	—	ka-ta-ta-še	—	—	ka-ti-še	—	ka-ti-še

Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post



## File Management and Organizing your Workflow

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Asset Name: mnh_8304.tif	Asset Name: mnh_8304.tif	Asset Name: mnh_8304.tif
Unique Asset Name: NMNH-mnh_8304	Unique Asset Name: NMNH-mnh_8304-000002	Unique Asset Name: NMNH-mnh_8304-00000
Asset Creator: National Anthropological Archives	Asset Creator: National Anthropological Archives	Asset Creator: National Anthropological Archives

*Above image: MNH 8304, Franz Boas posing for figure in USNM exhibit entitled "Hamats'a coming out of secret room" 1895 or before, National Anthropological Archives, Smithsonian Institution*



## File Management and Organizing your Workflow

75-7\_br02.jpg  
75-7\_br03.jpg  
75-7\_br04.jpg  
75-7\_br05.jpg  
75-7\_br06.jpg  
75-7\_br09.jpg  
75-7\_br15.jpg  
75-7\_br22.jpg  
75-7\_br26.jpg  
75-7\_br27.jpg  
75-7\_br29.jpg  
75-7\_br30.jpg  
75-7\_br33.jpg  
75-7\_br34.jpg  
75-7\_br42.jpg  
75-7\_br43.jpg  
75-7\_br47.jpg  
75-7\_br48.jpg  
75-7\_br49.jpg  
75-7\_br55.jpg  
75-7\_br60.jpg  
75-7\_br62.jpg  
75-7\_br64.jpg  
75-7\_br67.jpg  
75-7\_br68.jpg  
75-7\_x01.jpg  
75-7\_x02.jpg  
75-7\_x03.jpg  
75-7\_x04.jpg  
75-7\_x05.jpg  
75-7\_x06.jpg  
75-7\_x07.jpg  
75-7\_x08.jpg  
75-7\_x09.jpg  
75-7\_x10.jpg

### What are the Issues?

- filenaming
- formats/compatibility
- metadata
- Storage/use

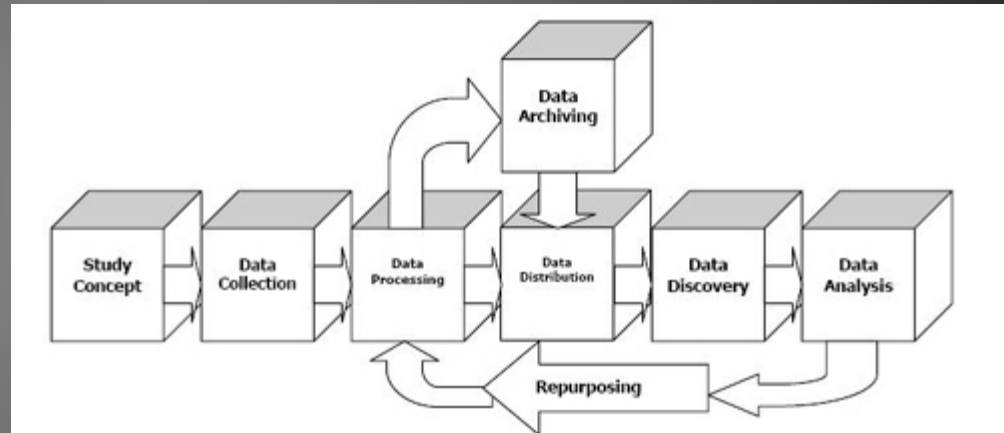
1882b\_1.jpg  
1882b\_2.jpg  
1882b\_3.jpg  
1882b\_4.jpg  
1882b\_5.jpg  
1882b\_6.jpg

1616\_075.jpg  
1616\_076.jpg  
1616\_077.jpg  
1616\_078.jpg  
1616\_079.jpg  
1616\_080.jpg  
1616\_081.jpg  
1616\_082.jpg  
1616\_083.jpg  
1616\_084.jpg  
1616\_085.jpg  
1616\_086.jpg  
1616\_087.jpg  
1616\_088.jpg  
1616\_089.jpg  
1616\_090.jpg  
1616\_091.jpg  
1616\_092.jpg  
1616\_093.jpg  
1616\_094.jpg  
1616\_095.jpg  
1616\_096.jpg  
1616\_097.jpg  
1616\_098.jpg  
1616\_099.jpg  
1616\_100.jpg  
1616\_101.jpg  
1616\_102.jpg  
1616\_103.jpg  
1616\_104.jpg  
1616\_105.jpg  
1616\_106.jpg  
1616\_107.jpg  
1616\_108.jpg  
1616\_109.jpg  
1616\_110.jpg  
1616\_111.jpg  
1616\_112.jpg  
1616\_113.jpg  
1616\_114.jpg  
1616\_115.jpg  
1616\_116.jpg  
1616\_117.jpg  
1616\_118.jpg  
1616\_119.jpg  
1616\_120.jpg  
1616\_121.jpg  
1616\_122.jpg  
1616\_123.jpg  
1616\_124.jpg



## File Management and Organizing your Workflow

- Project planning
- Project tracking
- Selection and Prioritization
- Image Capture
- Post-production and Quality Control
- Secure storage and migration

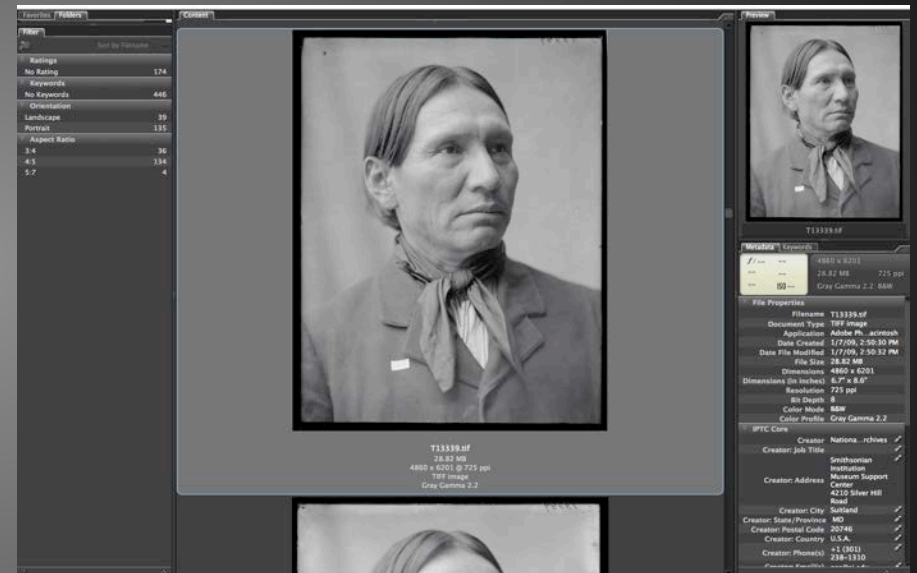


Source: DDI Structural Reform Group. "DDI Version 3.0 Conceptual Model." *DDI Alliance*. 2004. <<http://www.icpsr.umich.edu/DDI/committee-info/Concept-Model-WD.pdf>>.



## Project Planning: File Management and Organizing your Workflow

- What are the goals of the project?
- What does a basic workflow look like?





## Project Planning: File Management and Organizing your Workflow

### Sketch out a Workflow Diagram

- What are the specific materials you are going to digitize?
- Do you have enough storage space for the files?
- Begin to develop a filenaming schema – relate to collections naming as much as possible
- What metadata will you include?
- Lifecycle of the image
- Storage and migration





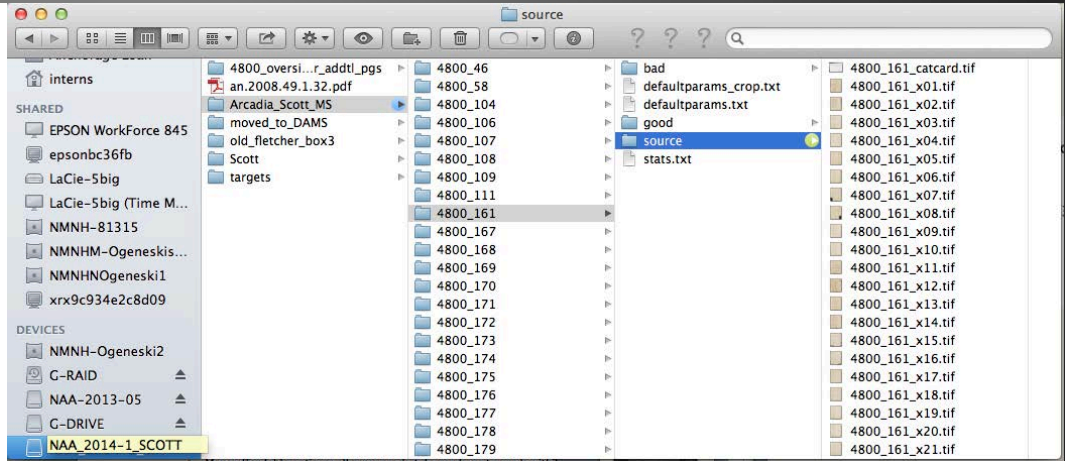


# Project Tracking: File Management and Organizing your Workflow

Folders, Folders, Folders: How you name them matters!

- BAE\_GN\_00001a-00599
- BAE\_GN\_00600a2-01199c
- BAE\_GN\_01200-01899
- BAE\_GN\_01900-02499
- BAE\_GN\_02500a-02899b13
- BAE\_GN\_02900a-03399b
- BAE\_GN\_03400a-03999
- BAE\_GN\_04000a1-04917
- BAE\_GN\_05766-t26537

- al\_fortpayne
- al\_stevenson
- al-ga\_rome
- ga\_dahlongena
- ga-sc\_walhalla
- ky-va-tn\_estillville
- legend
- nc\_cowee
- nc\_nantahalalah
- nc-sc\_pisgah
- nc-tn\_asheville
- sc\_pickens
- tn-chattanooga
- tn-loudon
- tn-nc\_knoxville
- tn-nc\_mtguyot
- tn-nc\_murphy
- tn-nc\_roan\_mountain
- va-tn\_bristol
- x\_ga-x\_walhalla
- x\_tn\_cleveland





## Pre-Digitization Work: File Management and Organizing your Workflow

Filenaming, Filenaming, Filenaming, Filenaming, Filenaming!

- Unique identifier; well defined
- Persistent and consistent
- Versioning

NAA Digital Filename Guidelines December  
2011

### Digital Filename Guidelines

National Anthropological Archives Smithsonian Institution

The National Anthropological Archives digital surrogates are named with unique persistent identifiers to help locate the surrogates throughout their lifecycle. The filenames are directly associated with the analog collection naming conventions in most cases, as to keep consistency between the analog item and its digital counterpart.

This document is divided into the following sections:

- A. General Filenaming Guidelines- FADGI
- B. NAA Specific Filenaming Guidelines
- C. Examples and Explanation of NAA Digital Filenaming Conventions (current and legacy practice)
- D. Directory and Disk Names (current and legacy practice)
- E. Batch Renaming (when necessary)
- F. Appendix: Legacy Required Neg. Number Prefixes

#### A. General Filenaming Guidelines

The National Anthropological Archives is participating in the [Federal Agencies Digitization Guidelines Initiative](#) (FADGI), a collaborative effort by federal agencies formed in 2007 to define common guidelines, methods, and practices to digitize historical content in a sustainable manner.

The [Technical Guidelines for Digitizing Cultural Heritage Materials](#), released by FADGI in 2010, is the master document that defines a set of guidelines for still images as recommended by the group.

The following section: VI. Metadata – Filenaming through Naming Derivative Files is taken from the August 2010 version of the above-mentioned guidelines. The portion on quality of metadata has not been included. Please refer to the document at the link above for further information. Verifying metadata routines are laid out in the NAA doc: Embedding Metadata and JPG Derivative Guide.

#### File Naming

A file-naming scheme should be established prior to capture. The development of a file naming system should take into account whether the identifier requires machine- or human-indexing (or both – in which case, the image may have multiple identifiers). File names can either be meaningful (such as the adoption of an existing identification scheme which correlates the digital file with the source material), or non-

Last updated: December 2011 Comments: [Stephanie Christensen](#)



## Pre-Digitization Work: File Management and Organizing your Workflow

Analog Collection Name	Digital Surrogate <u>File naming</u> Convention	Example
NAA MS 385	<u>Msnumber_number.tif</u> <sup>1</sup>	385_002.tif
Notecard within NAA MS 385	<u>Mnnumber_notecard.tif</u>	385_notecard_1.tif (use number if sequence)
Notebook within NAA MS 385	<u>Msnumber_ntbk_number.tif</u>	385_ntbk_001.tif (multiple notebooks: 385_ntbk_1_001.tif, 385_ntbk_2_001.tif, etc)
Multiple folders of NAA MS 4117-a	<u>MSnumber_foldernumber_number.tif</u>	4117-a_f1_001.tif, 4117-a_f2_001.tif, 4117-a_f3_001.tif, 4117-a_f4_001.tif
Multiple volumes of NAA MS 1795a	<u>Msnumber_volnumber_number.tif</u>	1795a_vol1_001.tif, 1795a_vol2_001.tif, 1795a_vol3_001.tif, 1795a_vol4_001.tif, 1795a_vol4_drawings_001.tif
Kinship chart or other identifiable materials within NAA MS 385	<u>Msnumber_kinship_chart.tif</u>	351_kinshipchart_no_1_ver_2.tif <sup>2</sup>



## Pre-Digitization Work: File Management and Organizing your Workflow

Note within NAA MS 385	<u>Msnumber_note.tif</u>	385_note.tif; 385_note_1.tif; 385_note_2.tif
Drawing within NAA MS 385	<u>Msnumber_drawing.tif</u>	385_drawing.tif*
Map within NAA MS 385	<u>Msnumber_map.tif</u> or <u>msnumber_map_identifiableinfo.tif</u>	385_map_001.tif; 385_map_east_facade.tif
Catalog cards	<u>Msnumber_catcard.tif</u>	385_catcard.tif; 385_catcard_v.tif <sup>3</sup>
MS 176,622	These types of numbering are also used with manuscripts and items that may contain INV numbers	0851501.tif
MS 166,032	Directory folder is 166032	Folder uses INV numbers 08704000
Ledger art	NMNHnaa_la_Annon_Plains_08531400_post.tif	File given out to public—drop prefix and suffix—distribute: 08513400.tif
Ledger art	May use INV numbers	<b>11005600.tif</b> <b>08624400.tif</b>
Ledger art	NMNHnaa_la_ms4653_08691300_post	Contained within MS 4653—drop prefix and suffix; distribute: 08691300.tif



## Pre-Digitization Work: File Management and Organizing your Workflow

Filenaming, Filenaming, Filenaming, Filenaming, Filenaming,  
Filenaming,...

- Prefix and suffix additions
- Examples of filename variations:
  - “v” = verso, the backside of a page
  - “x01” = numbering assigned during digitization
- “ntbk” – manuscript is a bound notebook
  - “front\_cover,” “back\_cover”
  - “title\_page”

### **MS 4490**

Main file to distribute/ combines all other files 4490.tif

These files illustrate Left, right, mid section and additional “detail views” that may have been shot

4490\_1\_detail.tif  
4490\_2\_detail.tif  
4490\_3\_detail.tif  
4490\_4\_detail.tif  
4490\_5\_detail.tif  
4490\_6\_detail.tif  
msnumber\_left\_1.tif; msnumber\_right\_1.tif



## Actual Digitization: File Management and Organizing your Workflow

### File Format

.tiff .jpg .mov .wav .mp3 .pdf

Are you saving in a file format that is compatible in the long term?

Hi- bit or low bit depth; Color space





## Actual Digitization: File Management and Organizing your Workflow

### Secure Storage and Migration

- Can you access your files easily?
- Do you understand your filenames months later after you've been away from the project?
- Do you have multiple copies in separate locations?
- What is your storage media?
- Do you have a long-term plan for the security of your data?
- If so you are on your way!





## File Management and Organizing your Workflow

### Additional Resources

- [Federal Agencies Digitization Guidelines Initiative](#)
- [University of Cambridge Library – Data Management, Choosing File Formats](#)
- [MIT Libraries: Data Management and Publishing](#)







Smithsonian  
*National Museum of Natural History*

# Digitization Project Management Selection and Prioritization, A Case Study: Breath of Life Archival Institute: 2013, 2015

Jeanine Nault  
Digital Imaging Specialist  
National Anthropological Archives  
National Museum of Natural History  
Smithsonian Institution  
naultj@si.edu

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



Smithsonian  
National Museum of Natural History

Breath of Life Archival Institute for Indigenous Languages | Facebook

www.facebook.com/breathoflifedc

Login Microsoft E... Web Access SIRIS - Smit...tion System Collections S... Institution Apple (277) Amazon News (3,467) PRISM: Smith...lan Intranet

facebook

Email or Phone  Password  Log In

Keep me logged in [Forgot your password?](#)

**Breath of Life Archival Institute for Indigenous Languages is on Facebook.**

To connect with Breath of Life Archival Institute for Indigenous Languages, sign up for Facebook today.

[Sign Up](#) [Log In](#)



**Breath of Life Archival Institute for Indigenous Languages** Like

204 likes

Education  
Breath of Life Institute connects Native language activists with archival and museum materials for use in language learning, teaching, and revitalization.

About Photos Likes

Highlights

**Breath of Life Archival Institute for Indigenous Languages** shared Recovering Voices's status. July 22

Grants available for community-based research at the Smithsonian!

Recent Posts by Others See All

**Jan Danek**  
What a wonderful two weeks!! Thanks to all the organ...  
June 21 at 3:36pm

**Wenona Gardner**  
Mohican-8 An intertribal circle for those interested ...  
February 25 at 5:10am

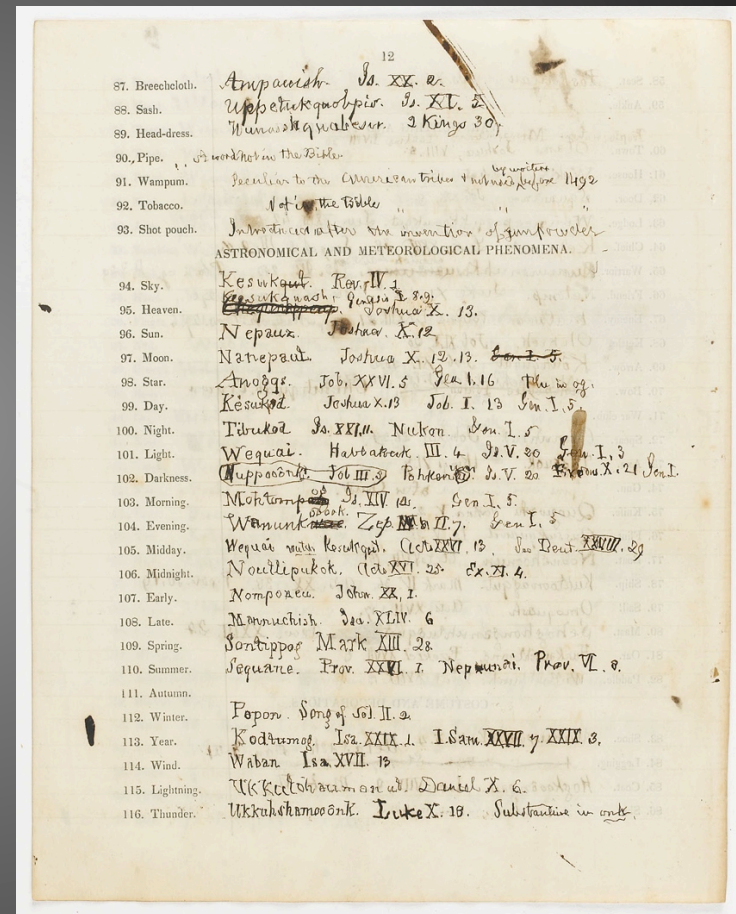
## What is Breath of Life

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# Goals of the Project

- Provide access to digitized materials for participants for their language revitalization efforts
- Create high-resolution digital surrogates for preservation of our collection



NAA MS 1827, "Massachusetts or Natick vocabulary"



# Collaborators

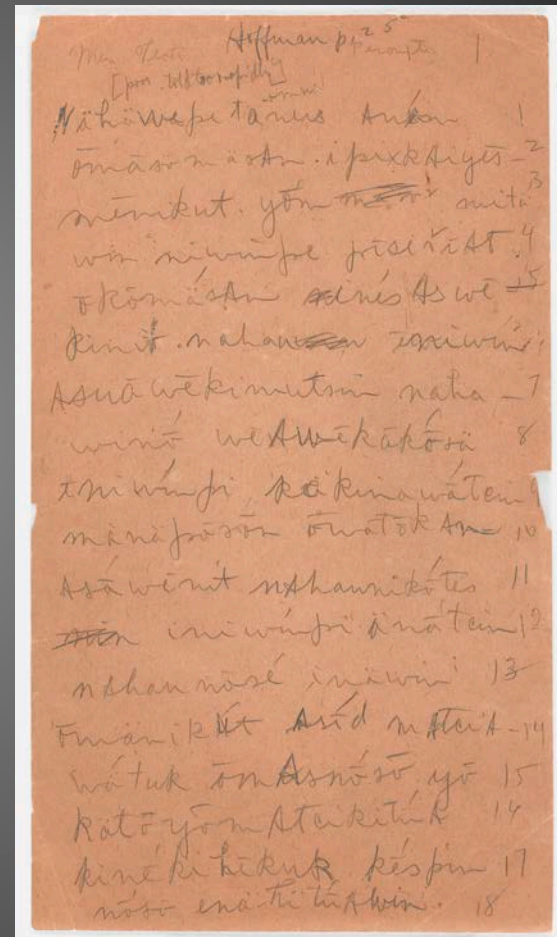
- Digital Lab Staff
  - Digital Program Manager
  - Digital Imaging Technician
  - Digital Imaging Intern
- NAA Staff
  - Reference Staff and Interns
  - Conservator
- Curator
- Program Staff
- Participants

The image shows two screenshots of Smithsonian websites. The top screenshot is the 'Recovering Voices' page, which features a navigation bar with links like Home, About Us, Research, Highlights, Events, Resources & Grants, and Connect & Support. It includes a photo of two people looking at a map, a tweet from @RecoverVoices, and text describing the program's mission to document and revitalize endangered languages. The bottom screenshot is the 'National Anthropological Archives (NAA) Human Studies Film Archives (HSFA)' page, which features a photo of an archive hallway and text explaining the collections and providing information on how to request an appointment to visit the archives.



# Breath of Life 2013 – By the Numbers

- 16 language teams
- 17 manuscripts, 20 text selections from the John P. Harrington collection, and 7 previously digitized manuscripts
- 1000 pages of archival manuscript material requested; over 3000 digitized in total
- Material ranges from manuscript pages, bound notebooks of various sizes and binding styles, to notecards and slipfiles



NAA MS 2797, "Menominee linguistic notes and texts collected by Truman Michelson, 1910 "



# Selection Process

BOL\_2013\_MS\_Digi\_Requests.xlsx

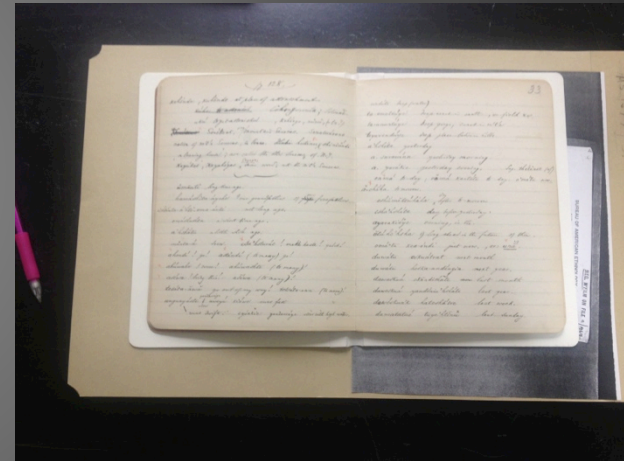
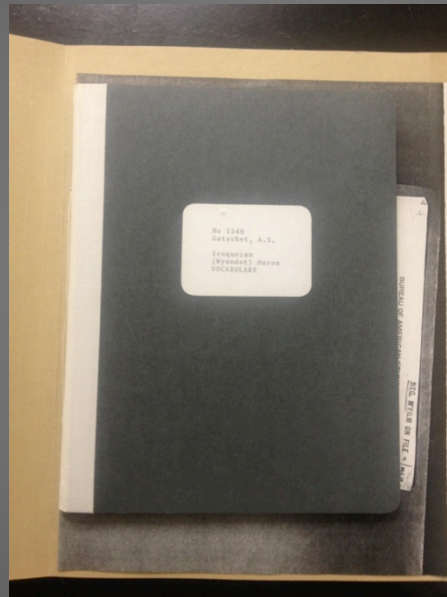
Language	NAA Local Control Number	Title (Siris Record)	Creator	Physical Description (size etc.)	Scanning Instructions	Observations made during preliminary review (box #)	Conservation Y/N	Digitization Request Date	# of pages for digitization	Digitization status	PDF complete	Notes
Menominee	NAA MS 2797	Texts	Michelson, Truman	13 pages	SELECTED FOR DIGITIZATION NOTE: the paper is very brittle, so for the purposes of conservation it would be ideal to scan all pages. If this is not possible, then follow these instructions: Please scan pp. 1-28. Also scan pp. 1-6 and 7-12.	Handwritten notes that appear to be of texts. Several dozen pages are in Menominee. Some pages are in English but it is not clear whether these are translations of the Menominee texts. The Menominee texts seems to be paginated in an odd way (or are out of order). The Menominee pages include some grammatical and lexical data as well. Perhaps we could digitize some texts. Conservation: paper is very brittle and breaks to the touch. Recommendation to BoL participants: Please review the digital surrogates against the physical manuscript and against MS 2824 to determine whether the English texts in these records are the translations of the Menominee texts.	Y	20-Mar	41	complete	x	
Maidu	NAA MS 646-b	Linguistic material 1877	Gatschet, Albert S. (Albert Samuel), 1832-1907 collector	84 pages	Please scan until page 1-33, 91-102, 129-131 for BoL. However, for conservation purposes, I recommend the entire manuscript be scanned.	Substantial sentential data, some vocabulary. I have selected this manuscript because our holdings of Konkow are modest, yet Chico Maidu is often considered to be the same language as Konkow (Golla 2011:138). Conservation: the ink seems to be fading or at least it is very light. Scanning may be recommended for conservation purposes.		26-Mar	45	complete	x	
Maidu	NAA MS 646-c	Supplementary Maidu vocabulary and sentences	ANONYMOUS	3 pages	Please scan all 3 pages	Konkan dialect. Limited data.	N	26-Mar	5	complete	x	
Hupa	NAA MS 1442	Hupa vocabulary December 1888-January 1889	Curtin, Jeremiah	107 pages	Please scan: p 82-83; 88-96; 109-125; 132-133.		N	26-Mar	28	complete	x	
Hupa	NAA MS 954	Weitspek (Pohlik Klamath) and Hopah dictionaries and ethnographic notes 1852	Gibbs, George	41 pages autograph document signed	Please scan entire notebook marked Shelf 71			2-Apr	29	complete	x	
Karok	123	Karok Medicine Formulas	Harrington, John Peabody 1884-1961		please scan first 11 page (marked) of 123-1		Y	2-Apr	11	complete	x	
Karok	130	Typescript of MS "Ethnobotany and Ethnozoology of the Karuk Indians of California" (former BAE MS#4553)	Harrington, John Peabody 1884-1961		please scan marked pages (10 pages) in 130-3 and marked pages (11 pages) in 130-2			2-Apr	21	complete	x	
Karok	122	Karok Stories and Myths	Harrington, John Peabody 1884-1961		Mockingbird and Robin story, Box 122, Folder 4, 11 pages separated by paper slip.		N	28-Mar	11	complete	x	
Fernandeño	563-1	Linguistic and Ethnographic Field Notes	Harrington, John		Please scan all 12 handwritten letter-size pages and the first 10 typewritten	The first slip on the data slip pack is somewhat damaged.	Y	28-Mar	17	complete	x	

Ready Sum = 92



Smithsonian  
National Museum of Natural History

# Breath of life 2013 – variations on a theme



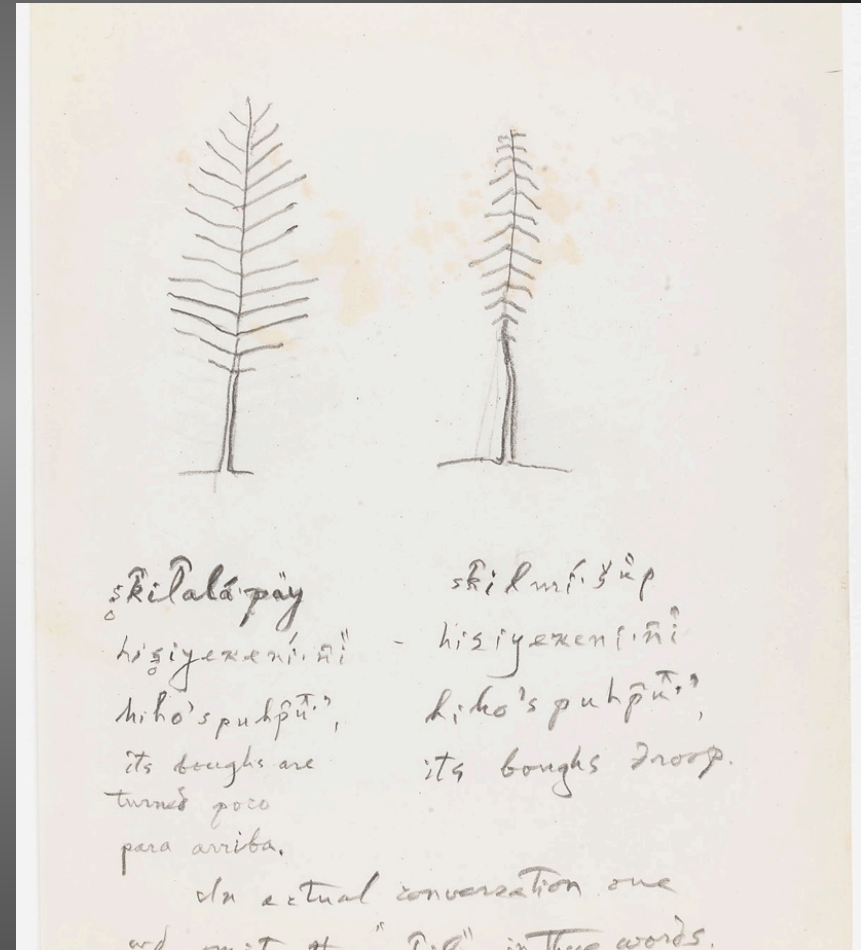
NAA MS 1549, "Wandot/Wyandot/ vocabulary January 8, February 7, March 15, 1881"

Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post



# The Harrington Issue

- Extremely large, not fully processed collection
- In the middle of another digitization project of the collection
- Approach differently than other materials
  - Not digitizing full folders or boxes for preservation purposes (access only)

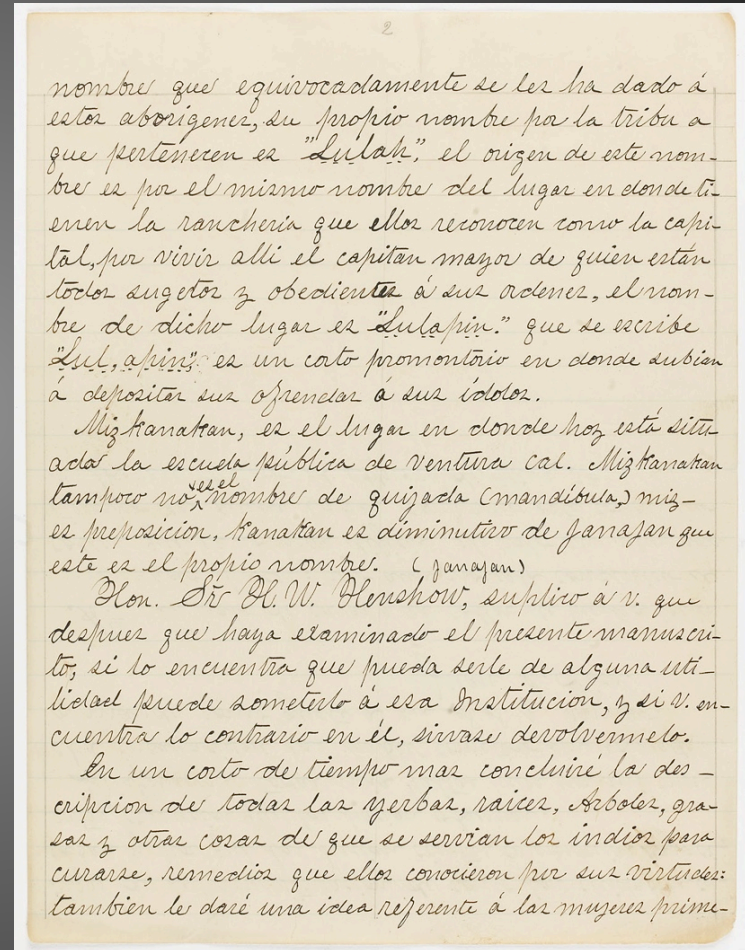
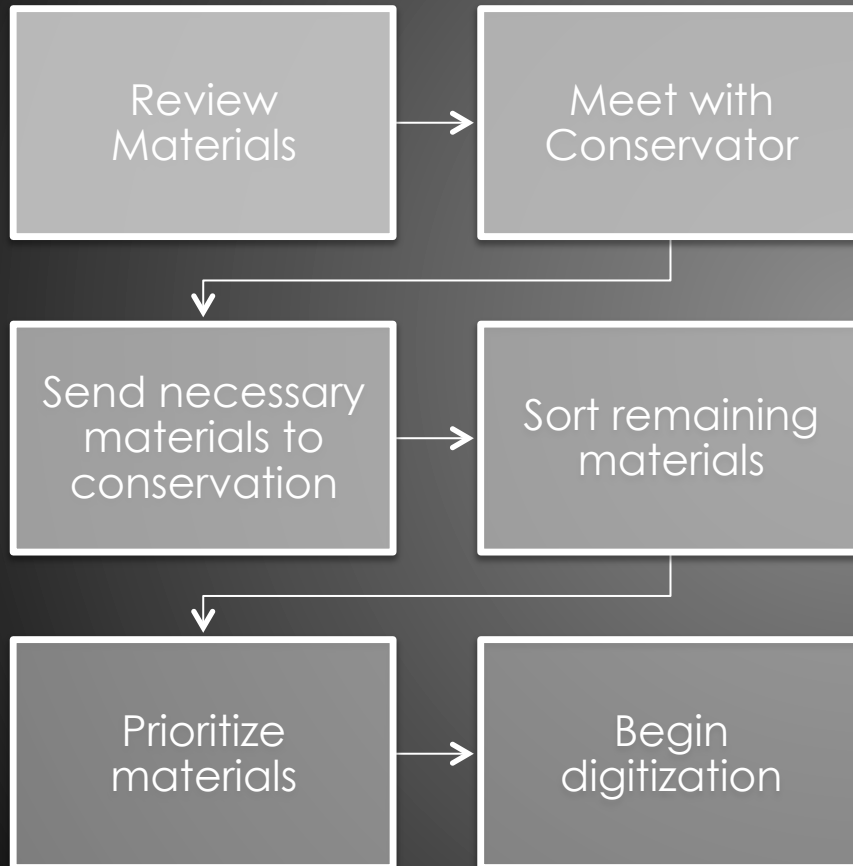


John P. Harrington collection Box 483, Folder 1





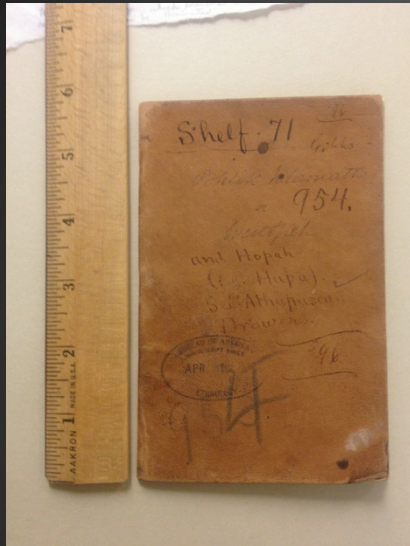
# Prioritization



NAA MS 3718, "Fourth of July, 1890 "



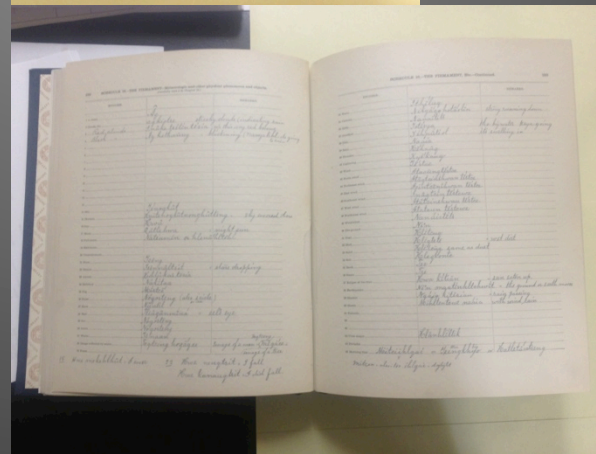
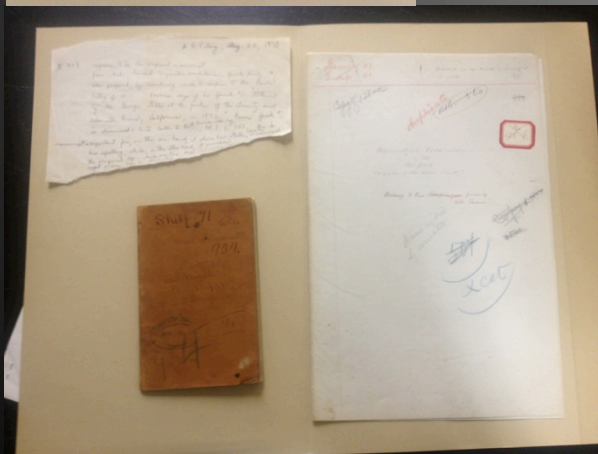
# Materials: What we're working with

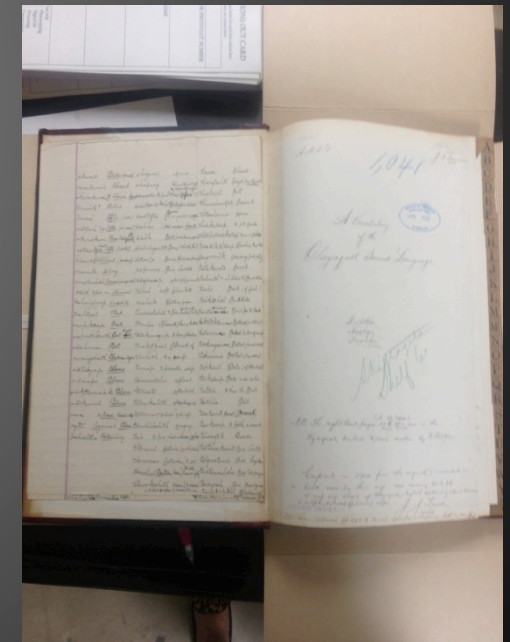
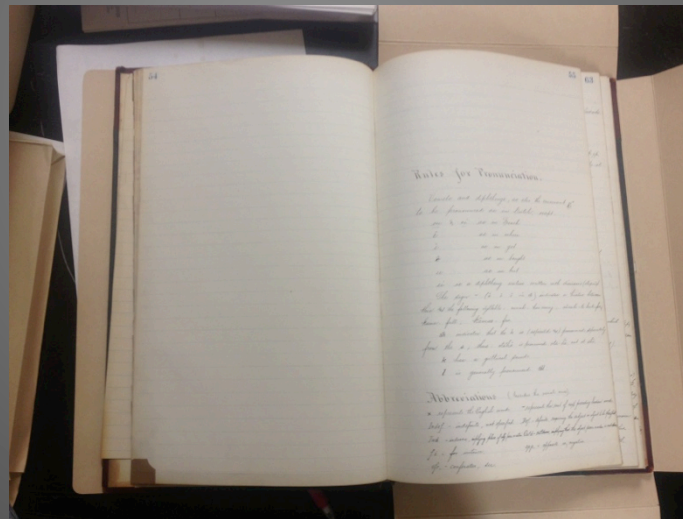
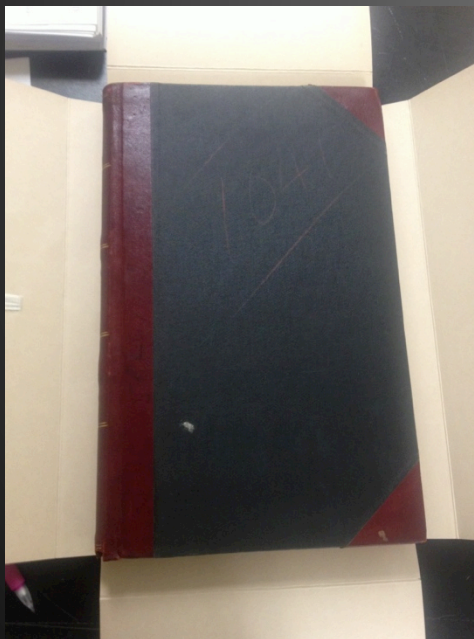
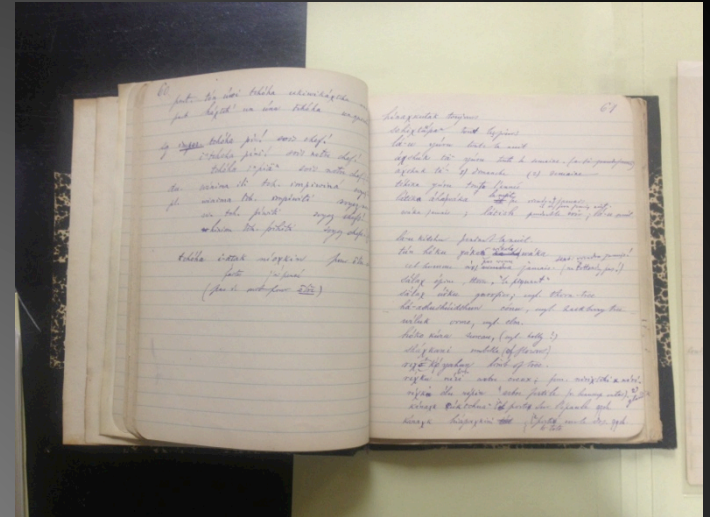
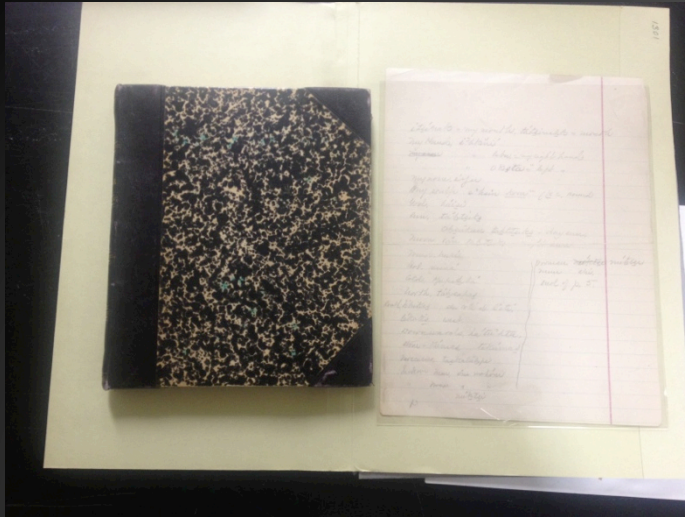


NAA MS 954



NAA MS 1442



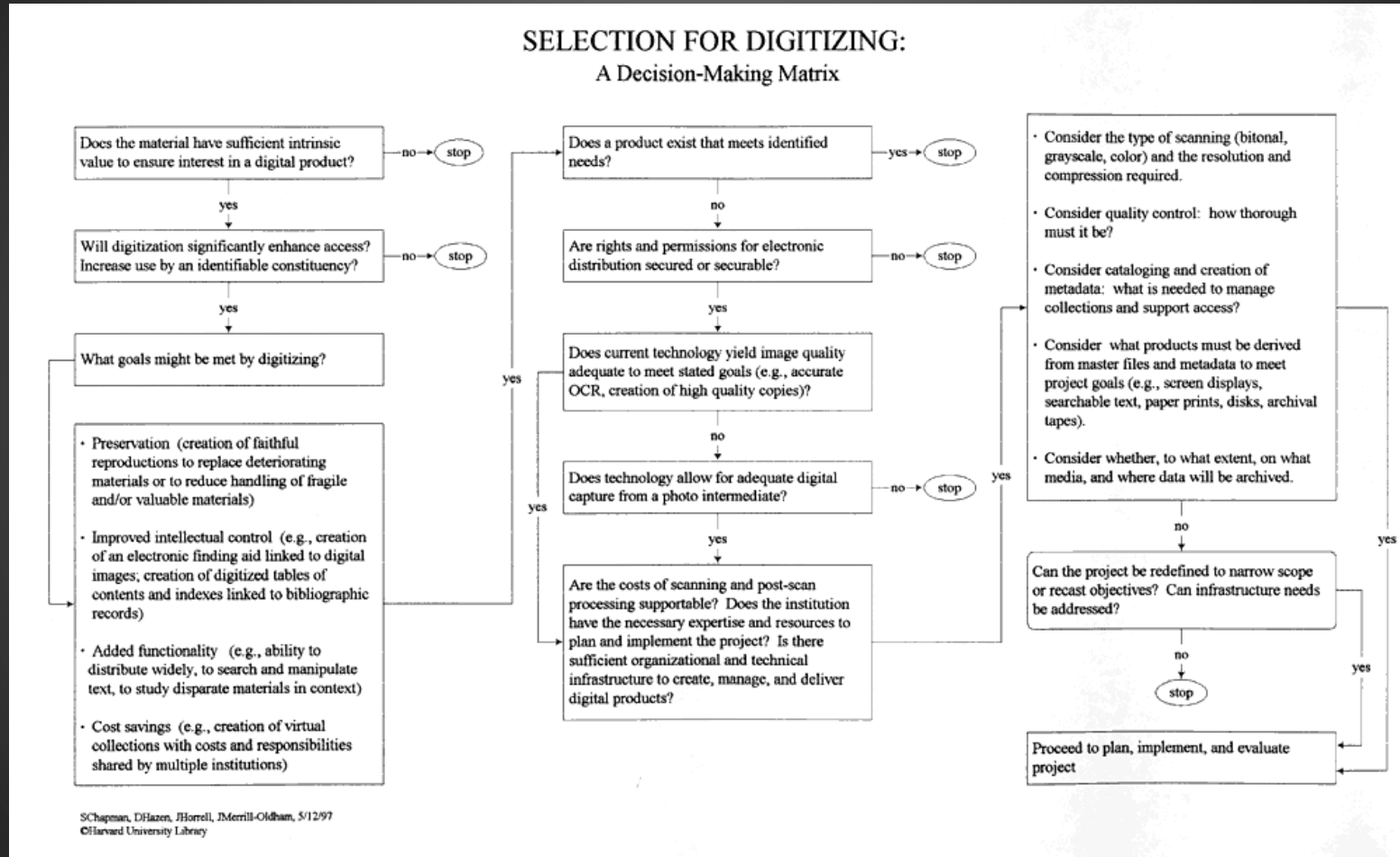


Above image from: NAA MS 1301 (T) and NAA MS 1041 (B)

Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post



# A textbook approach

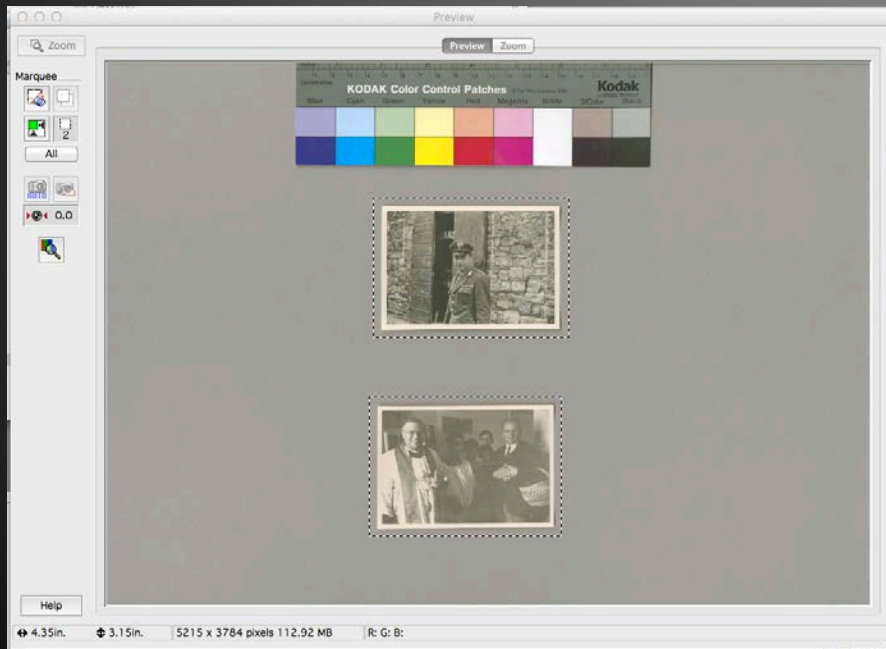


Hzen, Dan, Horrell, Jeffery, and Merrill-Oldham. (1998) "Selection for Digitizing: A Decision-Making Matrix," *Selecting Research Collections for Digitization – Full Report*.

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators , April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# Image Creation: Recommendations and Realities





# Questions to Consider

- What is being digitized?
- Where are these files going?
- Where will they be stored?
- Who will create them?
- What guidelines are being followed?  
Or created?
- What are the technical specifications?



Smithsonian  
*National Museum of Natural History*



## Federal Agencies Digitization Guidelines Initiative

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# (Even more!) things to consider

- Staff and Time
  - Expertise and Testing
- Imaging environment
  - Space
  - Lights
  - Monitor
- Hardware and Software
  - Scanner v. Camera
  - Platform (Mac v. PC)
  - Imaging Software
    - Proprietary with equipment
    - Adobe products
    - Image Science Associates

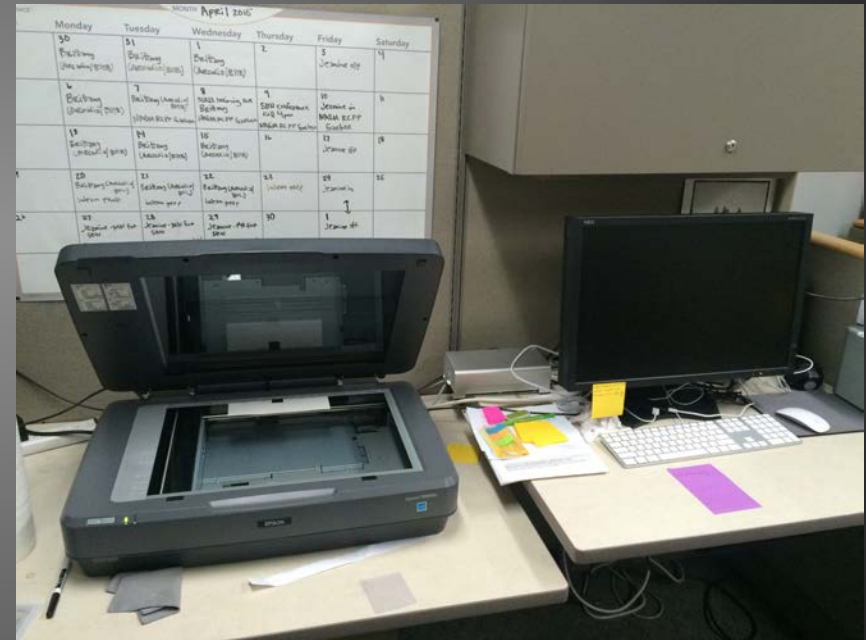
BOTTOM LINE: \$





Smithsonian  
National Museum of Natural History

# Scanner v. Camera



*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# Technical Specifications

- Bit depth
- Tone/Endpoints
- Color mode/profile
  - Resolution
  - File formats



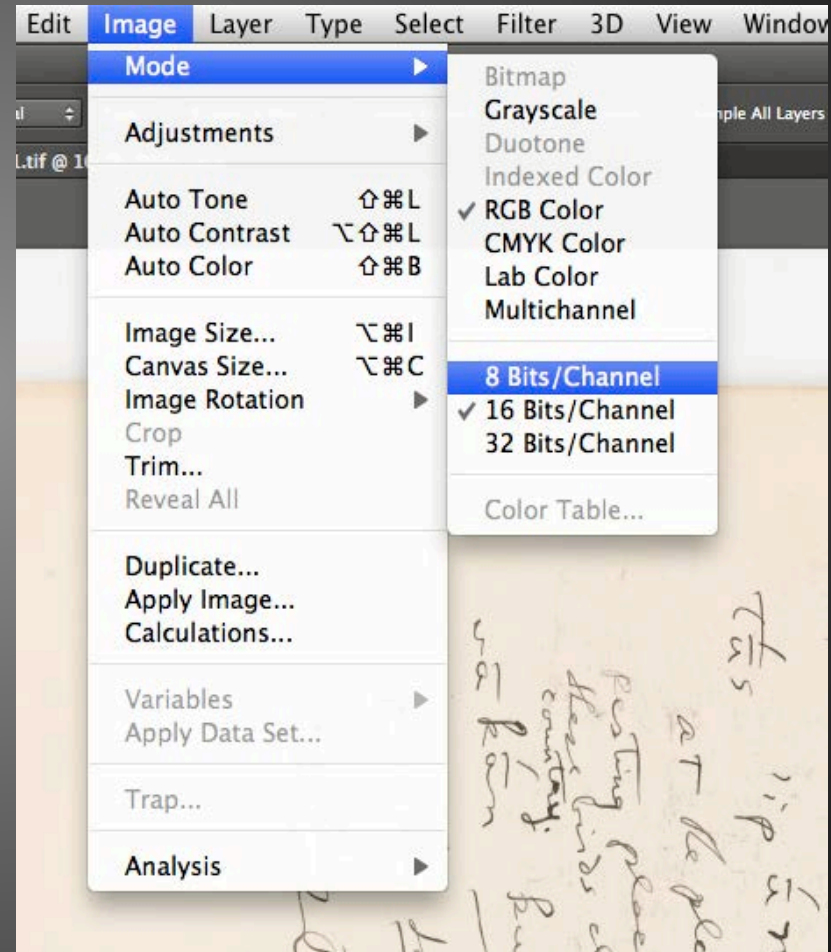
# Technical Specifications – Bit Depth

- Bit depth: the color information stored in an image; the higher the bit depth, the more colors an image can store. A 1 bit image only shows two colors – black and white
  - 8 bit:  $2^8 = 256$  colors
  - 16 bit:  $2^{16} = 65,536$  colors
  - 24 bit:  $2^{24} = 16,000,000+$  colors
  - 48 bit:  $2^{48} = 3,000,000,000$  colors



# Bit Depth – Recommendations

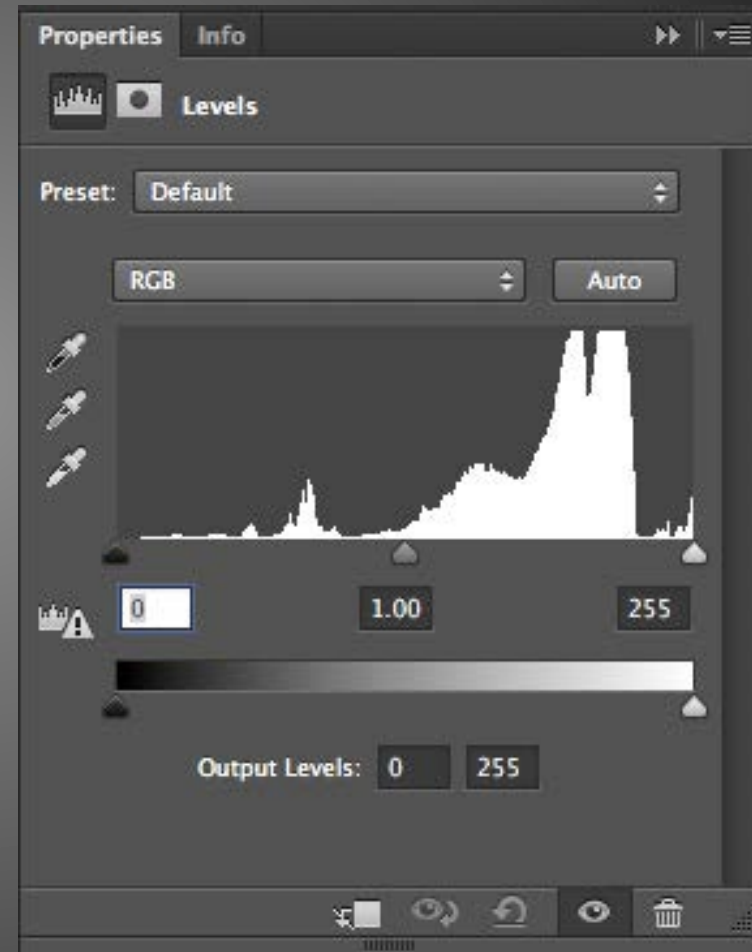
- Capture images in a high bit depth (16 bit or higher) then convert to a lower bit depth as needed
- Higher bit depth images will increase accuracy of representing the original BUT will create larger files





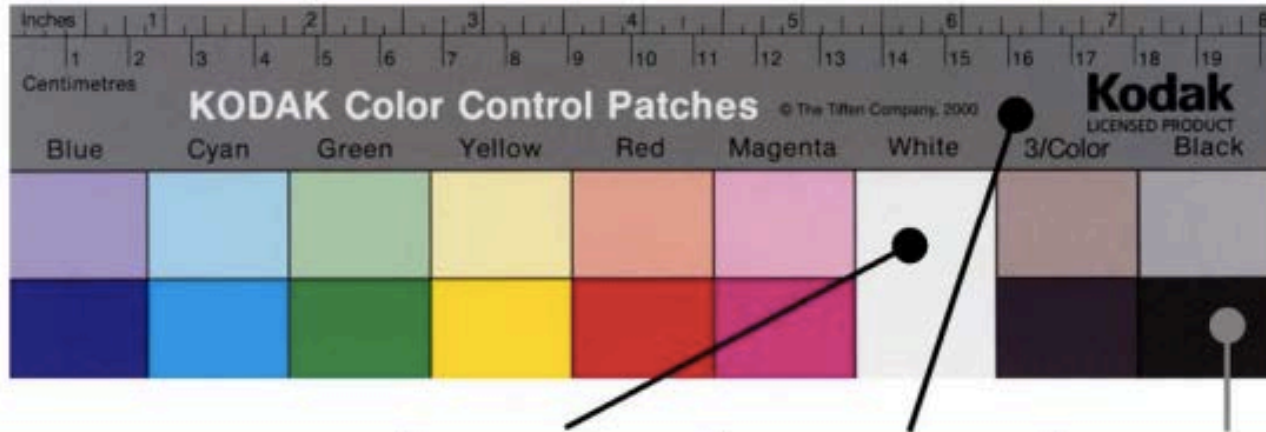
# Technical Specifications – Tone

- Each pixel in an image file is assigned a value between 0 and 255, with 0 being a “true” blackest black, and 255 being a “true” whitest white.
- Values below 9 (too black) or above 247 (too white) lose all detail





# Tone - Recommendations



		Neutralized White Point	Neutralized Mid Point*	Neutralized Black Point
Color Patch/Area		White	Gray Background	Single Color Black
Aimpoint	RGB Levels	237-237-237	102-102-102	23-23-23
	% Black	7%	60%	91%
Acceptable Range for Aimpoint	RGB Level	233 to 241	98 to 106	19 to 27
	% Black	5% to 9%	58% to 62%	89% to 93%

\*Aimpoint for mid point (MP) to be calculated from actual values for white point (WP) and black point (BP) using the following formula:  $MP = WP - 0.63(WP - BP)$



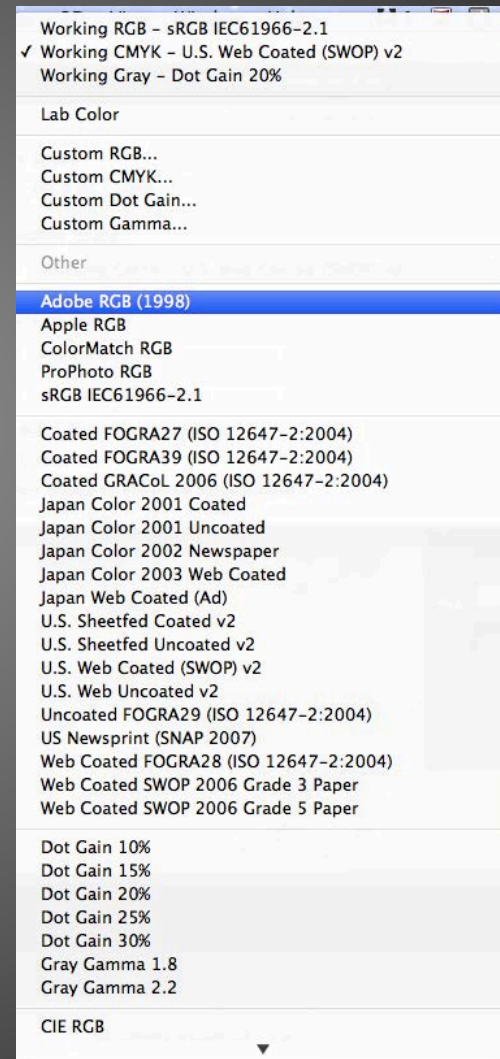
# Technical Specifications – Color Mode

- “Grayscale image files consist of a single channel, commonly either 8-bits (256 levels) or 16-bits (65,536 levels) per pixel with the tonal values ranging from black to white. Color images consist of three or more grayscale channels that represent color and brightness information. Common color modes include RGB (red, green, blue), CMYK (cyan, magenta, yellow, black), and LAB (lightness, red-green, blue-yellow). The channels in color files may be either 8- bits (256 levels) or 16-bits (65,536 levels). Display and output devices mathematically combine the numeric values from the multiple channels to form full color pixels, ranging from black to white and to full colors.”



# Color Mode – Recommendations

- Adobe RGB 1998 (or Adobe sRGB, alternatively) is the recommended color profile for master files







# Technical Specifications – Resolution

- Resolution: the number of pixels in each dimension that can be displayed
- DPI: dots per inch (more appropriate for printing)
- PPI: pixels per inch (more accurately describes images)
- “Resolution Threshold”: the point where no matter how many more pixels added per inch, no more information from original is gained



Smithsonian  
National Museum of Natural History

# Resolution – Recommendations

“All digital images created from reflective analog textual materials should have a minimum of 4,000 pixels along their longest dimension, and a minimum resolution of 400 dpi. Text on the reverse side of an object may be scanned at 200 dpi. All reflective photographic materials are at a minimum of 5,000 pixels on their longest dimension.

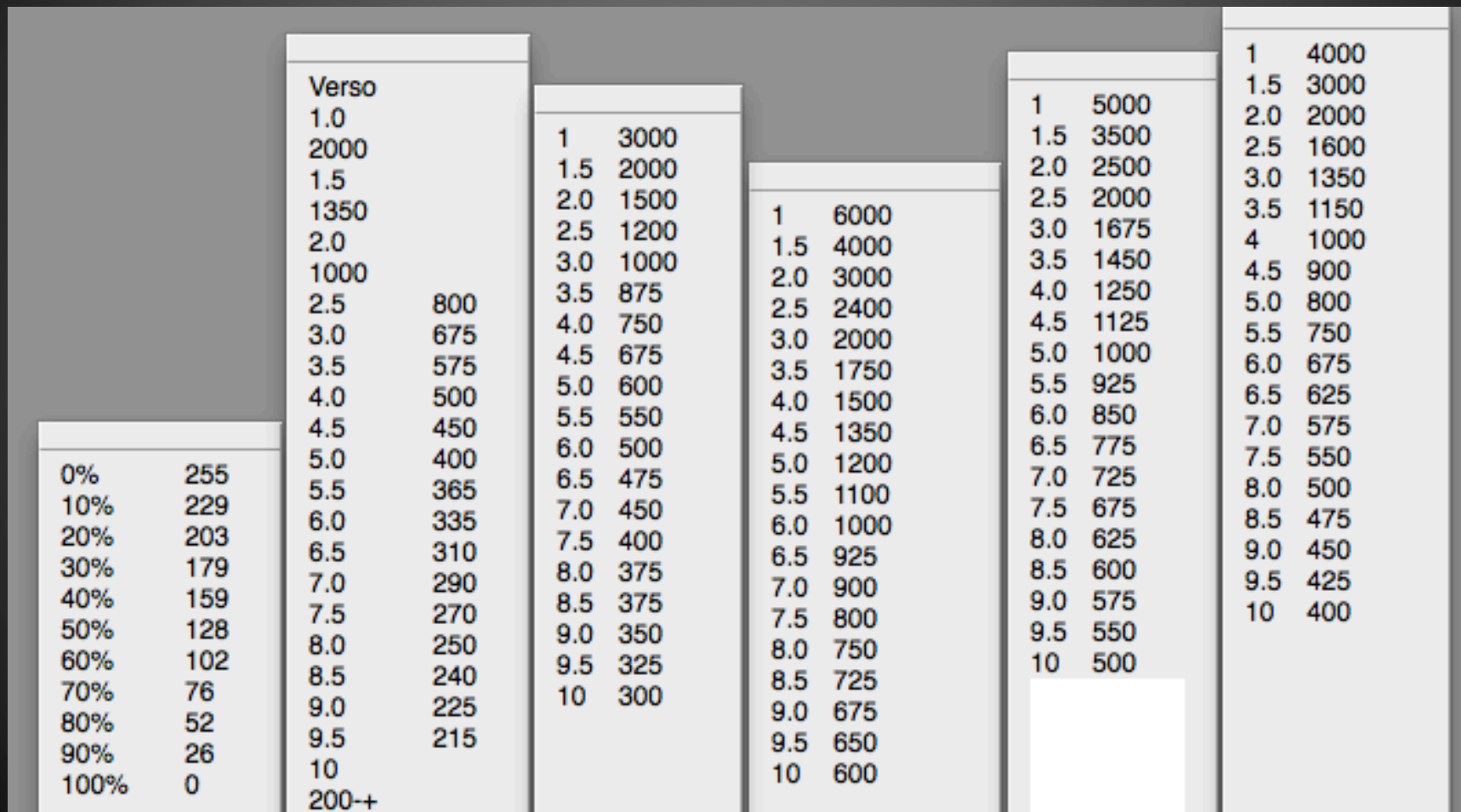
All digital images created from transmissive analog materials should have a minimum of 6,000 pixels along their longest dimension, and a minimum of 600 dpi.

For example, an 8 x 10-inch photograph should be scanned at 500 dpi; a 5-inch photograph at 1000 dpi, etc. However, photographic items should never be scanned at less than 500 dpi — even if the longest dimension of the object is greater than 10 inches. When measuring the longest dimension of an image, round down to the nearest half inch. When calculating resolution, round up to the closest multiple of 25. Calculations should reflect the dimensions of actual item to be scanned, not its matting or support.”

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# Resolution – Recommendations





# Technical Specifications – File Formats

File Format	Technical Considerations	Recommended Use
TIFF	<ul style="list-style-type: none"> <li>• “De facto” raster image format used for master files</li> <li>• Simply encoded raster-based format</li> <li>• Accommodates internal technical metadata in header/extensible and customizable header tags</li> <li>• Supports Adobe’s XMP (Extensible Metadata Platform)</li> <li>• Accommodates large number of color spaces and profiles</li> <li>• Supports device independent color space (CIE L*a*b)</li> <li>• Uncompressed; lossless compression (Supports multiple compression types for 1-bit files). JPEG compression not recommended in TIFF file</li> <li>• High-bit compatible</li> <li>• Can support layers, alpha channels</li> <li>• Accommodates large file sizes</li> <li>• Anticipate greater preservation support in repository settings; preferred raster image format for preservation</li> <li>• Widely supported and used</li> <li>• Long track record (format is over 10 years old)</li> <li>• Potential loss of Adobe support of TIFF in favor of PDF?</li> <li>• Not suitable as access file—no native support in current web browsers</li> </ul>	Preferred format for production master file
JPEG 2000	<ul style="list-style-type: none"> <li>• Increasingly considered as a viable format for master image files, but not yet widely adopted</li> <li>• More complex model for encoding data (content is not saved as raster data)</li> <li>• Supports multiple resolutions</li> <li>• Extended version supports color profiles</li> <li>• Extended version supports layers</li> <li>• Includes additional compression algorithms to JPEG (wavelet, lossless)</li> </ul>	Rapidly gaining acceptance as a format for production master file—however, not currently widely implemented (or TIFF files are also being saved alongside JPEG 2000 files as master formats (?))

JFIF/JPEG	<ul style="list-style-type: none"> <li>• Lossy compression, but most software allows for adjustable level of compression</li> <li>• Presence of compression artifacts</li> <li>• Smaller files</li> <li>• High-bit compatible</li> <li>• Longer decompression time</li> <li>• Supports only a limited set of internal technical metadata</li> <li>• Supports a limited number of color spaces</li> <li>• Not suitable format for editing image files—saving, processing, and resaving results in degradation of image quality after about 3 saves</li> </ul>	Access derivative file use only— not recommended for text or line drawings
PDF	<ul style="list-style-type: none"> <li>• Intended to be a highly structured page description language that can contain embedded objects, such as raster images, in their respective formats.</li> <li>• Works better as a container for multiple logical objects that make up a coherent whole or composite document</li> <li>• More complex format due to embedded/externally linked objects</li> <li>• Implements Adobe’s XMP specification for embedding metadata in XML</li> <li>• Can use different compression on different parts of the file; supports multiple compression schemes</li> <li>• Supports a limited number of color spaces</li> </ul>	Not recommended for production master files
PDF/A	<ul style="list-style-type: none"> <li>•</li> </ul>	
PNG	<ul style="list-style-type: none"> <li>• Simple raster format</li> <li>• High-bit compatible</li> <li>• Lossless compression</li> <li>• Supports alpha channels</li> <li>• Not widely adopted by imaging community</li> <li>• Native support available in later web browsers as access file</li> </ul>	Possible format for production master file—not currently widely implemented
GIF	<ul style="list-style-type: none"> <li>• Lossy (high color) and lossless compression</li> <li>• Limited color palette</li> <li>• 8-bit maximum, color images are dithered</li> <li>• Short decompression time</li> </ul>	Access derivative file use only— recommend for text records
[ASCII]	<ul style="list-style-type: none"> <li>• For image files converted to text</li> <li>• Potential loss to look and feel of document/formatting</li> </ul>	N/A
[XML]	<ul style="list-style-type: none"> <li>• For image files converted to text</li> <li>• Hierarchical structure</li> <li>• Good for encoding digital library-like objects or records</li> <li>• Allows for fast and efficient end-user searching for text retrieval</li> <li>• Easily exchanged across platforms/systems</li> </ul>	N/A



# File Format - Recommendations

File Format	Technical Considerations	Recommended Use
TIFF	<ul style="list-style-type: none"><li>▪ “De facto” raster image format used for master files</li><li>▪ Simply encoded raster-based format</li><li>▪ Accommodates internal technical metadata in header/extensible and customizable header tags</li><li>▪ Supports Adobe’s XMP (Extensible Metadata Platform)</li><li>▪ Accommodates large number of color spaces and profiles</li><li>▪ Supports device independent color space (CIE L*a*b)</li><li>▪ Uncompressed; lossless compression (Supports multiple compression types for 1-bit files). JPEG compression not recommended in TIFF file</li><li>▪ High-bit compatible</li><li>▪ Can support layers, alpha channels</li><li>▪ Accommodates large file sizes</li><li>▪ Anticipate greater preservation support in repository settings; preferred raster image format for preservation</li><li>▪ Widely supported and used</li><li>▪ Long track record (format is over 10 years old)</li><li>▪ Potential loss of Adobe support of TIFF in favor of PDF?</li><li>▪ Not suitable as access file—no native support in current web browsers</li></ul>	Preferred format for production master file



# Basic Scanning Workflow

1. Setup Scanning Workstation
2. Select Scanner Settings
3. Setup Scanner for Use
4. Scan
5. Save

Extra: Batch Scanning

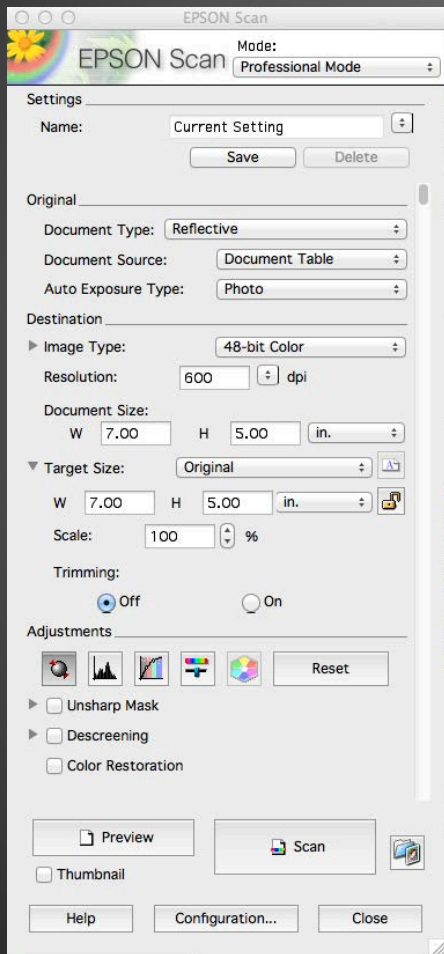


# 1. Setup Scanning Workstation

- Turn on computer and scanner
- Remove any loose articles of clothing or jewelry
- Wash hands or wear gloves (no lotions, no hand sanitizers)
- You will need:
  - A ruler
  - Staging area for original objects
  - Gloves and/or clean hands
  - Neutral background board



## 2. Select Scanner Settings



- Transmissive (negatives) v. Reflective (photographs, text)
- Resolution
- Color Profile
- Tonal Range
- Turn off adjustments!





## 3. Setting up Scanner for Use

- Wipe down scanner glass with lint free cloth
- Place color bar target in appropriate place (center top, for instance)
  - Utilizing a target is crucial for maintaining image quality and to serve as reference point if the integrity of images is every questioned





# 3. Setting up Scanner for Use

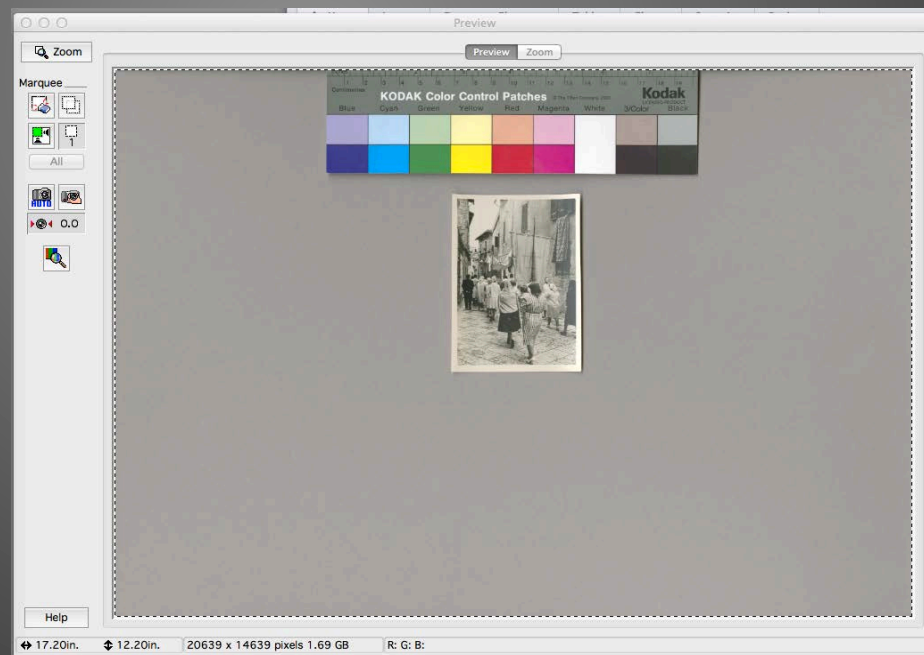
Measure the object you wish to scan. All digital images created from reflective analog textual materials should have a minimum of 4,000 pixels along their longest dimension. When measuring the longest dimension of an image, round down to the nearest half-inch. When calculating resolution, round up to the closest multiple of 25. Calculations should reflect the dimensions of actual item to be scanned, not its matting or support.

0%	255					
10%	229					
20%	203					
30%	179					
40%	159					
50%	128					
60%	102					
70%	76					
80%	52					
90%	26					
100%	0					
		Verso				
		1.0				
		2000				
		1.5				
		1350				
		2.0				
		1000				
		2.5	800			
		3.0	675			
		3.5	575			
		4.0	500			
		4.5	450			
		5.0	400			
		5.5	365			
		6.0	335			
		6.5	310			
		7.0	290			
		7.5	270			
		8.0	250			
		8.5	240			
		9.0	225			
		9.5	215			
		10				
		200+				
				1	3000	
				1.5	2000	
				2.0	1500	
				2.5	1200	
				3.0	1000	
				3.5	875	
				4.0	750	
				4.5	675	
				5.0	600	
				5.5	550	
				6.0	500	
				6.5	475	
				7.0	450	
				7.5	400	
				8.0	375	
				8.5	375	
				9.0	350	
				9.5	325	
				10	300	
				1	6000	
				1.5	4000	
				2.0	3000	
				2.5	2400	
				3.0	2000	
				3.5	1750	
				4.0	1500	
				4.5	1350	
				5.0	1200	
				5.5	1100	
				6.0	1000	
				6.5	925	
				7.0	900	
				7.5	800	
				8.0	750	
				8.5	725	
				9.0	675	
				9.5	650	
				10	600	
				1	5000	
				1.5	3500	
				2.0	2500	
				2.5	2000	
				3.0	1675	
				3.5	1450	
				4.0	1250	
				4.5	1125	
				5.0	1000	
				5.5	925	
				6.0	850	
				6.5	775	
				7.0	725	
				7.5	675	
				8.0	625	
				8.5	600	
				9.0	575	
				9.5	550	
				10	500	
				1	4000	
				1.5	3000	
				2.0	2000	
				2.5	1600	
				3.0	1350	
				3.5	1150	
				4	1000	
				4.5	900	
				5.0	800	
				5.5	750	
				6.0	675	
				6.5	625	
				7.0	575	
				7.5	550	
				8.0	500	
				8.5	475	
				9.0	450	
				9.5	425	
				10	400	



## 4. Scan

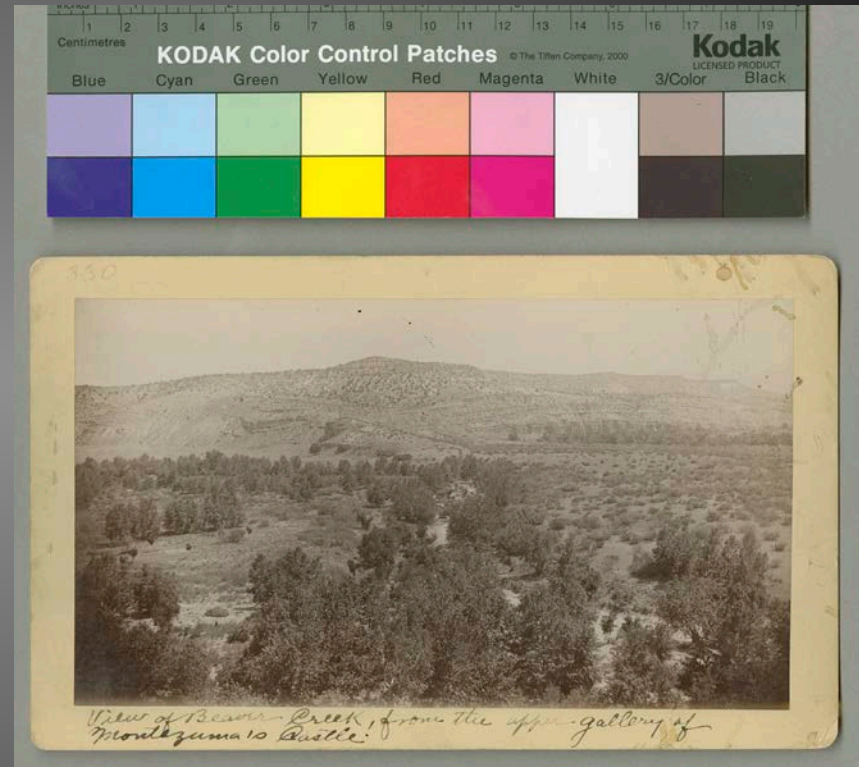
- Once you have determined the correct resolution, place the object on the glass, (for photographic material, face down) about an inch below, in the center of, the color bar. Place the background board over the scanner glass and carefully close the scanner lid.





## 4. Scan

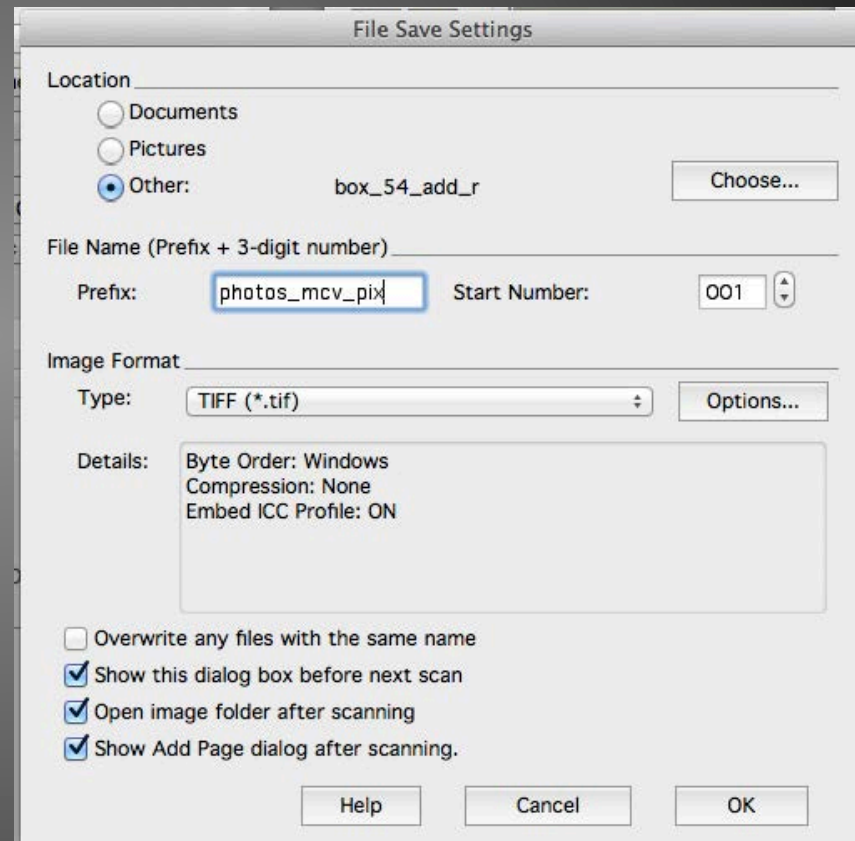
- Check your endpoints (white and black) to make sure you are within the correct range.
- If your endpoints are within the range, select the area to scan – you do not need to scan the entire scanner bed, just the object and the color bar target. Create a box with the dotted line around the color bar target and the object, with about 1/4 inch border of grey background around the edge.





# 5. Save

- Assign appropriate filename and save images into the appropriate folder location





# Extra – Batch Scanning

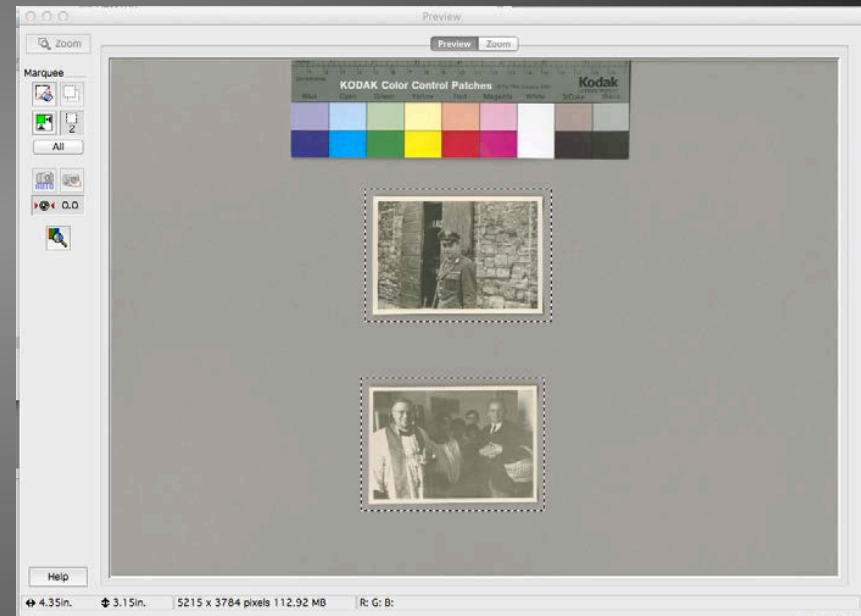
- If multiple objects need to be scanned, and more than one fit on the scanner at the same time, they can usually be batched scanned.
- **When determining resolution, set your resolution to align with the SMALLEST OBJECT ON THE GLASS, so that all objects reach the required pixel dimensions.**
- Arrange the objects on the glass far enough apart so you will be able to get a 1/4 inch border around each.





# Extra – Batch Scanning

- In the preview window, crop ALL images so there is a dotted border around each.
- When batch scanning, you do not need to include the color bar target in the scans.





Smithsonian  
National Museum of Natural History

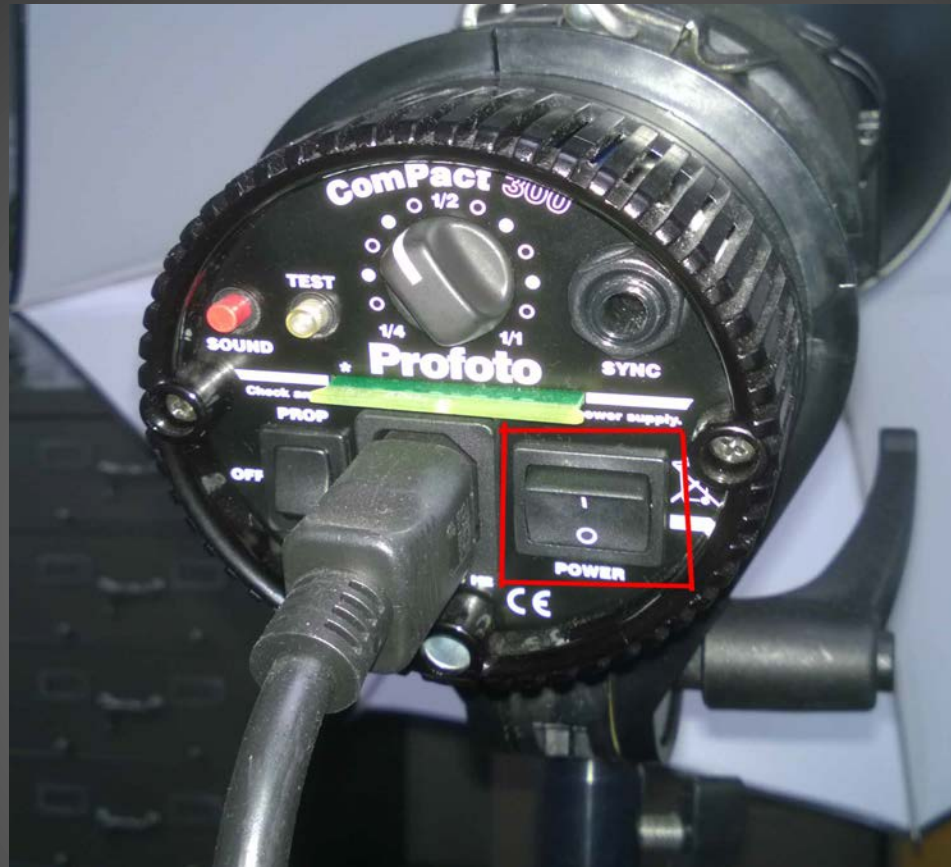


## Bonus: Abridged Camera Workflow

Turn on Nikon camera (slide control on the camera ).

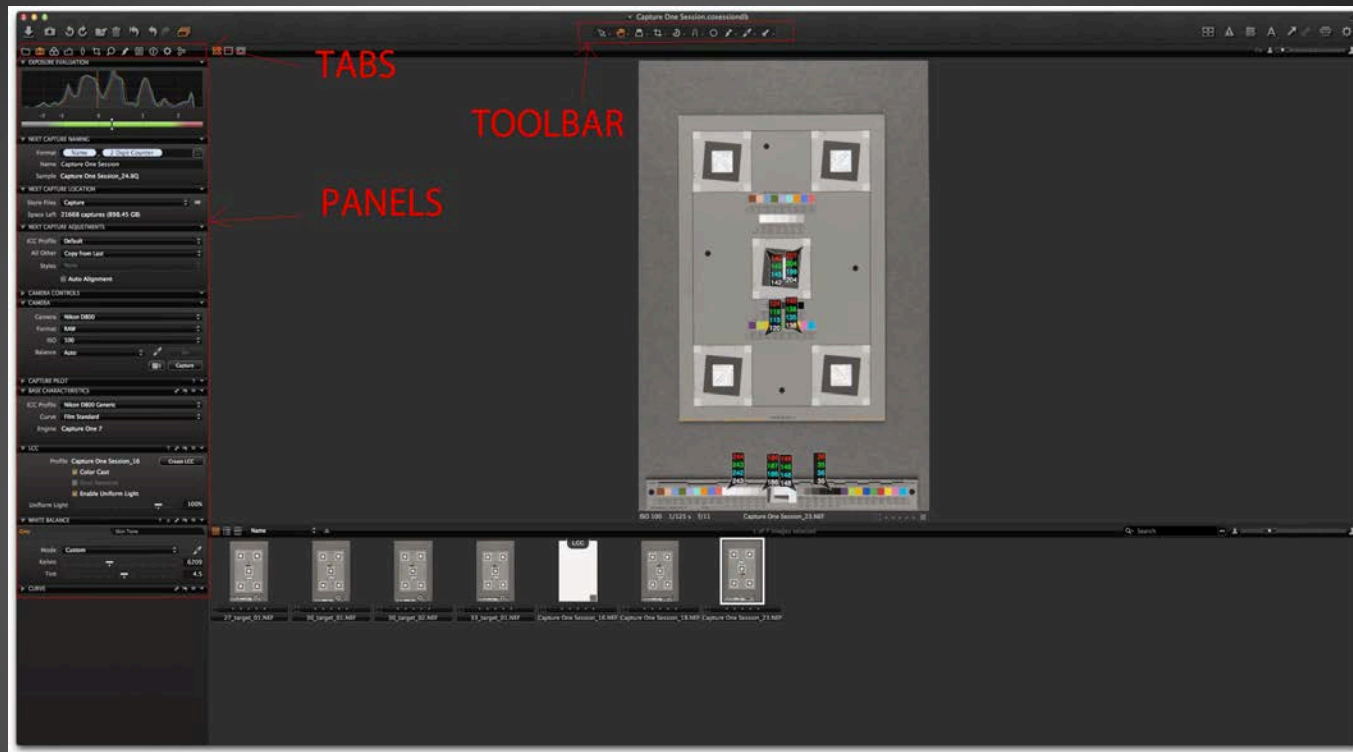
*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*





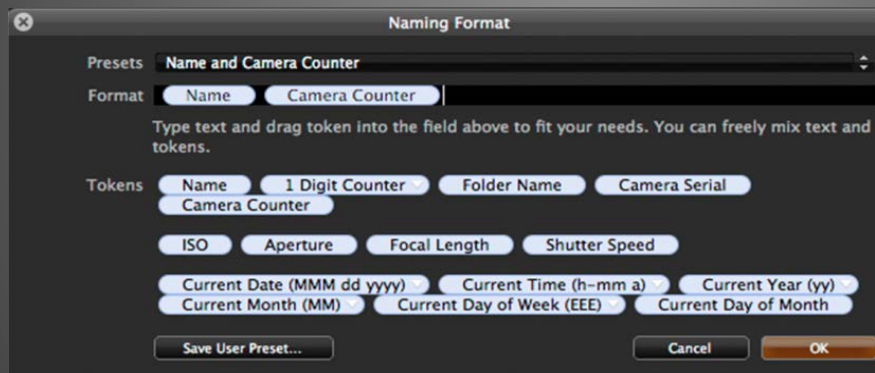
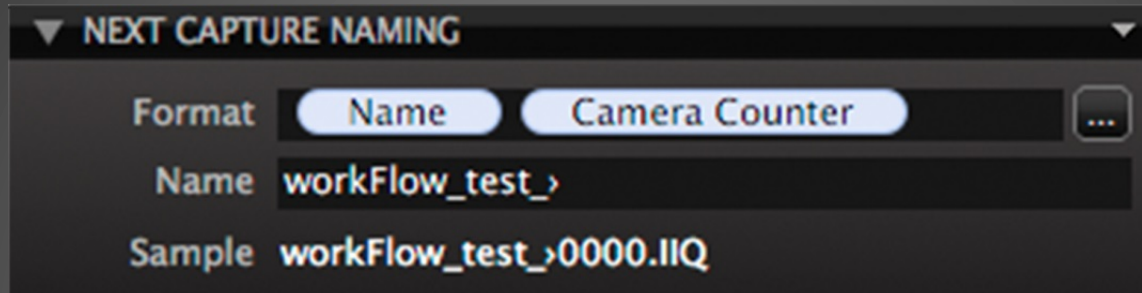
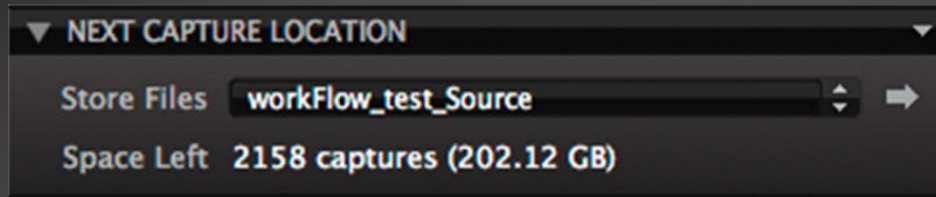
## Bonus: Abridged Camera Workflow

Turn on both Profoto lighting fixtures (black switches on the base near power cord)



## Bonus: Abridged Camera Workflow

Launch Capture One software



## Bonus: Abridged Camera Workflow

Setup camera counter and filenaming in Capture One



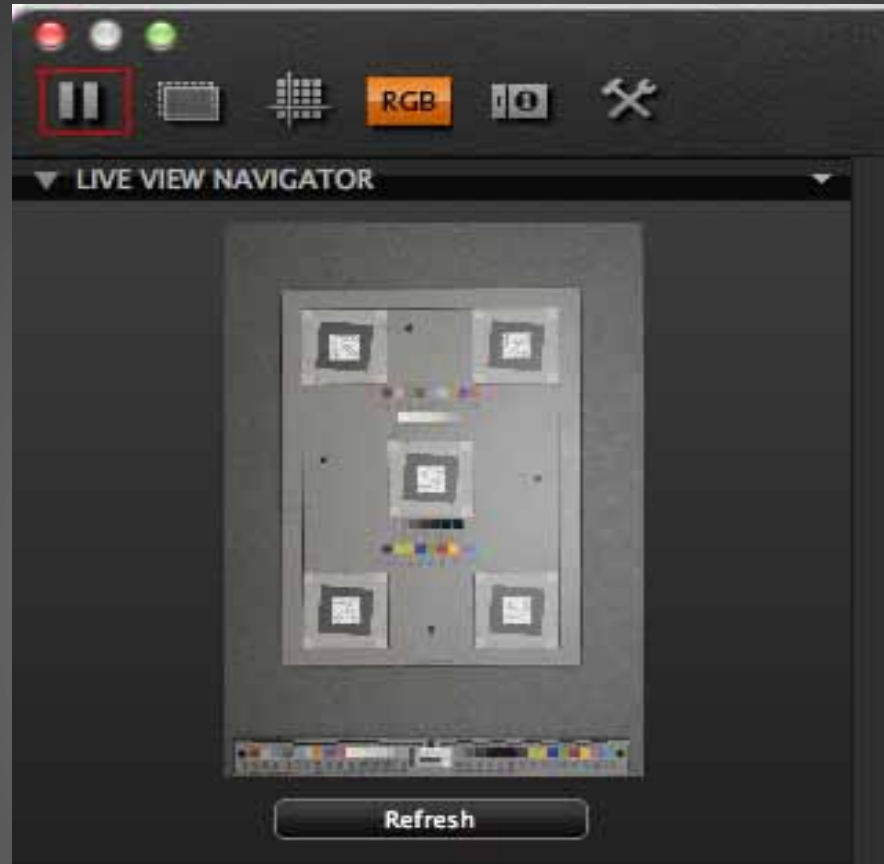
Smithsonian  
National Museum of Natural History



## Bonus: Abridged Camera Workflow

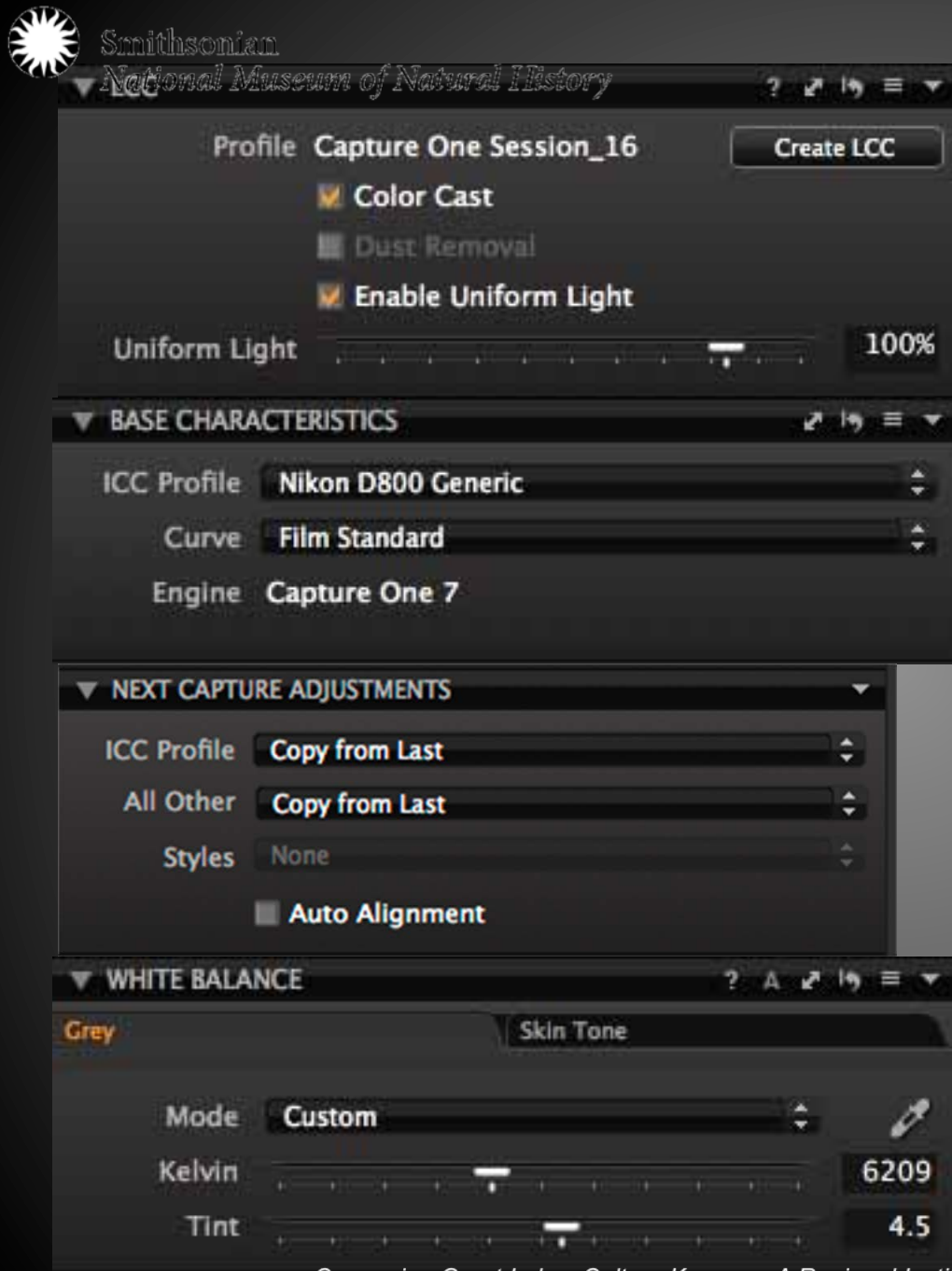
Rotate the 'Focus Ring' on the Nikon's lens to focus. Note that the image will be very pixelated even when it is focused. Once image is focused, hit command+K or Capture key

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



## Bonus: Abridged Camera Workflow

Using the Grayscale on the Device/Object Target, make sure your whites are between 242 and 248 or whatever is appropriate for your subject matter.



Create Lens Cast Calibration

Set Base Characteristics

Assign Capture Adjustments

Set White Balance



Smithsonian  
National Museum of Natural History

**SHARPENING**

Amount  93

Radius  0.2

Threshold  0.0

**NOISE REDUCTION**

Luminance  50

Color  50

Single Pixel  0

**NOISE REDUCTION ADVANCED**

Details  0

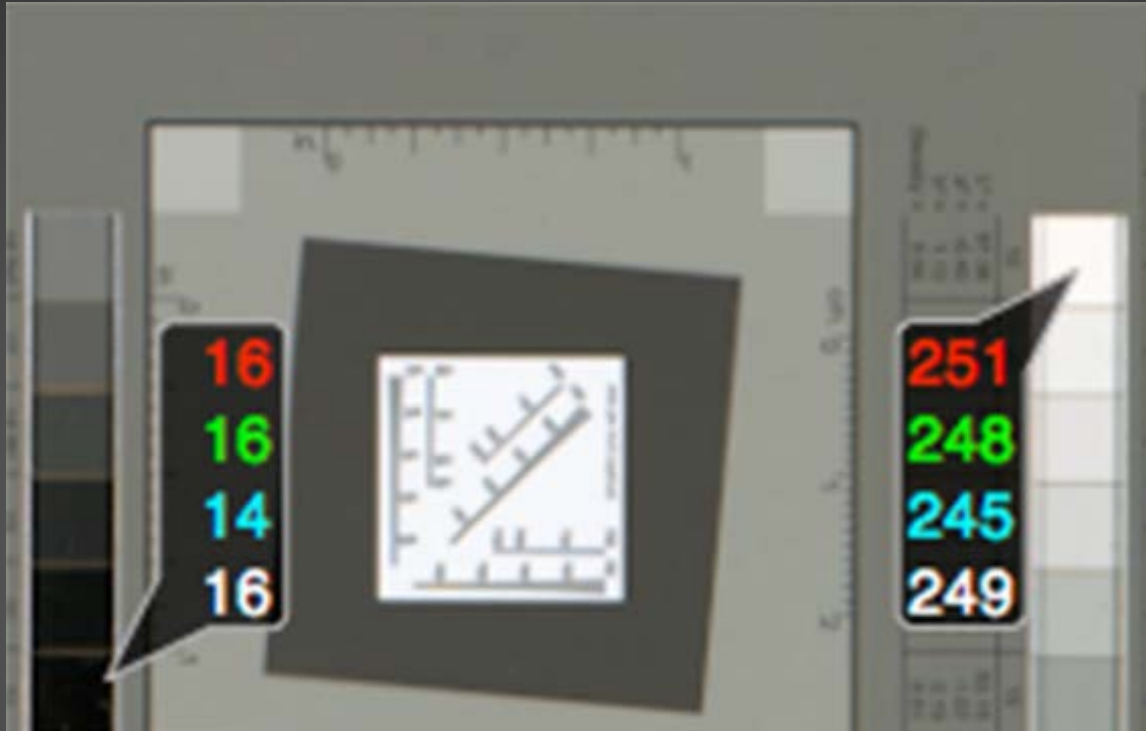
Fine Grain  0

**MOIRE**

Amount  0

Pattern  8

Bonus: Abridged Camera  
Workflow  
Settings in Details Tab



## **Bonus: Abridged Camera Workflow**

Place a series of RGB read outs by clicking various RGB patches on the Device Level and/or Object Level Targets

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*





# Image Creation – Conclusions

- Considerations:
  - Time
  - Staff
  - Hardware and Software
  - Imaging Environment
  - Bit depth
  - Tone
  - Color Mode
  - Resolution
  - File Formats

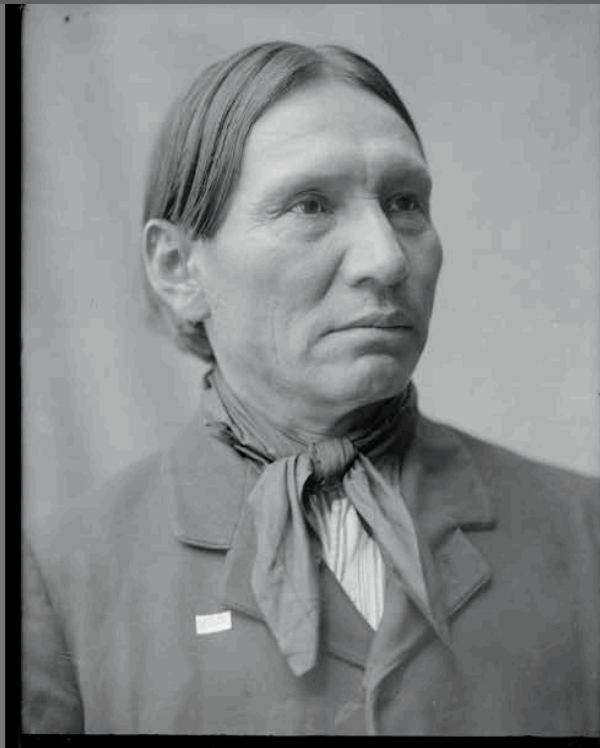
Don't forget to use a target!



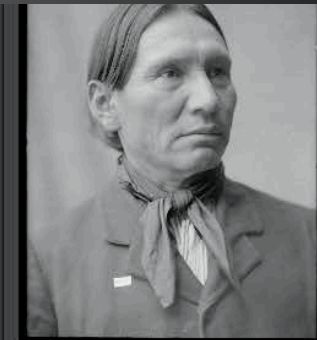


# More is More: The Role and Responsibility of Metadata

Ratings	
No Rating	174
Keywords	
No Keywords	446
Orientation	
Landscape	39
Portrait	135
Aspect Ratio	
3:4	36
4:5	134
5:7	4



T13339.tif  
28.82 MB  
4860 x 6201 @ 725 ppi  
TIFF image  
Gray Gamma 2.2



T13339.tif

Metadata	
f/---	4860 x 6201
---	28.82 MB 725 ppi
---	ISO---
Gray Gamma 2.2 B&W	

File Properties	
Filename	T13339.tif
Document Type	TIFF Image
Application	Adobe Ph...acintosh
Date Created	1/7/09, 2:50:30 PM
Date File Modified	1/7/09, 2:50:32 PM
File Size	28.82 MB
Dimensions	4860 x 6201
Dimensions (in inches)	6.7" x 8.6"
Resolution	725 ppi
Bit Depth	8
Color Mode	B&W
Color Profile	Gray Gamma 2.2

IPTC Core	
Creator	Nationa...rchives
Creator: Job Title	
Creator: Address	Smithsonian Institution Museum Support Center 4210 Silver Hill Road
Creator: City	Suitland
Creator: State/Province	MD
Creator: Postal Code	20746



# What is metadata?

- Literally “data about data”
- Descriptive information about collections
- value-added information used to arrange, describe, track, and enhance access to collections
- Metadata is more common in daily online life than you might think:
  - Photo tagging (Facebook, Instagram)
  - #Hashtags (Twitter, Instagram, Facebook)
  - Blog tags

# What are the different types of metadata?

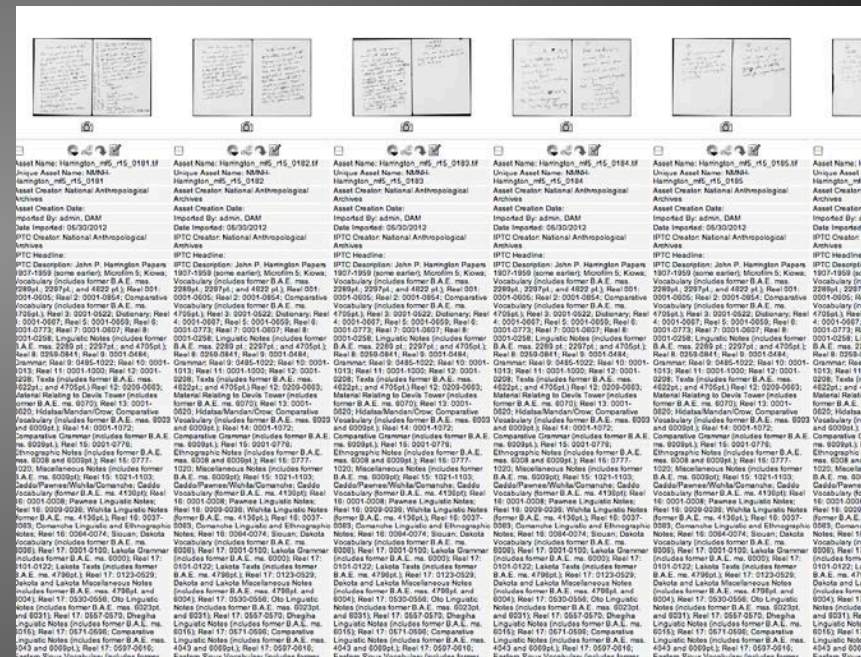
(Source: Source: Baca, Murtha, editor. Gill, Tony, Gilliland, Anne J., Whalen, Maureen, and Woodley, Mary S. (2008). "Setting the Stage," Introduction to Metadata Online Edition, Version 3.0, [http://www.getty.edu/research/publications/electronic\\_publications/intrometadata/index.html](http://www.getty.edu/research/publications/electronic_publications/intrometadata/index.html))

Type	Definition	Example
Administrative	Metadata used in managing and administering collections and information resources	<ul style="list-style-type: none"> <li>- Acquisition information</li> <li>- Rights and Reproduction tracking</li> <li>- Documentation of legal access requirements</li> <li>- Location information</li> <li>- Selection criteria for digitization</li> </ul>
Descriptive	Metadata used to identify and describe collections and related information resources	<ul style="list-style-type: none"> <li>- Cataloging Records</li> <li>- Finding Aids</li> <li>- Differentiations between versions</li> <li>- Specialized indexes</li> <li>- Curatorial information</li> <li>- Hyperlinked relationships between resources</li> <li>- Annotations by creators and users</li> </ul>
Preservation	Metadata related to the preservation management of collections and information resources	<ul style="list-style-type: none"> <li>- Documentation of physical condition of resources</li> <li>- Documentation of actions taken to preserve physical and digital versions of resources</li> <li>- Documentation of any changes occurring during digitization or preservation</li> </ul>
Technical	Metadata related to how a system functions or metadata behaves	<ul style="list-style-type: none"> <li>- Hardware and software documentation</li> <li>- Technical digitization information</li> <li>- Tracking of system response times</li> <li>- Authentication and security data</li> </ul>
Use	Metadata related to the level and type of use of collections and information resources	<ul style="list-style-type: none"> <li>- Circulation records</li> <li>- Physical and digital exhibition records</li> <li>- Use and user tracking</li> <li>- Content reuse and multiversioning information</li> <li>- Search logs</li> <li>- Rights metadata</li> </ul>



# Where is metadata?

- Different levels of metadata exist in different places
  - Embedded into files
  - Filing cabinets!
    - Accession Records
    - Catalog Cards
  - Collections Information Systems and Databases
    - Online, public-facing platform
    - Internal collections database system





# When is metadata?

- Collections staff has been working with metadata since the beginning – term not commonly utilized until collections started being digitized
- Metadata should be collected and applied to files as soon as possible – lag time between image creation and metadata enhancement could result in inconsistencies and missing information
- Build metadata into digitization workflow

<b>ETHNOLOGY</b> 2130	<b>NAME</b> Dressed Skin with Native Drawing	<b>NO. SPEC.</b> 1
<b>CAT. NO.</b>	<b>PEOPLE</b>	
	<b>LOCALITY</b> Upper Missouri	
<b>ACC. NO.</b>	<b>COLLECTOR</b> War Dept. U.S.	
	<b>ACQUIRED</b>	<b>DATE</b> Dec., 1866
	<b>PLACED</b>	<b>SIZE</b> L. 6' 6" W. 6'
<b>ORIG. NO.</b> MNH 1272-C 75-6963; -6964 (detail) 78-15894 78-15892 78-15890	<b>REMARKS</b> "2130 Indians of the Missouri A large Dressed Skin ornamented with color drawings by natives. US. War Dept." Hidatsa?	
<b>NEG. NO.</b> 78-15893 78-15891		
<b>SI-MNH-175A</b> 5-3-65	<b>INVENTORIED 1978</b>	<b>over:</b>
	<b>SEE HISTORY CARD FOR ADDITIONAL DATA</b>	



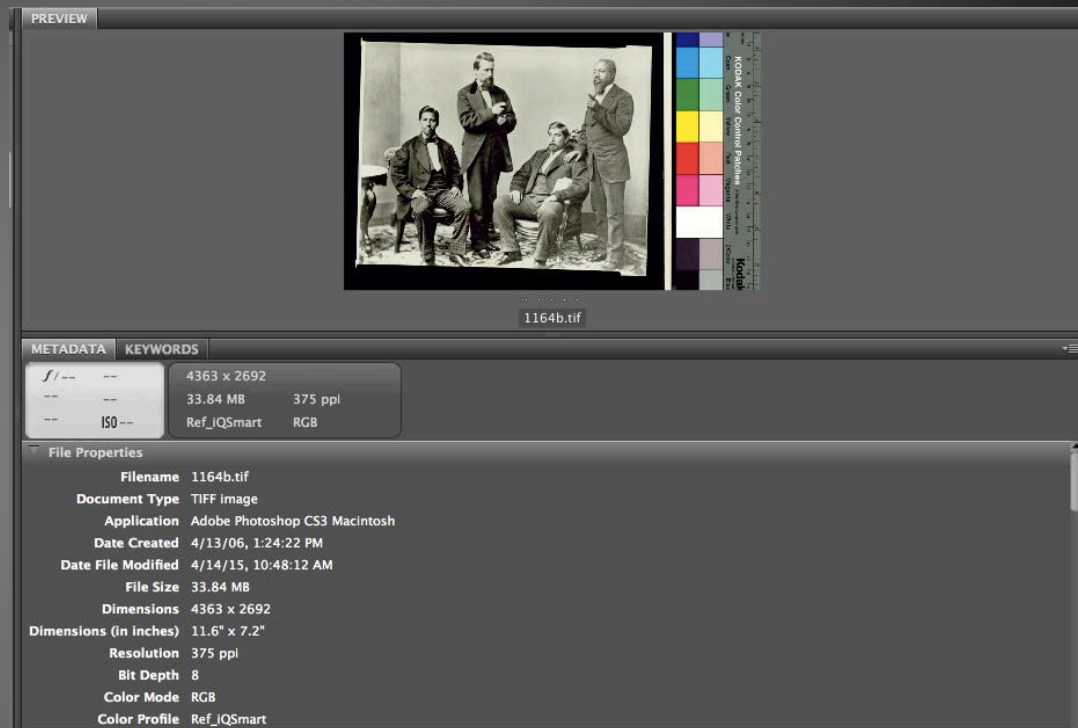
# Who is metadata?

- Standards bodies
  - JEITA
  - IPTC
  - PLUS
- Photographer trade associations
  - SAA
  - ASMP
  - APA
  - EP
  - NPPA
- Cultural Heritage and Government Institutions
  - AAT (Getty)
  - LOC
  - FADGI
- When working with metadata, some “who’s” to think about:
  - Who is supplying the metadata?
    - Curators
    - Donors
    - Collections staff (archivists, librarians)
    - Photographers/Imaging Technicians
  - Who is applying the metadata at its various levels?
    - Catalog records
    - Database entries
    - Embedded files
  - Who is using the metadata?
    - Staff
    - Researchers
    - Exhibitions
    - Publications



## Why metadata?

- Accessibility
- Discoverability
- Collections Management
- Usability
- Rights and Restrictions
- Standard Policies and Procedures







75-33.tif

Description | IPTC | Raw Data | IPTC Extension | Camera Data | GPS Data | Video Data | Audio Data | Mobile SWF | Categories | Origin

Document Title:

Author: National Anthropological Archives

Author Title:

Description:

Rating: ★ ★ ★ ★ ★

Description Writer:

Keywords:

ⓘ Semicolons or commas can be used to separate multiple values

Copyright Status: Unknown

Copyright Notice: The Smithsonian Institution continues to research information on its collections. Contact Smithsonian for current status.

Copyright Info URL:  Go To URL...

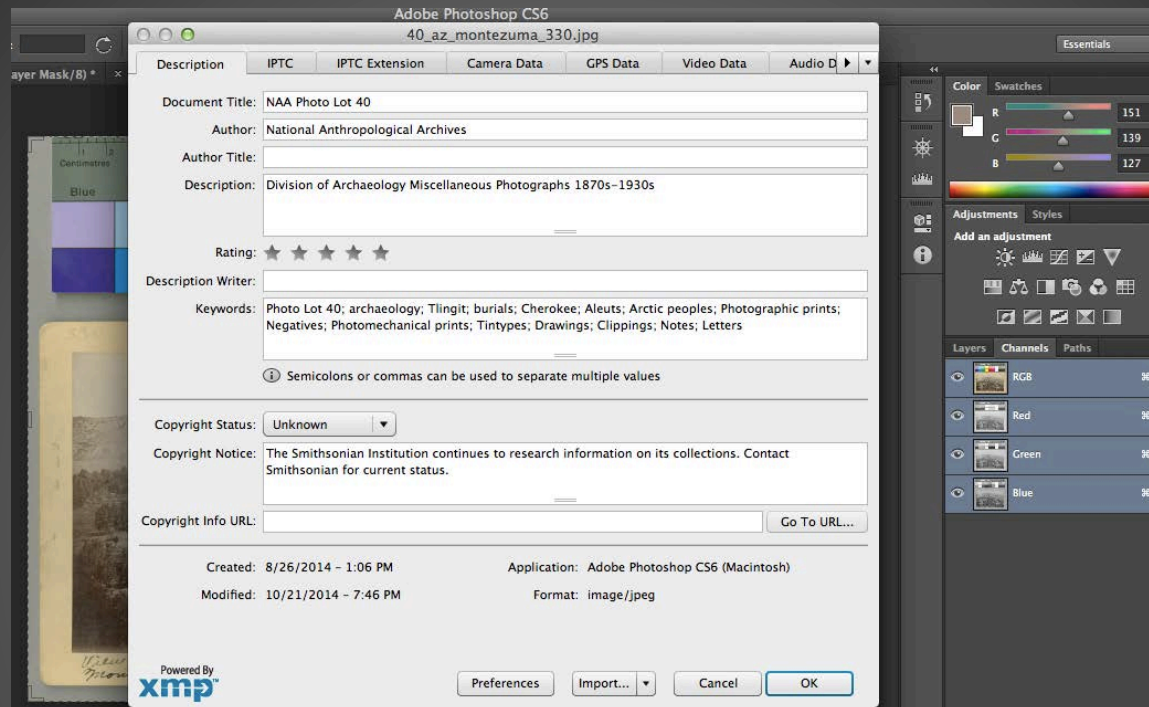
Created: 6/13/2000 - 10:33 AM      Application: Adobe Photoshop CS6 (Macintosh)  
Modified: 10/30/2013 - 10:45 AM      Format: image/tiff

Powered By **xmp**

Preferences | clean\_metadata\_NAA | Cancel | OK

## How metadata?

- IPTC: International Press Telecommunications Council
- Dublin Core
- Descriptive and Technical
- Embedded



## How metadata?

Basic Guidelines for Minimal Descriptive Embedded  
Metadata in Digital Images  
by EMDaWG (Embedded Metadata Working Group –  
Smithsonian Institution)  
April 2010



# How metadata? Requirements

## Required: Core Set of Embedded Metadata

Element Name	Definition	Sample Data Value	Notes	Character Limit*
<b>Document Title</b>	File number, Accession Number, Catalog Number, Negative Number, Unique Identifier root level etc.	LB016021-a 08596201 gn_03644 landes_photo_arizona_16 digital filename 23456.000 *catalog number  123457.000;1234568.000 *multiple objects in one image  P00001 *catalog number N00001 *catalog number T00001 *catalog number 20100121_01a_csf_ps_001.tif *file name. represents coverage of event at NMAI  SFF2009_Strauss_6-24_0004.dng		*IPTC fields have character limits depending on the application utilizing the standard and not all data may be displayed. (Add char limit info into the tables) With some applications data may be truncated at a given character limit.
<b>Copyright Notice</b>	Copyright Notice	This image was obtained from the Smithsonian Institution. Unless otherwise noted, this image or its contents may be protected by international copyright laws.  <i>*This statement is a default statement and you may use a more specific statement; check with OGC for a specific statement.</i>	Approved and suggested by OGC.	
<b>Source</b>	Name and Abbreviation of SI owning unit, Smithsonian Institution	NMAI- Natl. Museum of the American Indian, Smithsonian Institution NAA- Natl. Anthropological Archives, Smithsonian Institution		
<b>Creator</b> (*Note: Unit makes decision-document within unit how they reached this decision).  In the DAMS the IPTC Creator field is mapped to the Asset Creator field, which is the creator of the digital object.	Creator of digital object or Creator of original object	Edward F. Caldwell and Co.  Department of Anthropology  National Anthropological Archives  Photographer Name <i>*If author is not known then default to department name (refrain from using acronyms)</i>  Cynthia Frankenburg <i>*if Creator=name then Creators Title field is populated. Job Title=Photographer</i>  William Greene <i>*if Creator=name then Creators Title field is populated. Job Title=Scanner</i>  Woody Guthrie <i>*Creator original object</i>	*IPTC Creator Job Title field can be used to define the role of the creator.	



# How Metadata? Suggested

Element Name	Definition	Sample Value	Notes	Character Limit
<b>Date</b> <small>In the DAMS IPTC Date is mapped to Asset Creation Date, which is the creation date of the digital object.</small>	Date of Object or Date of Creation of Digital	07/01/1967	Description field can be used if date range or other date structure is being used.	
<b>Description</b>	Free narrative text	123457.000 (right); 123458.000 (left) *see above under Title example. This field is used to locate individual objects within an image that contains multiple objects.  Cultural Resources Center 2007 Powwow Open House, CRC Open House, CRC Exterior, Chief Joseph, Nez Perce		
<b>Keywords</b>	Free text field but should be used to store a list of standard term(s) separated by a common delimiter such as semicolon.	Lighting Archive; Electrification; Lighting; Lighting Fixtures; Architectural History; History of Architecture  Object; Publication; Our Lives <small>* taken from defined look-up list in database: <b>Object</b>=image of object in the collection; <b>Publication</b>= quality suitable for publication; <b>Our Lives</b>=imaged for <b>Our Lives</b> exhibition.</small>  Alice Fletcher; Francis La Flesche; 4558  Viento de Agua; plena; bomba  rugby; Wales; sports	This list can come from any existing controlled vocabularies like your unit CIS' Iconography lookup list, public resources such as Library of Congress Subject Headings, taxonomic checklists, etc. The goal is to be <b>consistent</b> as this is a field whose data is often used for searching. For instance, if you use singular form, stick with singular form, don't alternate between singular and plural. Don't alternate between variations, like US, USA, or United States. *(See below for links to controlled vocabularies for consideration.)	
<b>Credit/Provider</b>	What you would like to accompany the image in a publication. Ex: Image Number, SI owning Unit, Smithsonian Institution	Image Number, National Anthropological Archives, Smithsonian Institution  Papers of Ruth Landes, National Anthropological Archives, Smithsonian Institution  Ken Rahaim, Smithsonian Institution  ** Ernest L. Spybuck (Absentee Shawnee, 1883–1949), <i>Procession before War Dance</i> , ca. 1910. Watercolor on paperboard, 42.2 x 63.9 cm. Oklahoma. Photo by David Heald. 2/5735		
<b>Job Identifier</b>	Instructions or unit id for a job	MSC07-04608		
<b>Headline</b>	(Formally called Caption) A descriptive title or a caption.	Dr. J.E. Tallimage		



# Locating metadata content

The screenshot shows a web browser window displaying the SIRIS search results for the query 'naa ms 3718'. The page title is 'Fourth of July, 1890'. The search bar shows the query and a 'Refine Search' button. Below the search bar, there are navigation tabs for 'Search', 'Search Images', and 'About'. A secondary navigation bar includes 'Keyword', 'Browse', 'Finding Aids', 'Combined', 'Browse Images', 'Search History', and 'All Catalogs'. The search results are displayed in a table-like format with a 'Who else has...' sidebar on the left and a main content area on the right. The main content area includes a thumbnail image of a document and a table of 'Item Information' at the bottom.

**Search:** General Keyword  Refine Search

> You are only searching: Archives, Manuscripts and Photographs

**More Smithsonian Searches**

**Who else has...**

- Pico, Juan Estevan Chumash
- Chumash
- Indians of North America
- Chumash Indians
- Language and languages – Documentation

**Fourth of July, 1890** Add to my list

Creator: [Pico, Juan Estevan Chumash](#)

Title: Fourth of July, 1890

Phy. Description: 20 pages

Click to view image set:

Additional forms: Micro Reel 7; also separate 35 mm negative strip.

Summary: Letter in Spanish to H.W. Henshaw, San Buenaventura, California, April 21, 1891, written in parallel columns in the San Buenaventura (Ventureño Chumash) language and in Spanish. Typed English Translation, 19 pages.

Language Note: Spanish and San Buenaventura in parallel columns.

Cite as: Manuscript **3718**, National Anthropological Archives, Smithsonian Institution

Publication: Published by Heizer, University of California Anthropology Records, 15:2, 1955, pages 187-193.

Culture: [Chumash Ventureño](#)  
[Indians of North America – California](#)  
[Chumash Indians](#)

Subject-Topical: [Language and languages – Documentation](#)

Repository Loc: National Anthropological Archives, Smithsonian Museum Support Center, Suitland, Maryland

Local Number: **NAA MS 3718**

Co-Creator: [Henshaw, Henry Wetherbee](#) [address](#)

Item Information	
Repository	Call No.
National Anthropological Archives	NAA MS 3718 <span style="float: right;">Add Copy to MyList</span>

Format:  HTML  Plain text  Delimited

Subject:



# Creating a metadata template

- Open a file from the folder you're working on, and go to File→File Info in Photoshop. Using the EMDAWG guidelines, the following fields are populated:
- Document Title: The number of the manuscript (from the Local Number field in SIRIS).
- Author: National Anthropological Archives (EMDAWG – Creator)
- Description: The title of the manuscript, also from the SIRIS record.
- Keywords: The creator of the manuscript; a co-creator (if one exists); MS and its number; the culture the manuscript pertains to; and the subject-topical phrases. These all come from the SIRIS record
- Copyright Status: UNKNOWN. It is extremely important that the copyright status says UNKNOWN.
- Copyright Notice: "The Smithsonian continues to research information on its collections. Contact Smithsonian for current status." This is the same for all material.
- In the IPTC tab, make sure the Source field is populated. It should say, "NAA-Natl. Anthropological Archives, Smithsonian Institution."

3718\_a\_01v.tif

Description | IPTC | IPTC Extension | Camera Data | GPS Data | Video Data | Audio D | ▾

Document Title: NAA MS 3718

Author: National Anthropological Archives

Author Title:

Description: Fourth of July, 1890

Rating: ★ ★ ★ ★ ★

Description Writer:

Keywords: Juan Estevan Chumash Pico; Henry Wetherbee Henshaw; MS 3718; Chumash Ventureño; Chumash Indians; Language and languages; Documentation

ⓘ Semicolons or commas can be used to separate multiple values

Copyright Status: Unknown ▾

Copyright Notice: The Smithsonian continues to research information on its collections. Contact Smithsonian for current status.

Copyright Info URL: Go To URL...

Created: 4/10/2013 - 1:44 PM Application: Adobe Photoshop CS6 (Macintosh)

Modified: 4/30/2013 - 3:08 PM Format: image/tiff

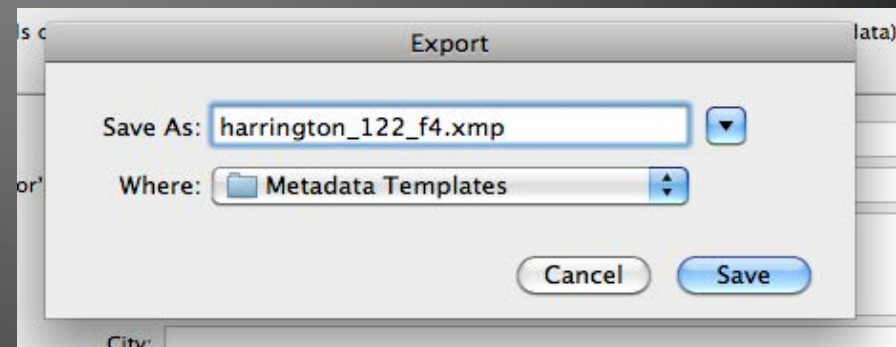
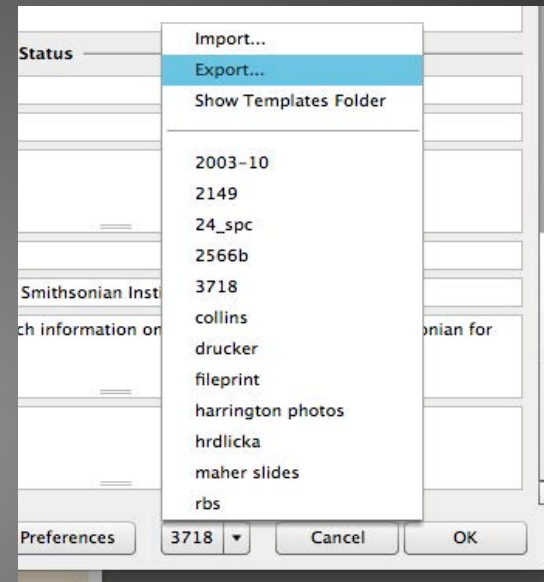
Powered By **xmp**

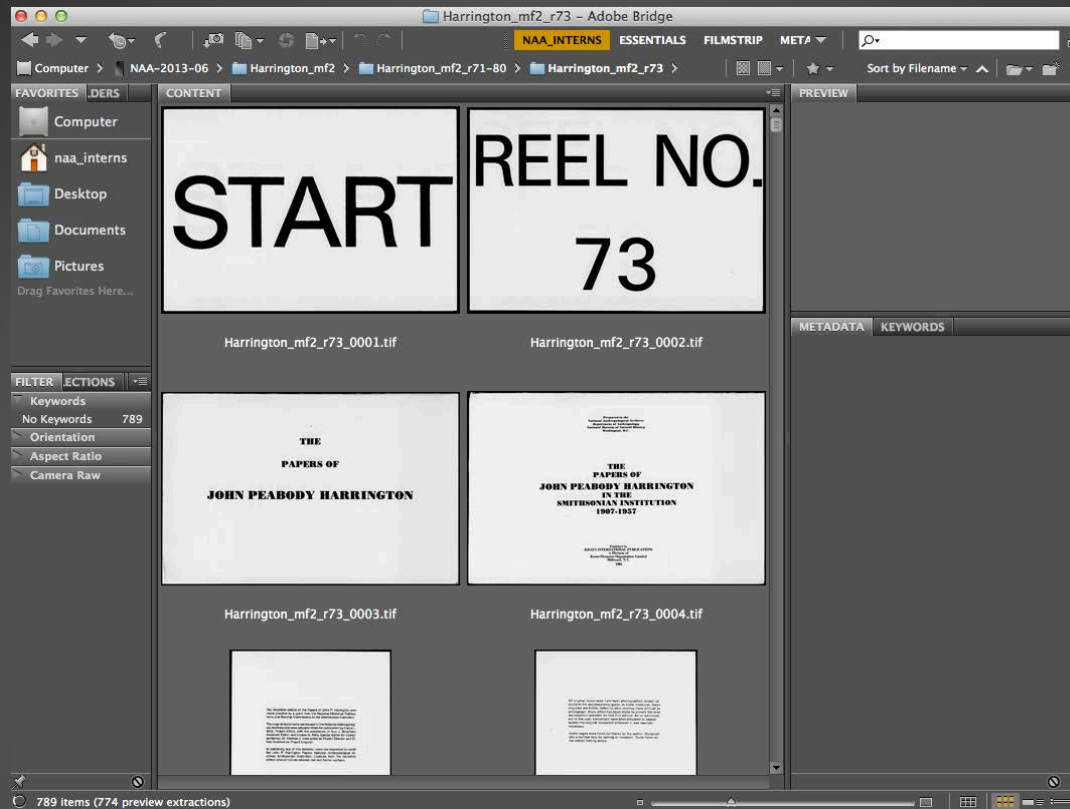
Preferences 3718 ▾ Cancel OK



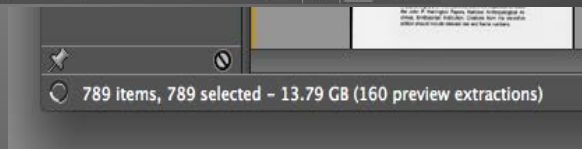
# Exporting Metadata

- Once you have created the metadata template, you will want to Save (or Export) it for later use.
- In the metadata window, click Export. A new Save window will pop up, and it will automatically fill in the filename of the file currently open.
- You may leave it as such, or make it shorter and simpler, like just using the number of the manuscript for which the metadata applies.
- Click save. You have created a metadata template that can be used on a collection of images as needed.





## Embedding Metadata



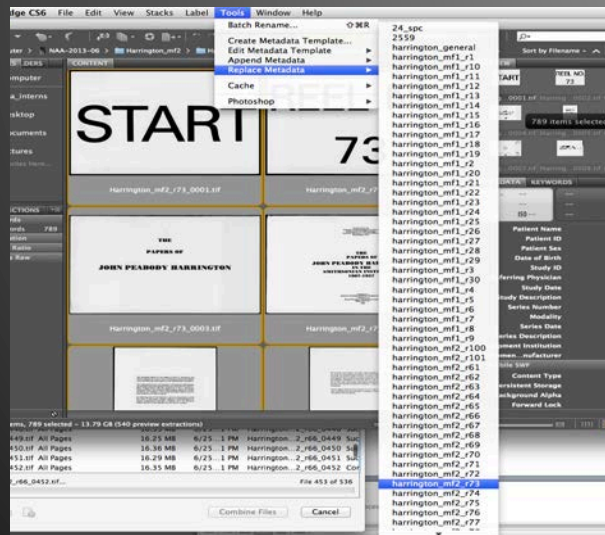
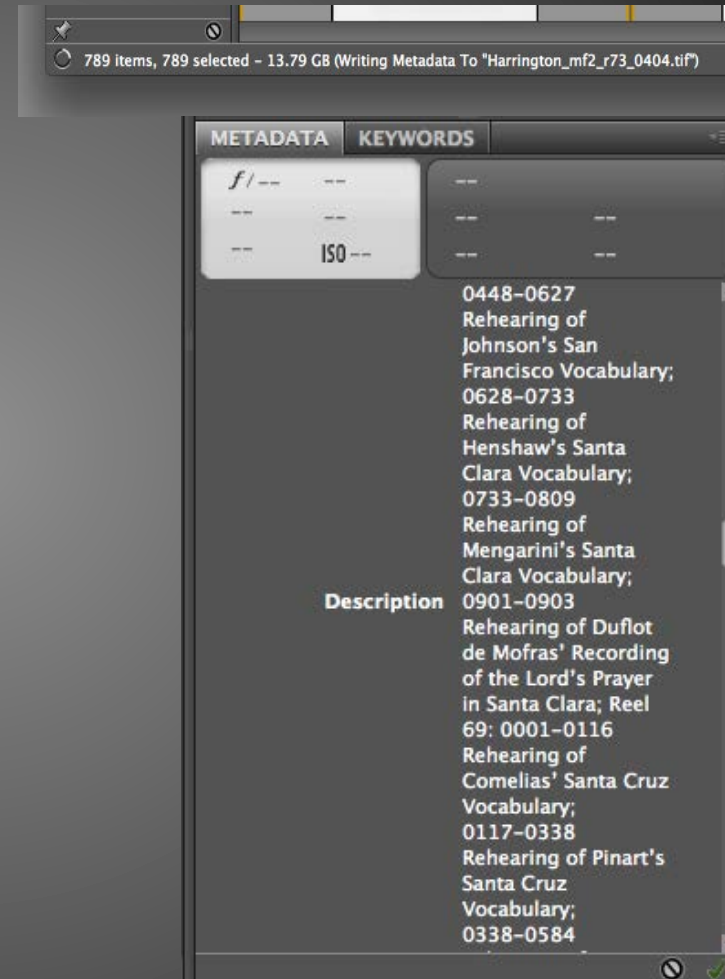
Metadata can be embedded more quickly embedded using Adobe Bridge  
Open the files in Adobe Bridge. The files may take several minutes to load and go through the preview extraction.  
Once all the files have loaded, select all files (command + A) then go to Tools → "Replace Metadata"





# Embedding Metadata

- Select the corresponding metadata template. Embedding will begin automatically and may take a few moments to complete. Once completed, bridge will automatically save the metadata.
- On the lower right hand side, confirm that the metadata has been embedded.
- Confirm that metadata has been embedded by opening the files in Adobe Photoshop and checking the “File Info.”





# Metadata - Conclusions

- Metadata is already an integral part of Collections Management
- Embedding and Utilizing Metadata will increase accessibility, usability, discoverability, organization, alleviate rights and restrictions issues





Smithsonian  
*National Museum of Natural History*

# Quality Control and Post-Processing

How much is too much? How  
much is enough?

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# What is Quality Control?

- Series of processes done after digitization to ensure you are meeting your predetermined standards and benchmarks, while maintaining accuracy, quality, and consistency.



# Quality Control Checklist

Kenney, Anne R. and Oya Y. Reiger. "Establishing a Quality Control Program," Chapter 4, pp.61-73. (2000).  
Moving Theory into Practice: Digital Imaging for Libraries and Archives. Mountain View, CA: Research  
Libraries Group.

## Prerequisites for QC

- Identify products and goals
- Agree on Standards
- Determine a Reference Point
- Understand the Limitations of Current Knowledge, Practice and Technology

## Setting up your QC Program

- Identify Scope
- Determine Methods
- Evaluate System Performance
- Codify Inspection Procedures
- Control QC Environment
  - Hardware configuration
  - Image-display software
  - Monitor setup
  - Color quality control
  - Color management
  - Viewing conditions
  - Human characteristics

## Assessing Image Quality

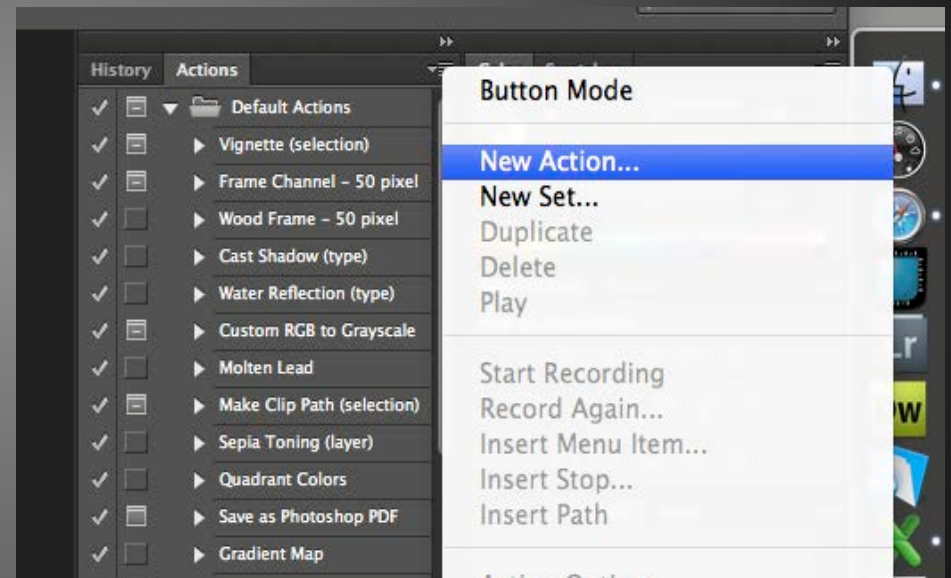
- Evaluate:
  - Resolution
  - Color and Tone
  - Overall appearance



# Photoshop - Actions

- Post-processing needs like conversion (16-bit to 8-bit), rotation, embedding metadata, curve adjustments can be performed in batch actions in Photoshop.

Open Photoshop. Go to the Action tab on the right side and click the dropdown arrow. Click "New Action."

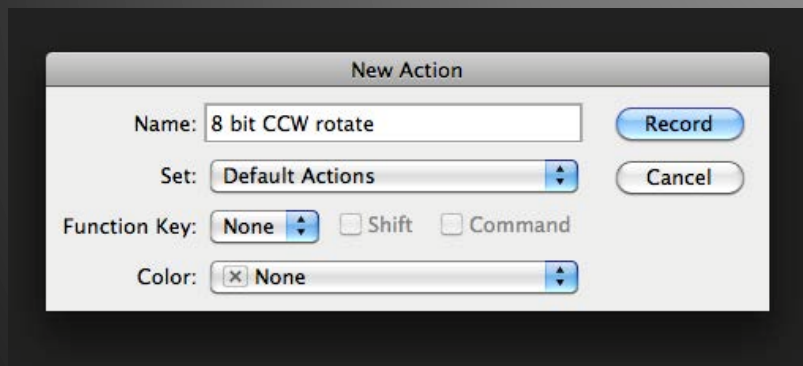




# Creating an Action

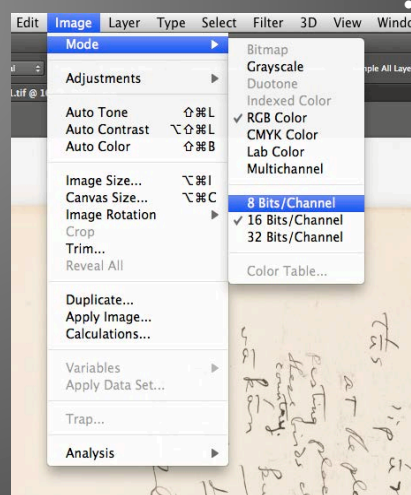
## Naming the Action

- A new naming window will pop up. Create a simple, explanatory name for the action, such as “8 bit and CCW rotate.”
- Click “record.” Photoshop will record all of the steps you complete and save them as a single action that can be applied to an entire folder of images for faster, more efficient processing.

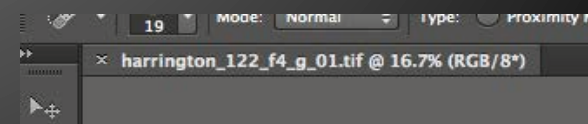


## Running the Action

- Go to File → Open and open the first image in the folder you are processing.
- Go to Image → Mode → 8 Bits/Channel



Once you have converted the image to 8 Bits/Channel, you may notice that Photoshop will reflect that in the tab where the filename of the image is (RGB/8\* indicates this):

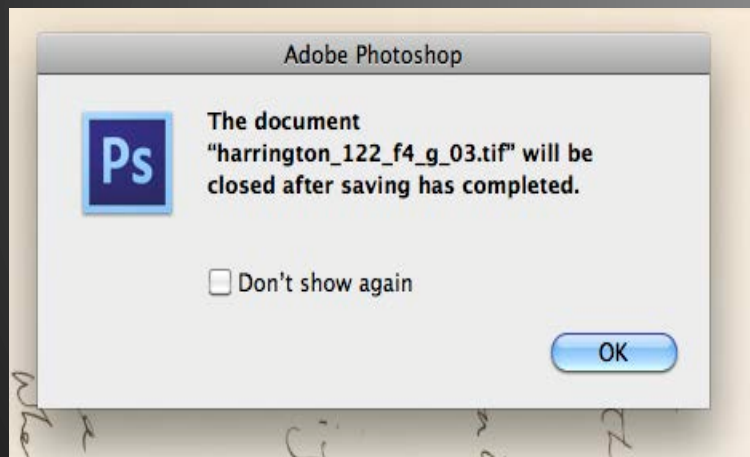




# Completing an Action

## Saving the Changes

- Go to File → Save. Once the dropdown menu has disappeared, go to File → Close. If you get the following popup window, just click Ok.



## Stopping the Action

- After you have completed all the steps for the action, you will have to stop recording the action. Click the "stop button," the small square at the bottom of the Action menu.



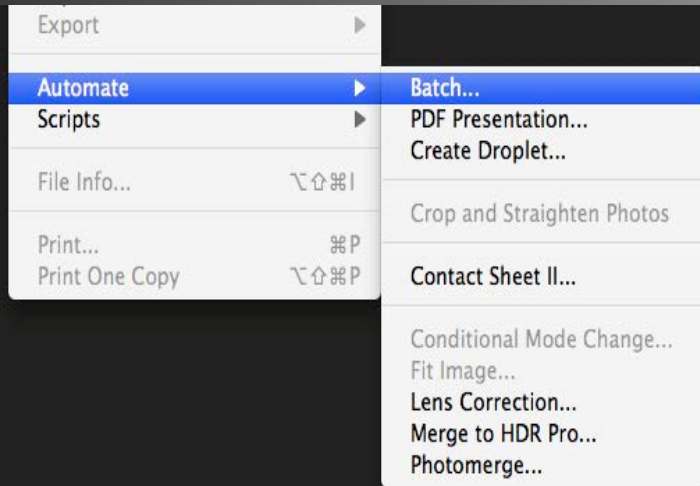




# Running an Action

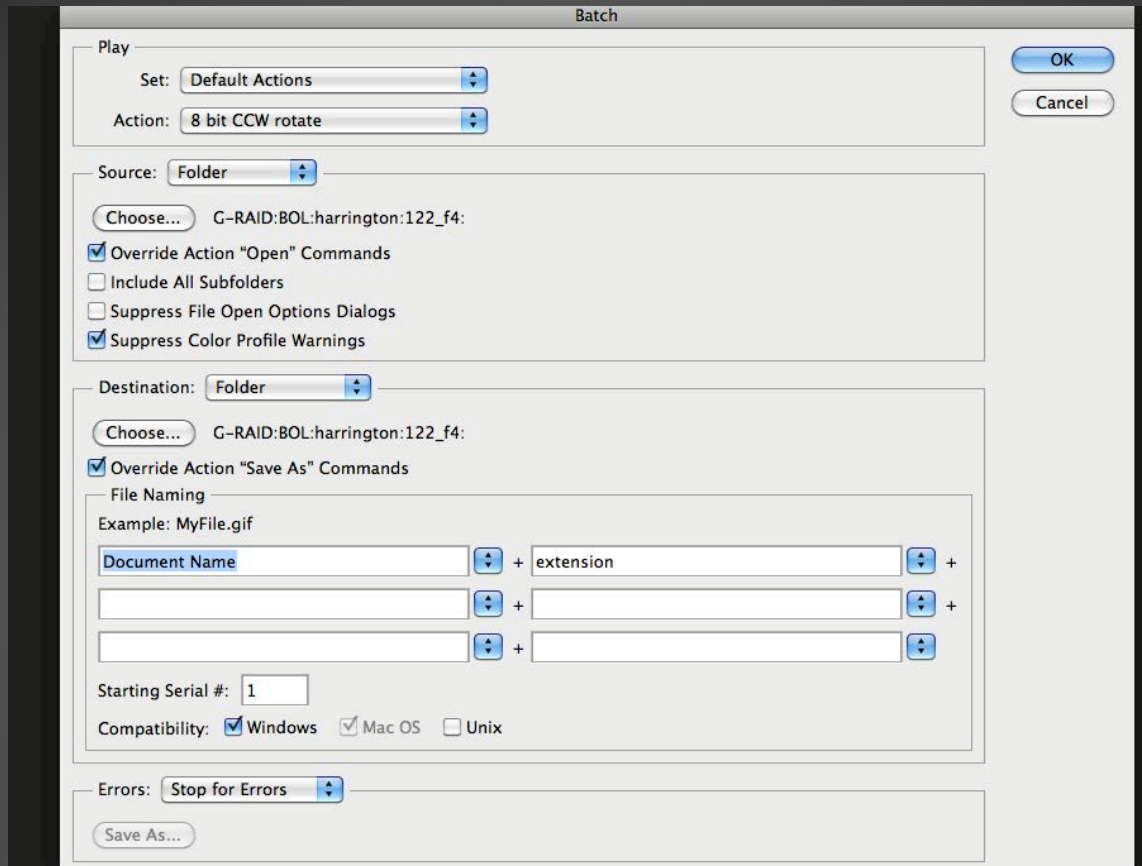
## Playing the Action

- Now that you have created the Action, you can “play” it so the steps are applied to the entire folder of images. Go to File → Automate → Batch



## Setting the Action

- A new window called “Batch” will pop up. In the “Play” section select the action you have just created. In the “Source” section click on Choose; locate the folder containing the files you want to run the batch action on; select it. In the Destination section (where the files will end up), do the same as the step before. The window should look like this:

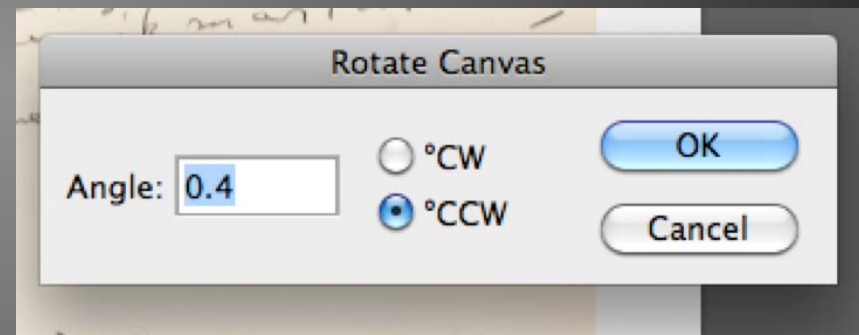
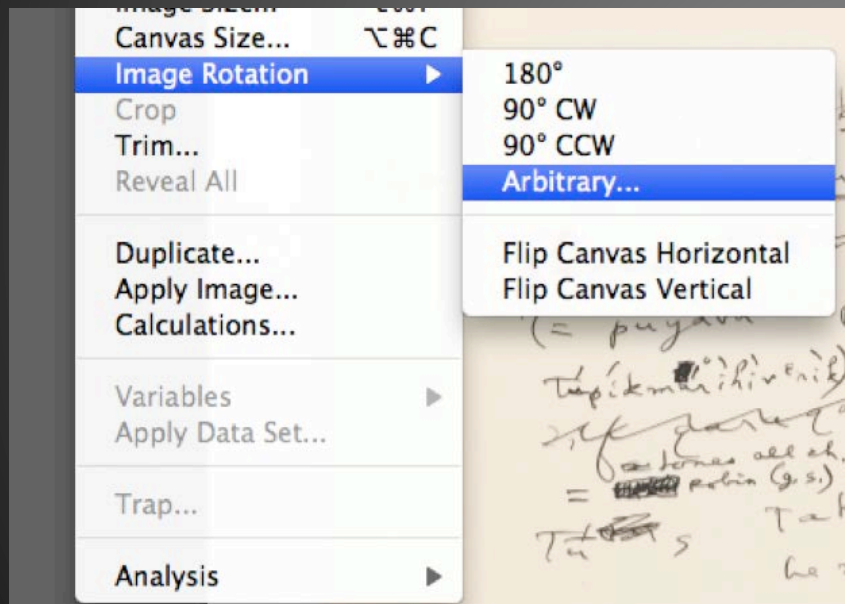


## Completing an Action

- Click "Ok." The batching will automatically begin and run through the files contained within the folder, which will take anywhere from a few seconds to a few minutes, depending on the amount and size of the files.

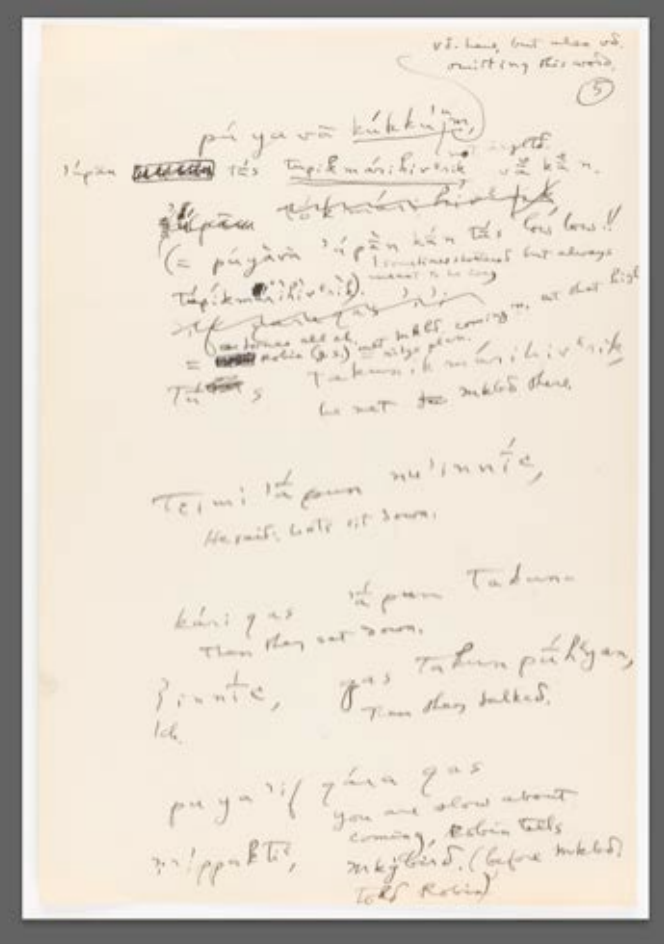
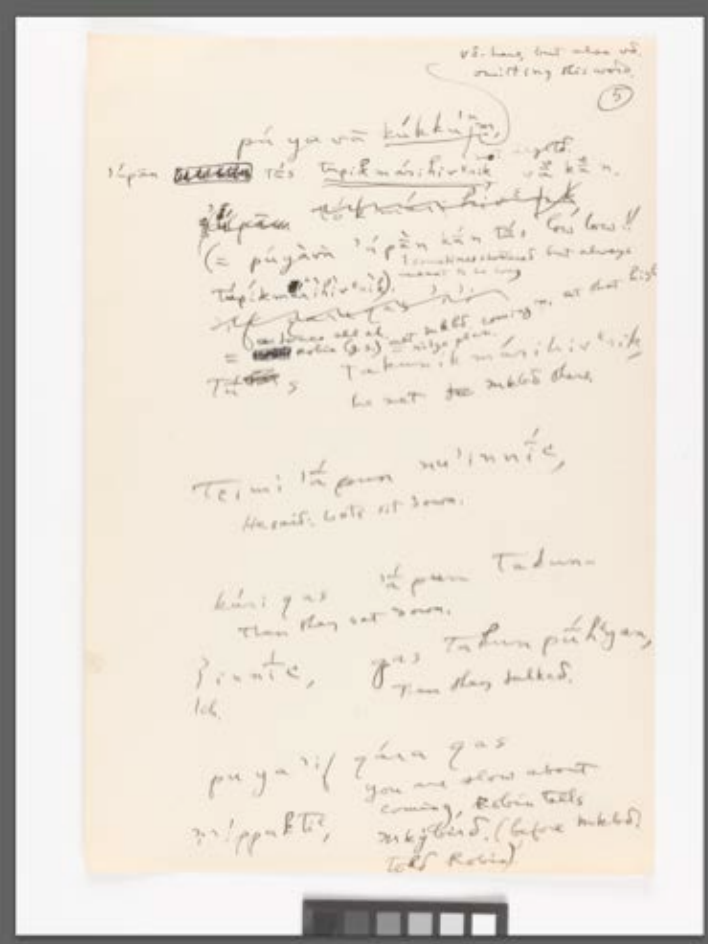


# De-skew images as needed





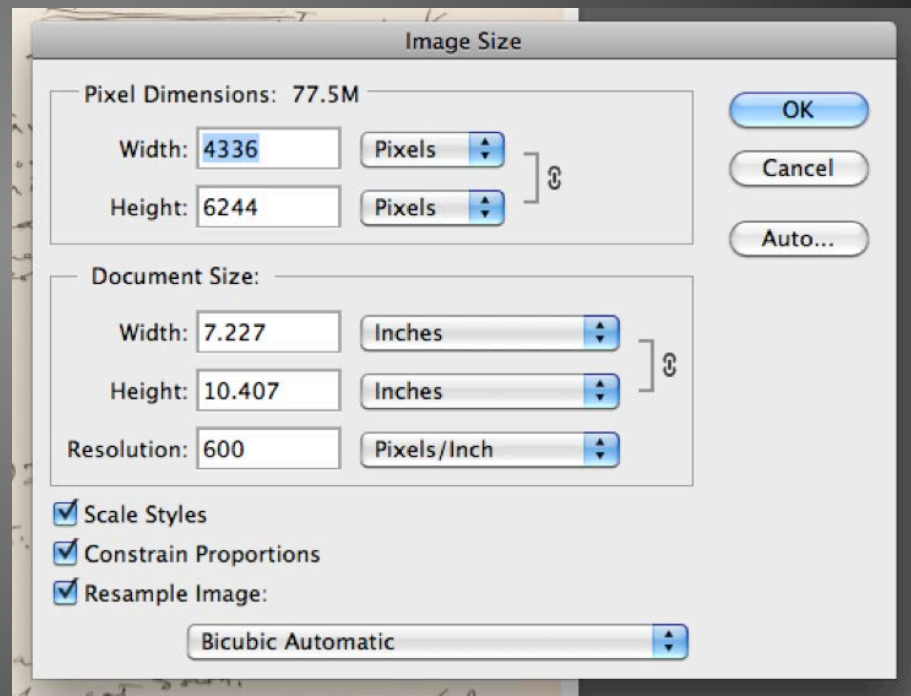
# Cropping Images





# Cropping Images

- Once you have cropped the first image, go to Image→Image Size and note the pixel dimensions. You will want to make a note of them so you can set the crop area to these proportions so all the images in a given collection will be the same size.





# Cropping Images

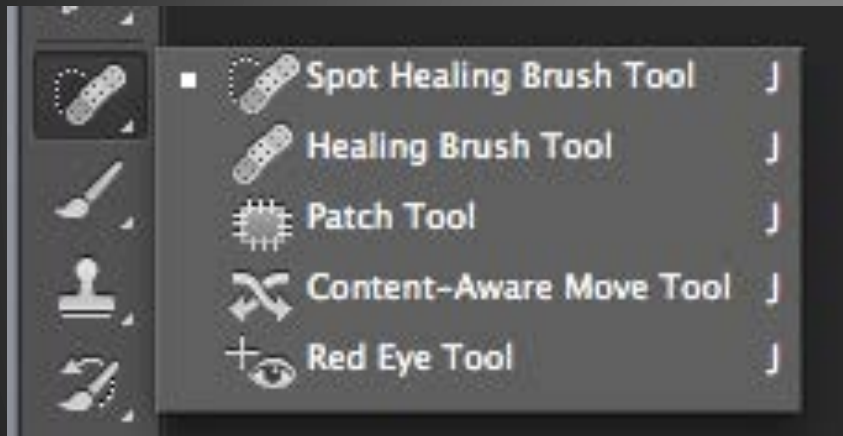
Instead of cropping it “freehand” you will set the crop area based on the pixel information from the previous image.

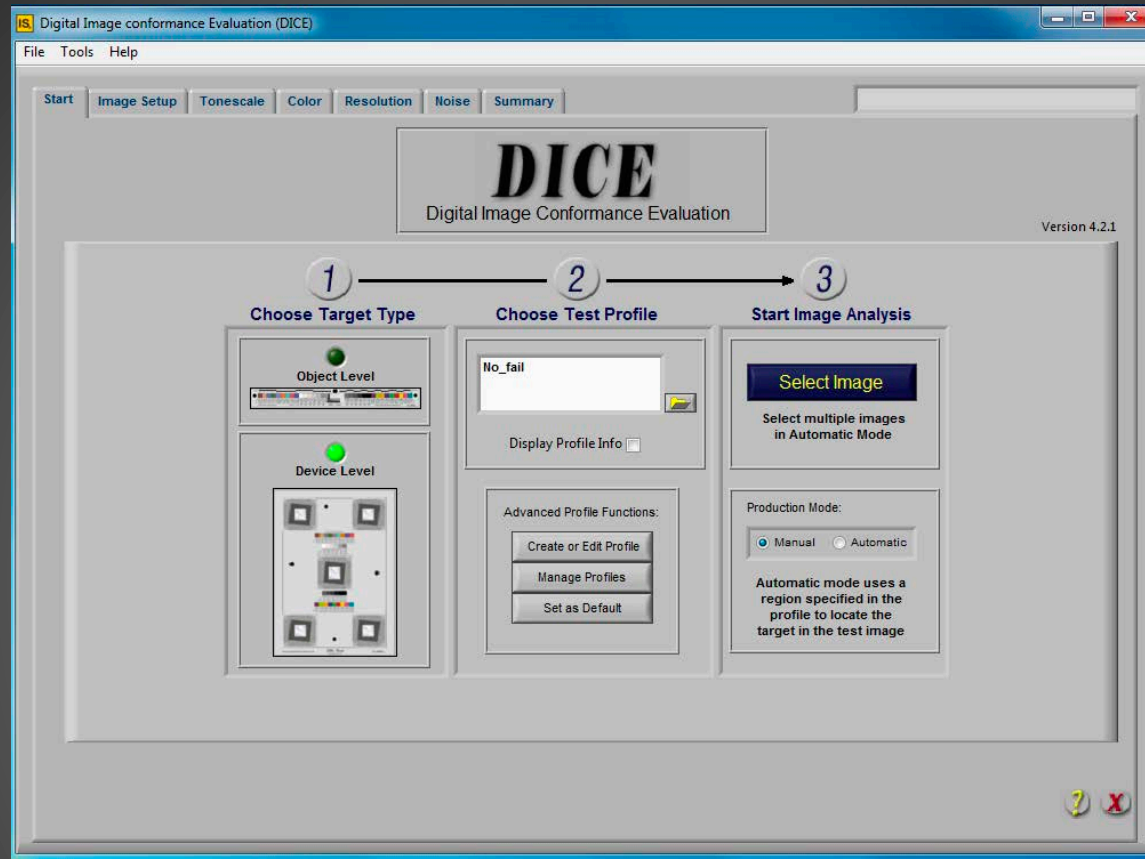


When you go to crop the image, you will notice that the crop area is constrained to the proportions you've set. Move the crop area around so it fits snugly around the image, while still leaving some background. Once you have cropped it to the correct size, hit enter. The image will be cropped. Click File→Save and File→Close.



# Inspection of 10% of images at 100% magnification

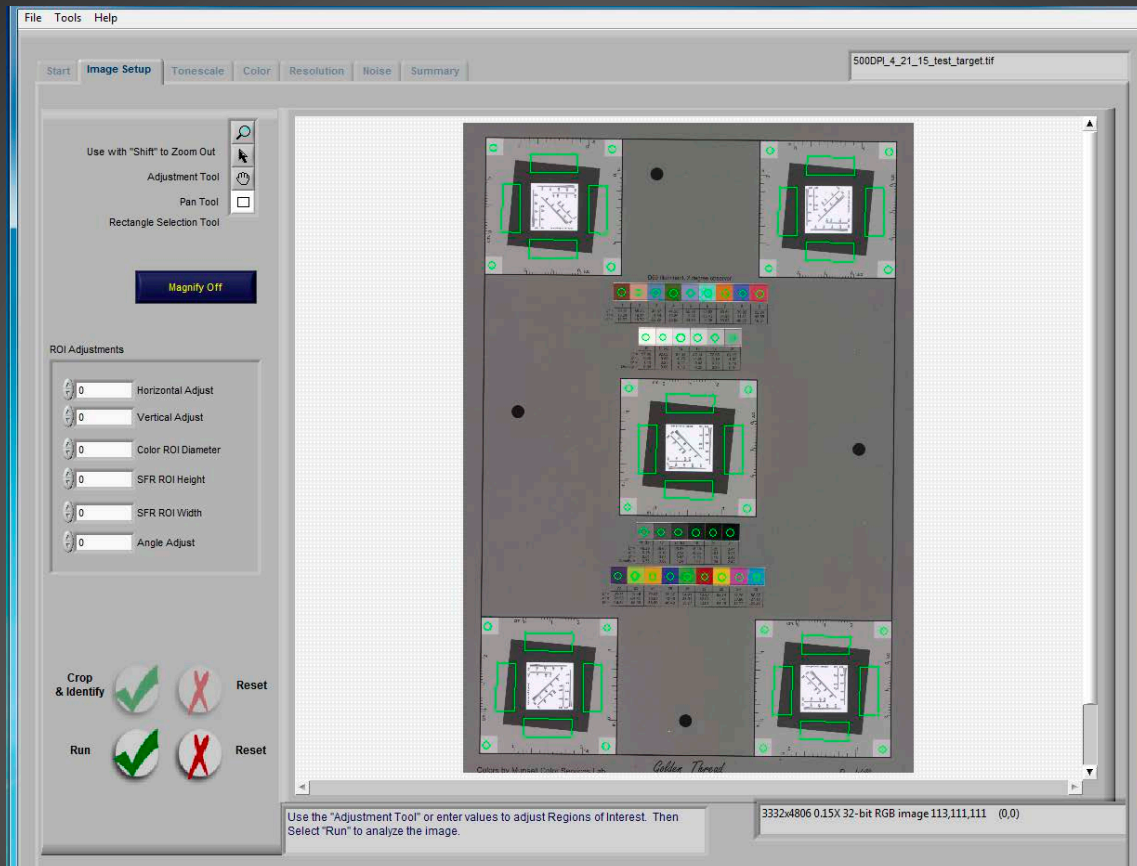




# Advanced Quality Control

DICE (Digital Image Conformance Evaluation)





# Advanced Quality Control

## DICE (Digital Image Conformance Evaluation)

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# Advanced Quality Control: GoldenTouch Auto-Cropping Software

**Parameters**

Source Directory: T:\arcadia\in\_process\leach\leach\_addtnl\_trobriand\_mat\source\

Result Directory: T:\arcadia\in\_process\leach\leach\_addtnl\_trobriand\_mat\examples\

Failure Directory: T:\arcadia\in\_process\leach\leach\_addtnl\_trobriand\_mat\bad\

TIFF

JPG

Gamma: 0   Match Green

Gain: 0

Offset: 0

Load Densities:

Load Intensities:

Calibration Mode - Calibration Object Type

Per Image 9"  Per batch

Per Image 18"  No Target

Per Image 4.5"

Calib Image:

Non-stop Batch

16 Bit Input

16 input / 8 output

16 input / 16 output

Target Position

Top

Left  Right

Bottom

Crop Output

Remove Target  Black Background

Crop Background  White Background

Threshold: 200

Load Params Save Params

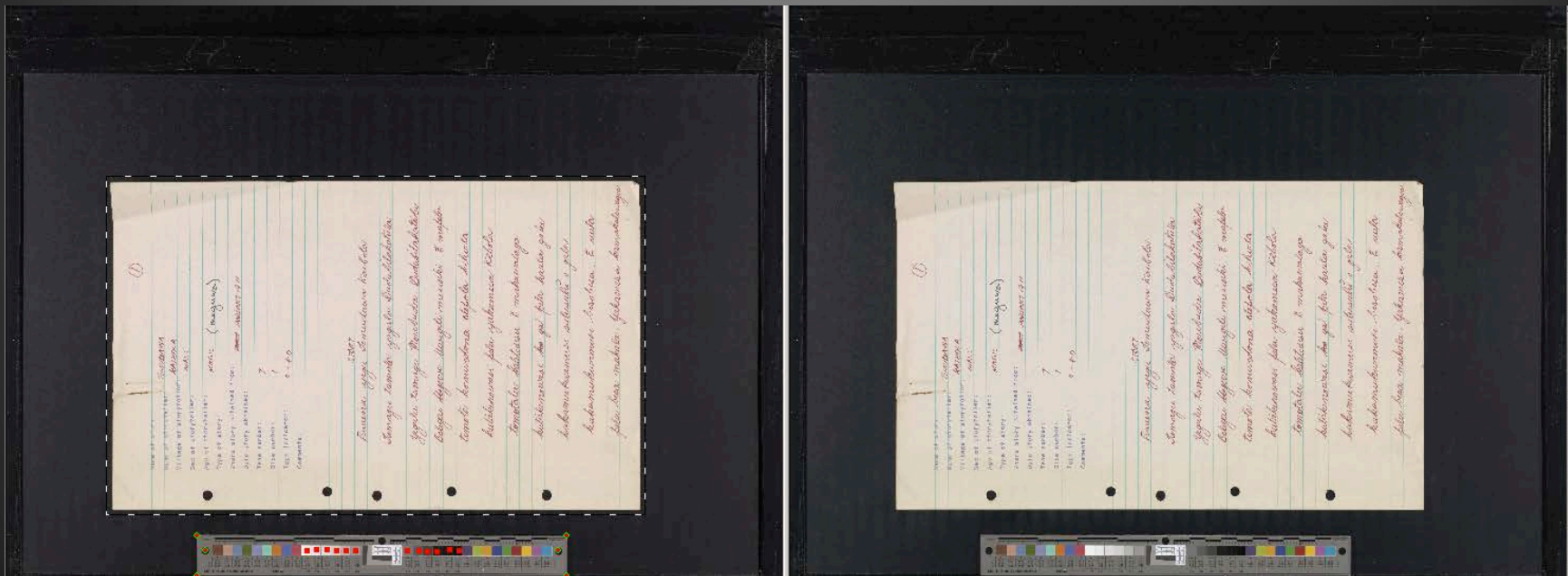
OK Start Batch OK No Start Cancel

```

defaultparams.cs - Notepad
File Edit Format View Help
// Batch Parameters
Source Directory      : T:\arcadia\in_process\leach\leach_addtnl_trobriand_mat\source\
Destination Directory : T:\arcadia\in_process\leach\leach_addtnl_trobriand_mat\examples\
Failure Directory     : T:\arcadia\in_process\leach\leach_addtnl_trobriand_mat\bad\
// Allowed file format: 0 - no, 1 - yes
TIFF Output          : 1
JPG Output           : 1
// Output file name format: "" - original name, "gamma" - gamma added to name
Output File Name     :
// Create output text files (LUT,gamma etc.): 0 - no, 1 - yes
Output Text Files    : 0
// Enabled 16bit depth input, processing and output: 0 - no, 1 - yes
16bit Input         : 0
Non-stop mode       : 1
Enabled batch non-stop (no interaction) processing mode: 0 - no, 1 - yes
Parameter File      : T:\arcadia\in_process\leach\leach_addtnl_trobriand_mat\bad\defaultparams.txt
// Tonal Curve Parameters: gamma, gain, offset
Gamma               : 0.000000
Gain                : 0.000000
Offset              : 0.000000
// Match tonal curves to green curve: 0 - no, 1 - yes
Match green curve   : 1
// ASCII file with target densities
Densities File      :
// ASCII file with target code values (intensity values)
Intensities File    :
// Target Parameters
// Target type used: object - object-level 9 inch, object18 - object-level 18 inch, object4.5 - object-level 4.5 inch, Device - device-level, none = no object
Target Type         : object18
// If device-level target, image with device-level target
Calib Image         :
// Search target in position
Right Target Position : 1
Left Target Position  : 1
Top Target Position   : 1
Bottom Target Position : 1
// Post Processing
Crop Target          : 1
Crop Background      : 1
Crop White           : 0
Crop Threshold       : 200
Crop Margin          : 68
Crop Shadow          : 2.500000
// QA Parameters
QA Neutrality        : 4.000000
QA Aim               : 5.000000
// System Parameters
Win DPI              : 300.000000
Max DPI              : 600.000000
  
```

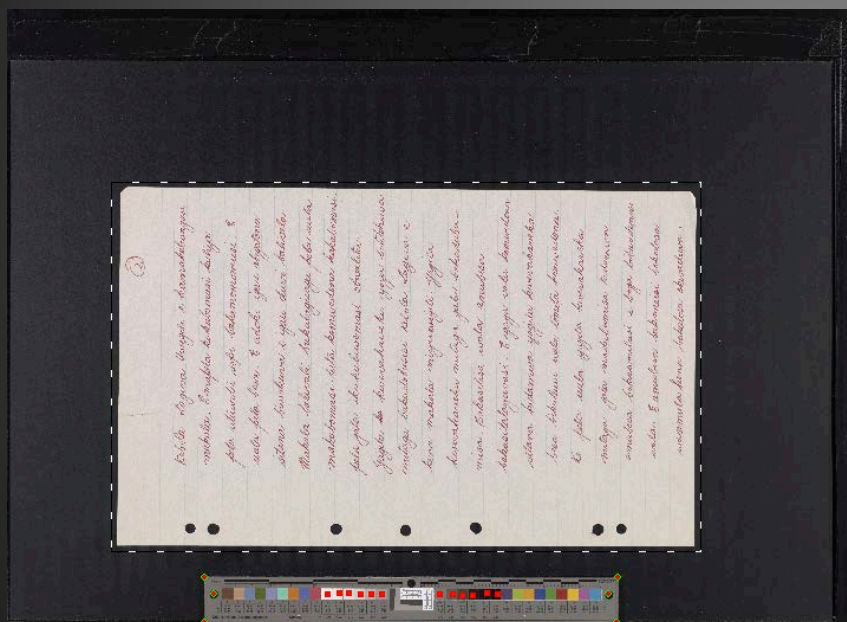


# Advanced Quality Control: GoldenTouch Auto-Cropping Software





# Advanced Quality Control: GoldenTouch Auto-Cropping Software





# Quality Control Checklist

Kenney, Anne R. and Oya Y. Reiger. "Establishing a Quality Control Program," Chapter 4, pp.61-73. (2000).  
Moving Theory into Practice: Digital Imaging for Libraries and Archives. Mountain View, CA: Research  
Libraries Group.

## Prerequisites for QC

- Identify products and goals
- Agree on Standards
- Determine a Reference Point
- Understand the Limitations of Current Knowledge, Practice and Technology

## Setting up your QC Program

- Identify Scope
- Determine Methods
- Evaluate System Performance
- Codify Inspection Procedures
- Control QC Environment
  - Hardware configuration
  - Image-display software
  - Monitor setup
  - Color quality control
  - Color management
  - Viewing conditions
  - Human characteristics

## Assessing Image Quality

- Evaluate:
  - Resolution
  - Color and Tone
  - Overall appearance



Smithsonian  
National Museum of Natural History

# Digital Asset Management, Storage!

OPEN TEXT

OPEN TEXT  
Media Management™

Log in to Media Manager ? ⚙

User name  
|

Password

Remember me

Forgot your password?

V 7.2.1.5-5

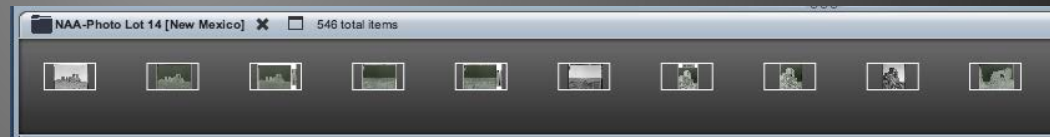
*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# Strength of Storage

- What are your storage needs?
  - Throughput
- What are your storage capabilities?
  - Staff
  - Budget
- What is the level of IT support?
  - Infrastructure

```
----- HOT FOLDER IMPORT SUMMARY -----  
Total files to import : 1114  
Number of files imported : 1114  
Number of files not imported: 0  
Number of warnings : 0  
-----
```





Smithsonian  
National Museum of Natural History

# STORAGE WARNING

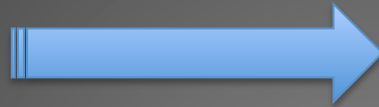
- DIGITAL FILES NEED TO BE BACKED UP IN MULTIPLE LOCATIONS
- DIGITAL FILES NEED TO BE MIGRATED
- DIGITAL FILES NEED SPACE TO LIVE





# Storage Options

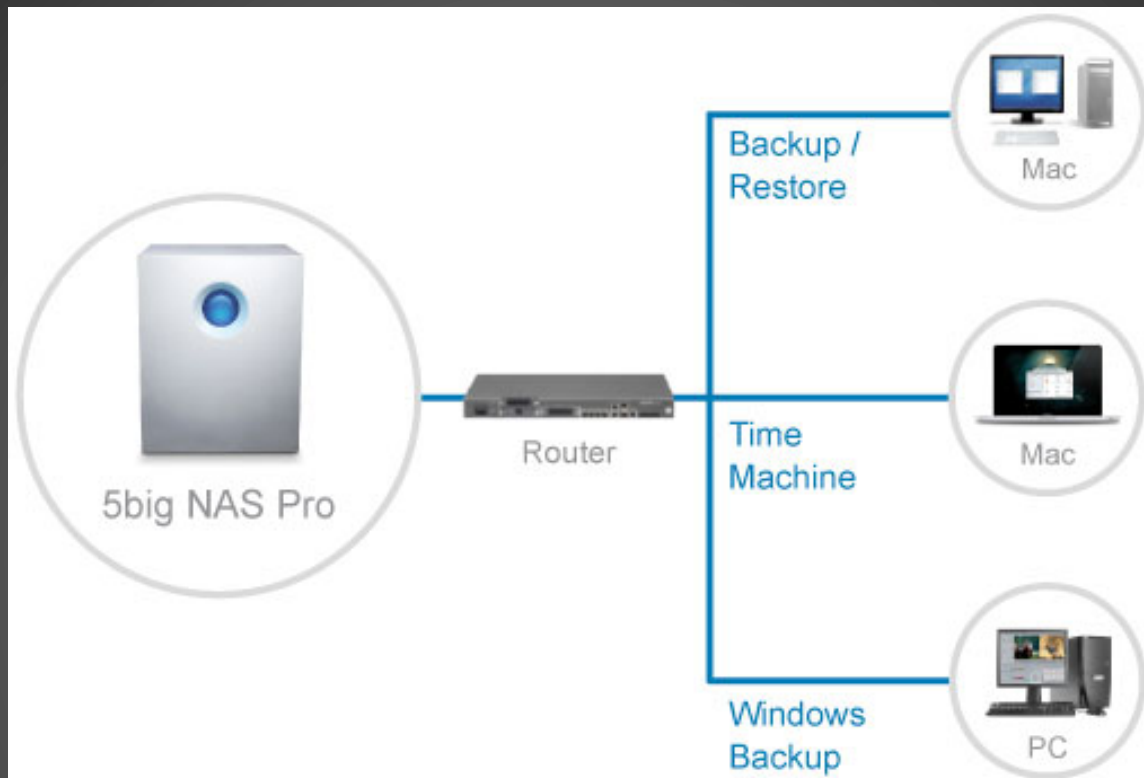
## Good, Better



- Cloud-based Storage
- Network Attached Storage (NAS)
- Locally managed servers
- Local external storage
  - RAID Hard Drives

## Best

- Trusted Digital Repositories
- Digital Asset Management Systems
  - Desktop
  - Workgroup
  - Enterprise



## Local Storage

Requires constant maintenance and upgrading

Hard drives fail!!

Migration and offsite backup



# Cloud Storage

- Institutional
  - DuraCloud
  - Chronopolis
  - Digital Preservation Network
- Commercial
  - Amazon
  - Mozy
  - Google Drive
  - iCloud





# DSpace

- Free, open-source software to build open digital repositories



**DSpace**  
Open Source "turn-key" institutional repository application  
Brought to you by: aschweer, benbosman, bluyten, bollini, and 19 others

Summary | Files | Reviews | Support | Mailing Lists | External Link ▾

★ 5.0 Stars (43)  
↓ 952 Downloads (This Week)  
📅 Last Update: 2015-03-16

[Download](#)  
dspace-5.1-src-release.zip

[Tweet](#) 3 | [g+](#) 4 | [Like](#) 11

[Like](#)

[Browse All Files](#)

**Description**

Open Source Digital Asset Management system that enables services for access, provision, stewardship and re-use of digital assets with a focus on educational and research materials

[DSpace Web Site](#)

[Follow @dspacetweets](#)

**Categories**  
Archiving, Digital preservation, Research

**License**  
BSD License



# Trusted Digital Repositories

- “A trusted digital repository is one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future.” (Trusted Digital Repositories: Attributes and Responsibilities An RLG-OCLC Report)
- Digital Preservation Coalition
- HathiTrust (libraries)





# DAM useful

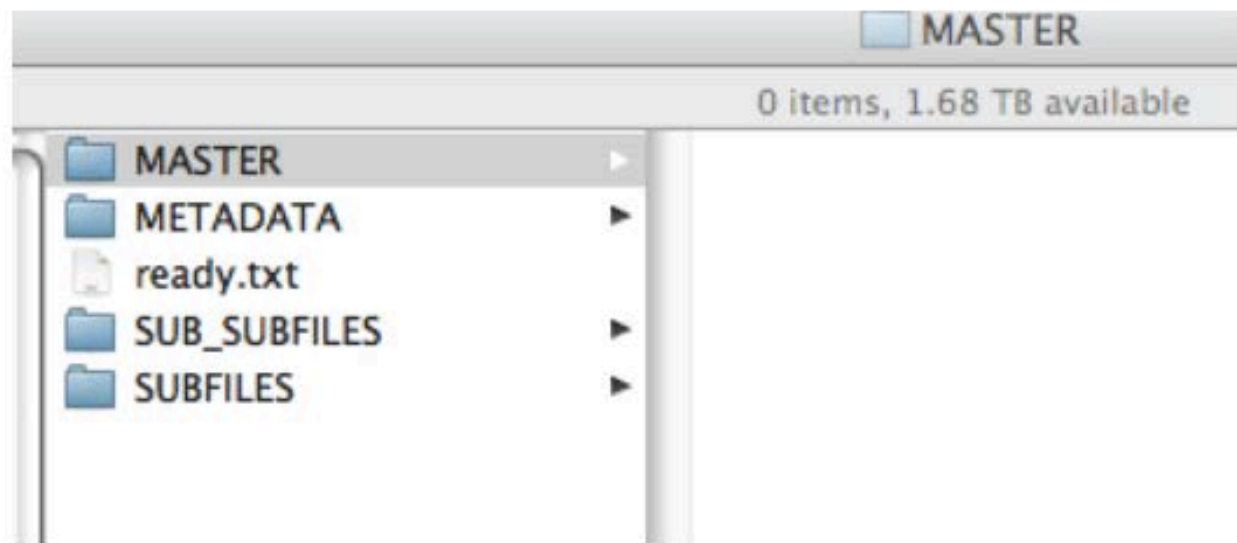
- *“Instituting digital asset management (DAM) in cultural heritage institutions tends to be a major IT initiative. What is often overlooked is that DAM in isolation merely provides a resource intensive organizational tool for digital assets within the institution. The more important aspects of a DAM implementation are the workflow processes and procedures that are integrated into the application, the links the application makes internally and externally to other institutional systems, and ultimately how the implementation changes and enhances the institution’s business processes surrounding the use of digital assets.”*
  - Howard Goldstein (Center for Digital Imaging, Inc.) and Rob Hendriks (The Rijksmuseum, Amsterdam)



# DAMS workflow - 1

## B. Upload digital assets to the DAMS through the Hot Folder.

1. Sign into the Hot Folder [Go → Connect to Server]. You will see the following folder structure:



2. You will drag files directly into the Hotfolders. Where you copy the files will depend on what you have:



# DAMS workflow - 2

- **NEGATIVES:**
- XX\_neg folder-> MASTER
- XX\_post folder-> SUBFILES
- XX\_crop folder-> SUB-SUBFILES
- XX\_jpg folder-> do not ingest
- **FILES WITH CROPS AND VERSOS**
- XX.tif file -> MASTER
- XX\_crops folder and/ or Versos -> SUBFILES
- XX Verso Crops (if they have been renamed) -> SUB-SUBFILES
- **FILES WITH CROPS AND VERSOS** [Alternative, usually Photo Lots]
- XX.tif file (recto) -> MASTER
- XX.tif file (verso) -> MASTER
- XX crop (recto) -> SUBFILE
- XX crop (verso) -> SUBFILE
- **FILES WITH CROPS AND NO VERSOS**
- XX.tif file -> MASTER
- XX\_crops folder -> SUBFILES
- **FILES WITH VERSOS AND NO CROPS**
- XX.tif file -> MASTER
- XX Verso Crops -> SUBFILES







# DAMS workflow - 3

- Confirm ingest
- Track ingest

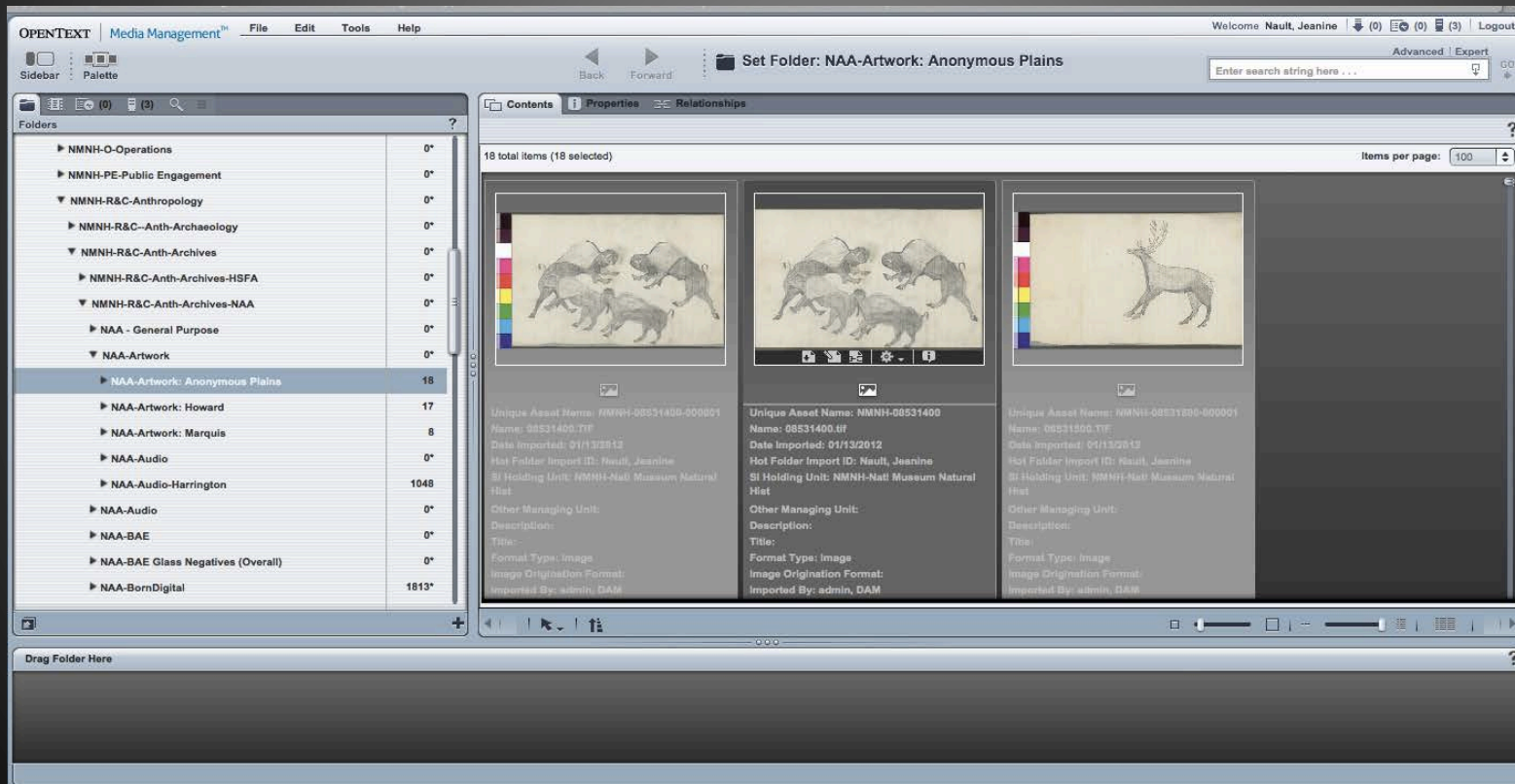
SI DAMS Hot Folder Import - Report for Metadata Profile  
no-reply@dams.si.edu

	Collection	Added - by	Number of Assets Ingested	Number of Assets Verified	Search Method in DAMS	Folder Created	Origination Format	Original Content Format	DAMS Level Metadata Applied	Notes
134	MS 4365-d	8/19/14 JN	89+46+17+12+5+7+2+1+7+6 PDFs	192		x		text		when exporting, export PDFs for e folder, but create image viewer of ENTIRE MS
135	MS 4800 [23]	8/20/14 SV	36+1 PDF	37		x		artwork	x	Link to 4800 [Drawings] folder eventually
136	MS 4800 [105]	8/20/14 SV	18+1 PDF	19		x		artwork	x	Link to 4800 [Drawings] folder eventually
137	MS 4800 [190]	8/20/14 SV	70+1 PDF	71		x		artwork	x	Link to 4800 [Drawings] folder eventually
138	MS 4800 [191]	8/20/14 SV	9+1 PDF	10		x		artwork	x	Link to 4800 [Drawings] folder eventually
139	MS 2790	8/27/14 JN	45+1 PDF	46		x		Text	x	
140	MS 2829	8/27/14 JN	12+1 PDF	13		x		Text	x	
141	MS 2994	8/27/14 JN	61+1 PDF	62		x		Text	x	
142	MS 2999	8/27/14 JN	211+1 PDF	212		x		Text	x	
143	MS 1791	9/8/14 JN	108+1 PDF	109		x		text	x	
144	MS 1825	9/8/14 JN	8+1PDF	9		x		text	x	
145	MS 2792	9/8/14 JN	118+1 PDF	119		x		text	x	
146	MS 3219	9/8/14 JN	1+62+47+2 PDFs	112		x		text	x	
147	MS 3220	9/8/14 JN	98+1 PDF	99		x		text	x	
148	MS 3342	9/8/14 JN	69+1 PDF	70		x		text	x	
149	MS 3343	9/8/14 JN	18+1 PDF	19		x		text	x	



# DAMS workflow - 4

- Search and verify assets





# DAMS workflow - 5

- Apply DAMS level metadata

OPENTEXT | Media Management™ | File Edit Tools Help | Welcome Nault, Jeanine | (0) (0) (0) | Logout

**Edit Properties : p114\_nm\_129.tif**

Metadata Security Categories

Type: SI Core Image

Administrative

SI Holding Unit: NMNH-Natl Museum Natural Hist

SI Department: NMNH-DoA-National Anthro Archives-33202X

Administrative Content Type: SD 600

For Public Use: No

Save and Close Close Without Saving Data Metadata Editor

- Apply DAMS security policies

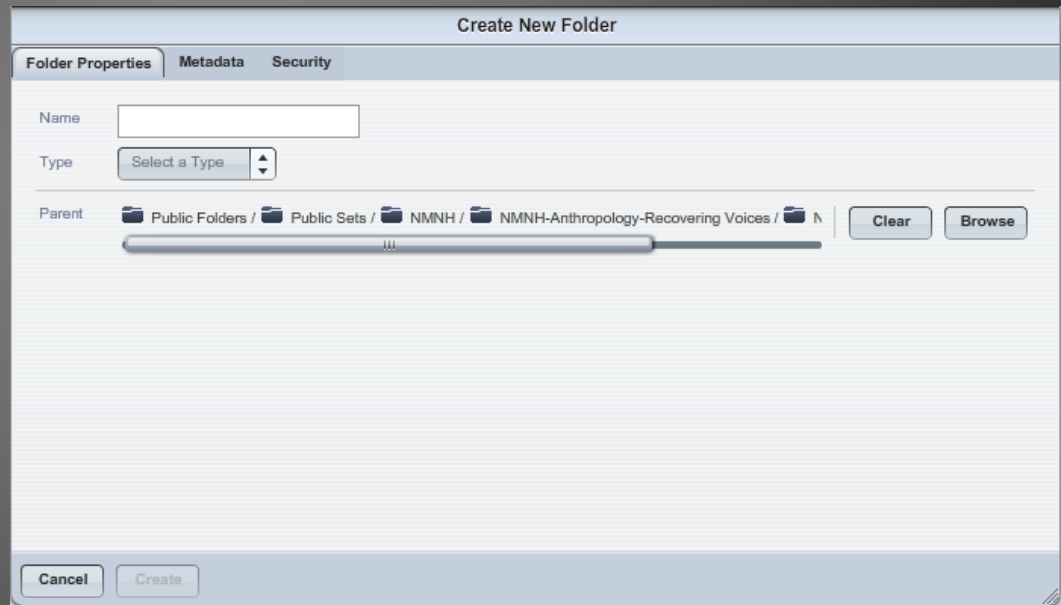
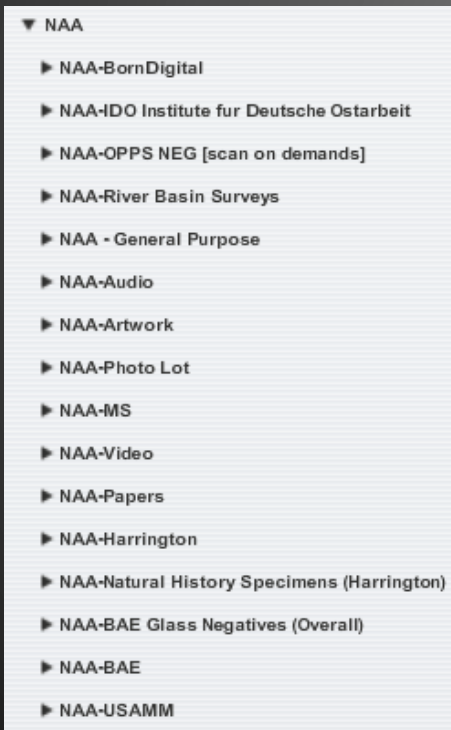
NMNH-R&C-DoA-Natl Anthro Archives SecPol  
NMNH-R&C-DoA-Natl Anthro Archives policy to allow all perms for NAA assets to NAA users but only view thumbnail & summary data to all SI-DAMS users.

Metadata Security Categories



# DAMS workflow – 6

- Create a Folder with Collection Name and Number





# Storage Conclusions

- Data needs to be backed up in multiple locations, migrated, and takes up space
- Decide which storage and/or digital asset management system is right for you
  - infrastructure, support, staff, time, money and NEEDS



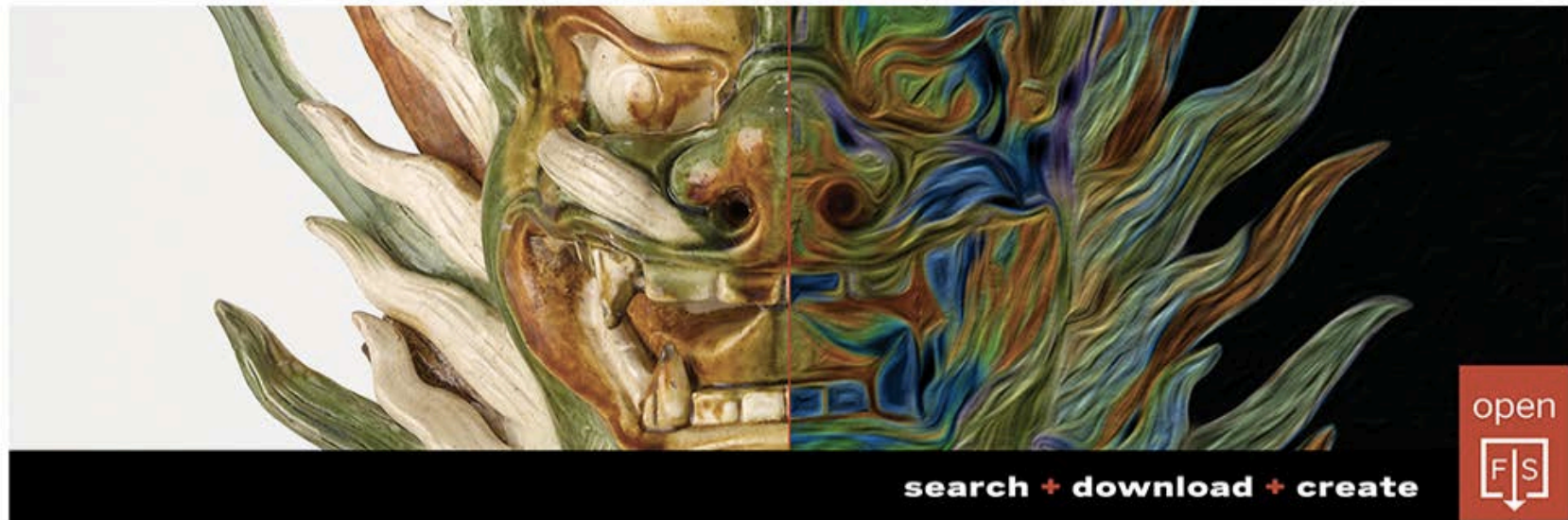
Smithsonian  
National Museum of Natural History

# All About Access

PLAN YOUR VISIT   EXHIBITIONS   EVENTS   **COLLECTIONS**   EXPLORE + LEARN   RESEARCH   SUPPORT F|S   CONNECT

[Home](#) > [Collections](#) > [Open F|S](#)

## Open F|S



Welcome to Phase 1 of Open F|S, the complete digitized collections of the Freer and Sackler Galleries and the Freer Study Collection. With more than 40,000 works being made available for high-resolution download—expanding regularly with our [new acquisitions](#)—you can explore the Smithsonian's museums of Asian art from anywhere in the world, whenever you like. Images can be used for all [non-commercial purposes](#), from [desktop wallpapers](#) to artistic gifts for family and friends. Send

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# Allowing Access

## Why?

- Mission driven
- Reach newer, bigger audiences
- Expectation of public
- Feasibility
- Relevance
- Goodwill relationship with source communities
- Reduce handling, increase access to fragile collections

## How?

- Who are your users?
- How will they access collections?
- How will you provide access?
- Open access, freely available? Or fee for service?
- Rights and Restrictions
- Copyright issues
- Staff, budget, time



Smithsonian  
National Museum of Natural History

# Open Access



[The Getty Open Content Program](#)

[Digita Vaticana](#)



**VERBA VOLANT, SCRIPTA MANENT, DIGITA SUNT.**

Converting 40 million pages into more than 45 quadrillion bytes

Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post





# Increasing Access

## [John P. Harrington web portal](#)

- Digital access to various collections
  - Manuscript and Microfilm
  - Photographs
  - Sound Recordings
  - Botanical specimens




Smithsonian  
National Museum of Natural History

# Increasing Access – Harrington web portal

**Natural History Specimens by Specific Culture**



[Chumash Natural History Specimens](#)



[Salinan Natural History Specimens](#)



*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*



# Increasing Access – transcription.si.edu

Smithsonian  
HOME PROJECTS ABOUT TIPS

SMITHSONIAN DIGITAL VOLUNTEERS: TRANSCRIPTION CENTER SIGNUP LOGIN

FEATURED  
ALL PROJECTS  
VIEW PROJECTS

ink on it. then again, painted all with  
Bainbridge white, no medium and added ultramar  
The shadows and had to paint much into it  
with the colors, to which linseed and a little l  
umber varnish was added. All this is rusatic  
tory, since no real white and brilliancy is  
ossible. The Cadmium yellows are even much  
more luminous than the white and in the house  
the doors I had to add much oil is sides an

JOIN US!  
LEARN HOW TO TRANSCRIBE  
Become a Smithsonian Digital Volunteer  
and unlock our stories as you transcribe  
our collections.

BROWSE PROJECTS  
Select a category below to begin browsing  
projects.  
Select a Category

HOME > PROJECTS > NATIONAL ANTHROPOLOGICAL ARCHIVES > COPIES BY SWAN OF DRAWINGS BY HAIDA INDIANS OF MYTHOLOGICAL ANIMALS, SOME DATED 1873  
VIEWING PAGE 1 OF 10

Oolala 2  
[figure]  
This is the mythological being of the belief of the Haidah Indians of Queen  
Charlotte Islands, half man and half bird supposed by them to live on the  
mountains and to live on whales, or Indians, a Skookum and Evil Spirit. It is  
similar to the [Thenkloots?] or [Bockerally?] or the Makah.  
Copied from a drawing made by Geneskelos, brother of Kitkun of the Carvers and  
Tattoers of the Haidah tribe, May 1873.

colala 2

This is a mythological being of the  
belief of the Haidah Indians of Queen  
Charlotte Islands, half man and half bird  
supposed by them to live on the mountains and  
to live on whales, or Indians, a Skookum and  
Evil Spirit. It is similar to the Thenkloots  
or Makah, or the Makah.  
Copied from a drawing made by  
Geneskelos, brother of Kitkun, one of the  
Carvers and Tattoers of the Haidah tribe  
May 1873.

feedback

Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post



# Increasing Access – Arcadia project

Archives, Manuscripts, Photographs Catalog  
Smithsonian Institution Research Information System (SIRIS)

Search Search Images About

Keyword Browse Finding Aids Combined Browse Images Search History All Catalogs

Search: General Keyword NAA 99.2707

You are only searching Archives, Manuscripts and Photographs

More Smithsonian Searches

Who else has...  
 • Michelson, Truman  
 • Arapaho Indians  
 • Indians of North America  
 • Rites and ceremonies  
 • Psychology  
 • Games  
 • Arapaho language  
 • Language and languages – Documentation  
 • Linguistics  
 • Ethnology  
 • Friday  
 • Fitzpatrick, Truman  
 • Field notes  
 • Vocabulary

**Arapaho ethnological and linguistic notes collected by Truman Michelson, 1910**

Creator: Michelson, Truman, 1879-1938  
 Title: Arapaho ethnological and linguistic notes collected by Truman Michelson, 1910  
 Contained in: NAA99.ms.2707.1350a:1990a.1000a.kst01  
 Phys. Description: 98 pages  
 Click to view image set:

Digital Reference:

Additional forms:  
 Summary: Digital images are available online.  
 Language Note: In Arapaho and English.  
 Cite as: Manuscript 9997, National Anthropological Archives, Smithsonian Institution  
 Local Notes: Title changed from "Miscellaneous ethnological notes, linguistic notes, vocabulary Appendix 1910; certain pages dated '1910' 4/15/2014.

Four Digital Content (PDF format) buttons are displayed.

Archives, Manuscripts, Photographs Catalog  
Smithsonian Institution Research Information System (SIRIS)

Search Search Images About

Keyword Browse Finding Aids Combined Browse Images Search History All Catalogs

Search: General Keyword NAA 99.3

You are only searching Archives, Manuscripts and Photographs

More Smithsonian Searches

Who else has...  
 • Gaines, DeMark  
 • Illinois Indians  
 • Indians of North America  
 • Language and languages – Documentation  
 • Illinois language  
 • Miami language (Ill. and Okla.)  
 • Sound recordings

**Kaskaskia vocabulary and recorded interview with Louis E. Myers 1991**

Creator: Gaines, DeMark, interviewer  
 Title: Kaskaskia vocabulary and recorded interview with Louis E. Myers 1991  
 Phys. Description: 1 sound cassette; 7 pages  
 Digital Reference:

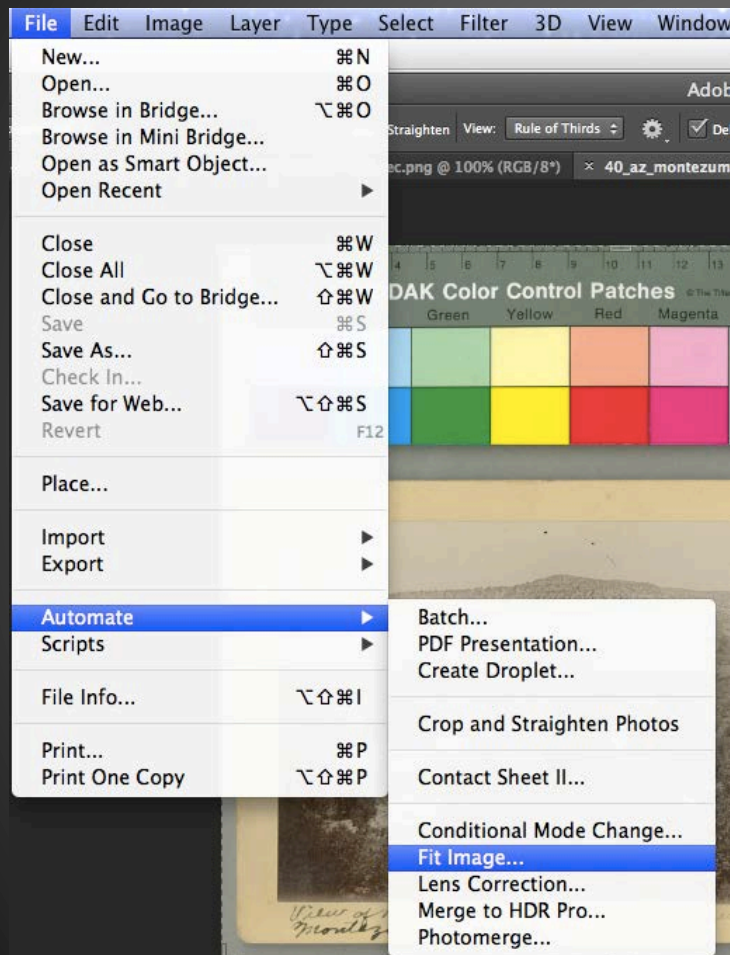
Online Sound (Sound Recording) Online Sound (Sound Clip)

Summary: Taped interview with Myers, then chief of the Peoria tribe, May 21, 1991. Interview conducted by DeMark Gaines, Director of the Northeastern Oklahoma A & M College Heritage Program and Orono Valley, Peoria Indian. Recording is accompanied by written Kaskaskia vocabulary list.  
 Language Note: In Myasme (Kaskaskia) and English  
 Cite as: Manuscript 1999-45, National Anthropological Archives, Smithsonian Institution  
 Funding Note: Digitization and preparation of these materials for online access has been funded through generous support from the Arcadia Fund.  
 Culture: Illinois Indians  
 Add: KW Sub: Miami-Myasme  
 Subject-Topic: LANGUAGE AND LINGUISTICS -- Documentation  
 Form / Genre: Sound recordings  
 Repository Loc: National Anthropological Archives, Smithsonian Museum Support Center, Suitland, Maryland  
 Local Number: NAA MS 1999-45

- Freely available, downloadable PDFs of 20,000 pages of endangered language manuscript material and 15,000 pages of associated text related to sound recordings
- Online mp3's of 3500 sound recordings



# How to create access – derivatives

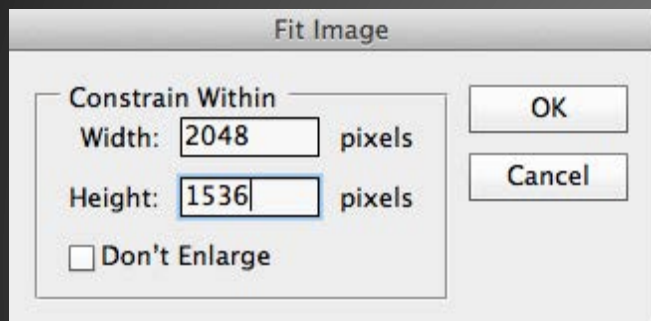


Go to File > Automate >  
Fit Image

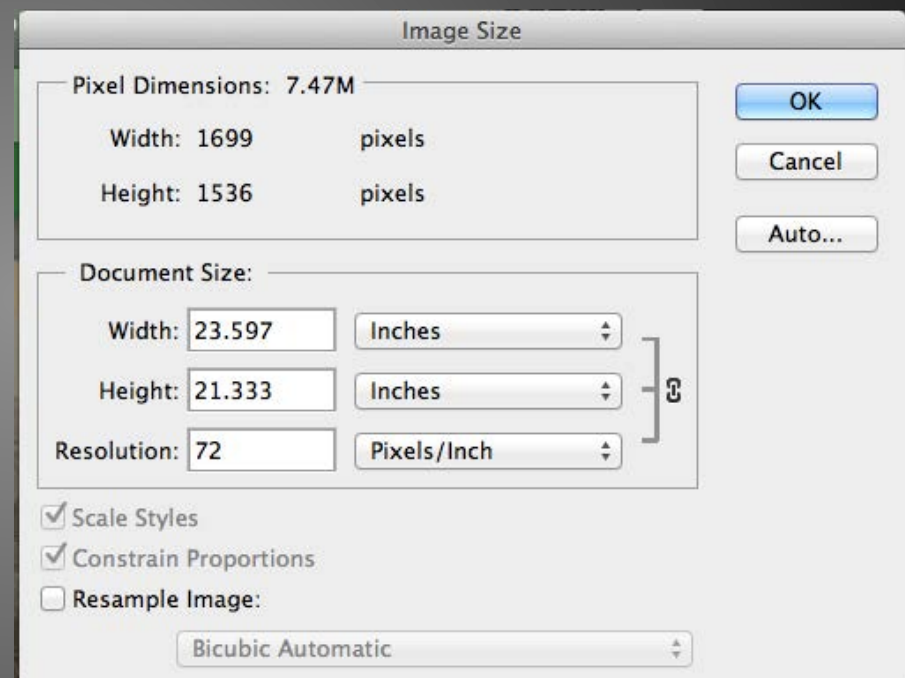


# How to create access - derivatives

A window will pop up – enter the follow dimensions into the appropriate boxes – click OK



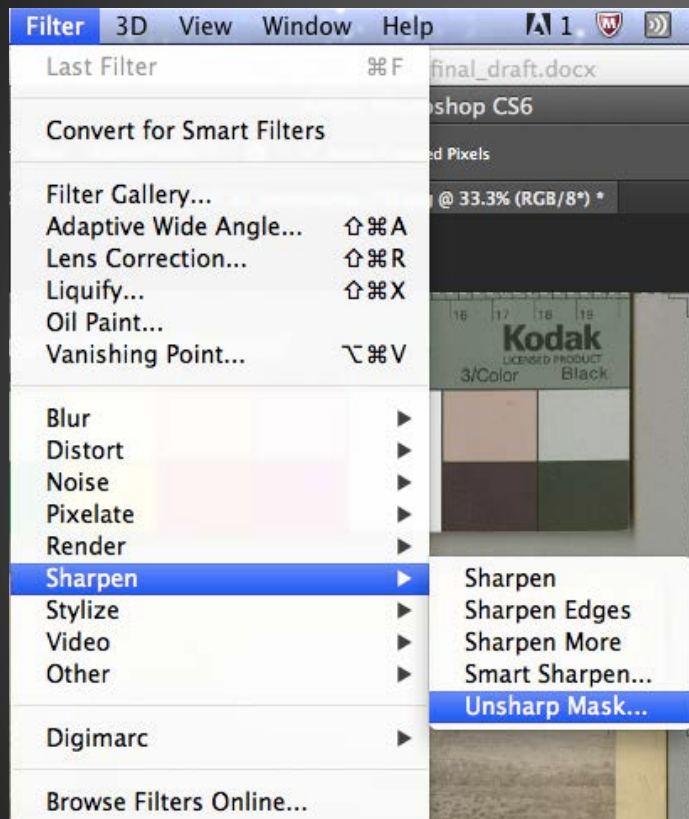
Go to Image > Image Size. An Image Size window will pop up.



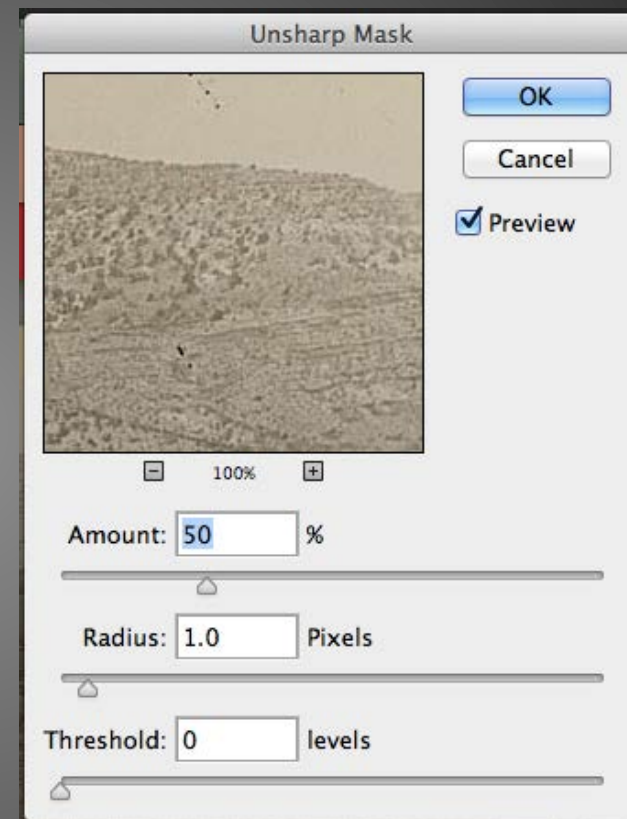
In the resolution field, enter 72. Make sure pixels/inch is selected. Make sure box for “Resample Image” is unchecked. Click ok.



# How to create access - derivatives



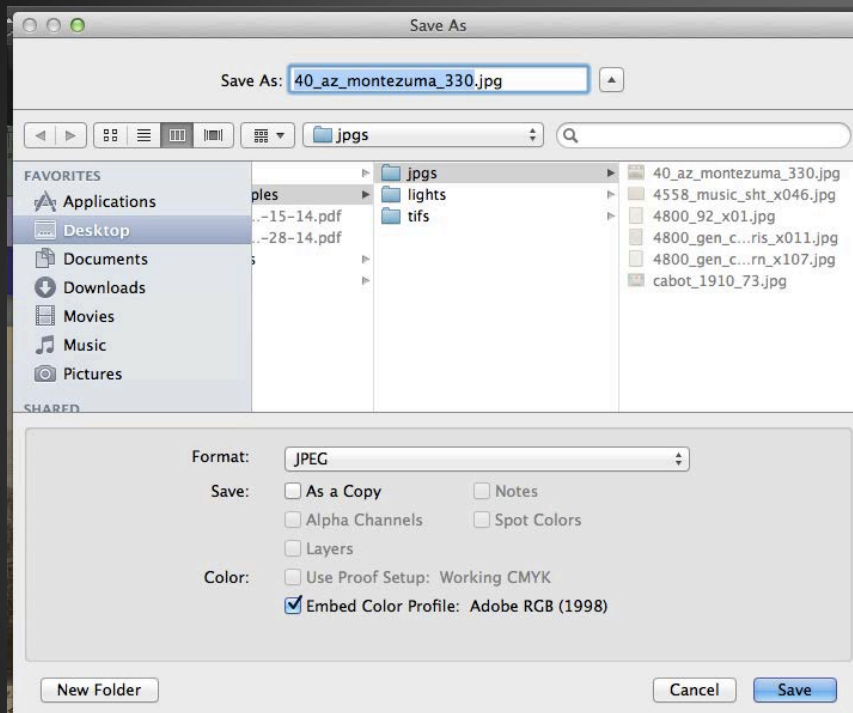
Go to Filter > Sharpen > Unsharp Mask



In the Unsharp Mask window enter the above in each field. Click Ok.

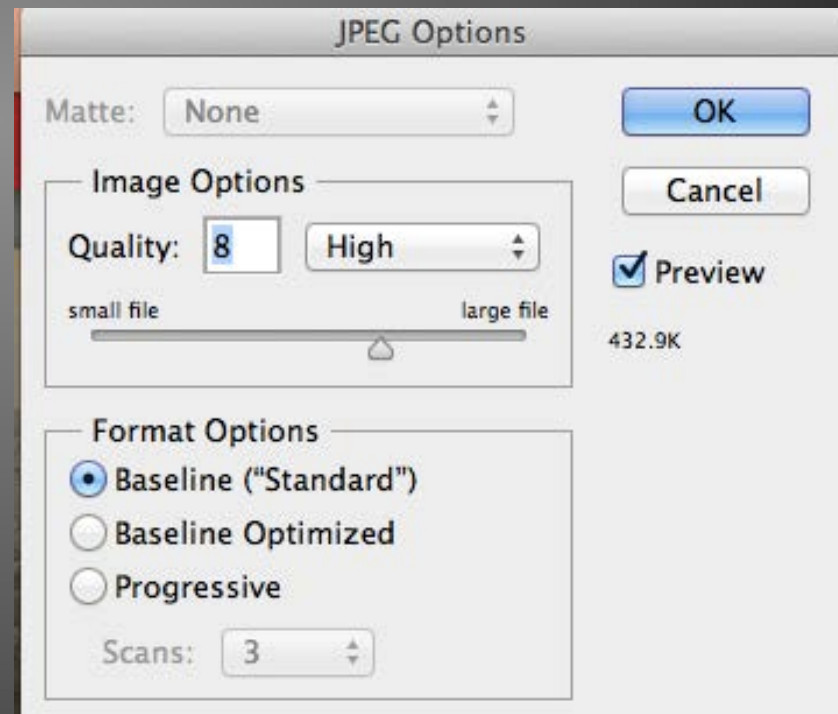


# How to create access - derivatives



Go to File > Save As. In the Save As window, make sure the format is “jpg,” and the box for “Embed Color Profile” is checked. Click Save.

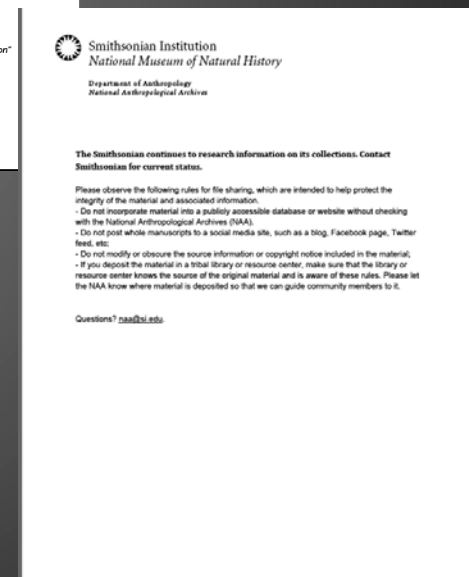
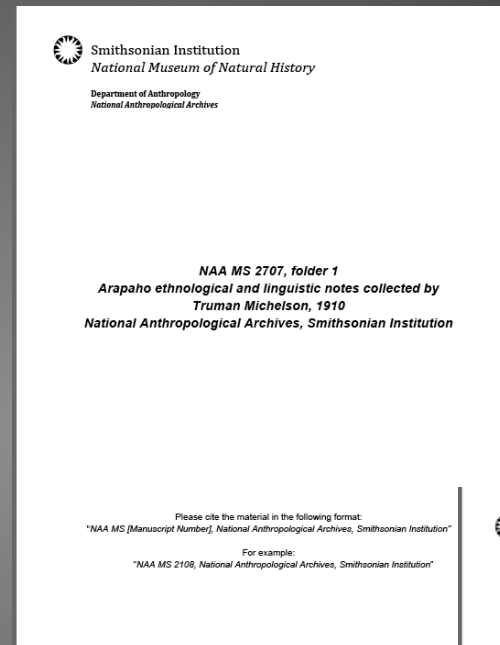
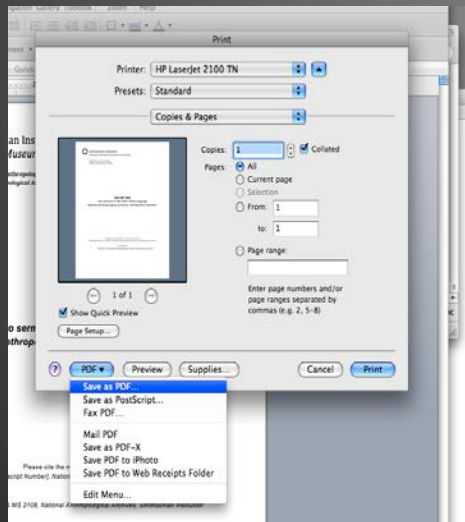
In the JPEG options window, make sure the options are:  
Quality: 8  
Baseline: Standard  
Click Ok. Close window.





# How to create access - derivatives

You may want to create a Word file of the designated PDF Cover Sheets. There are both generic cover sheets and project or material specific cover sheets.

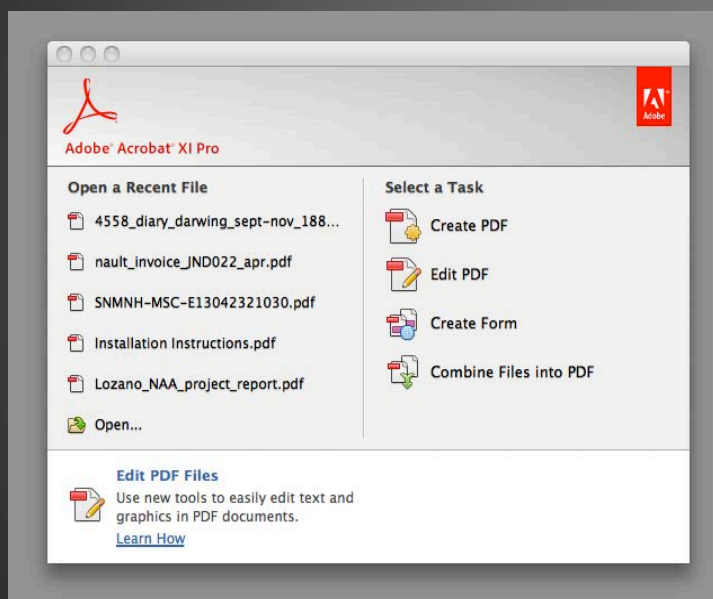




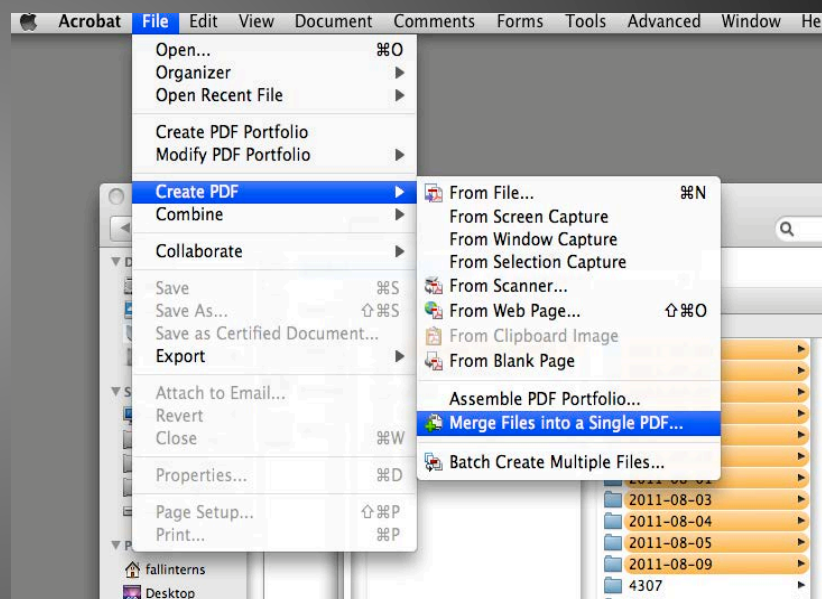
# How to create access - derivatives

## Combining Files and Cover Sheet into PDF file.

Open Adobe Acrobat Professional.



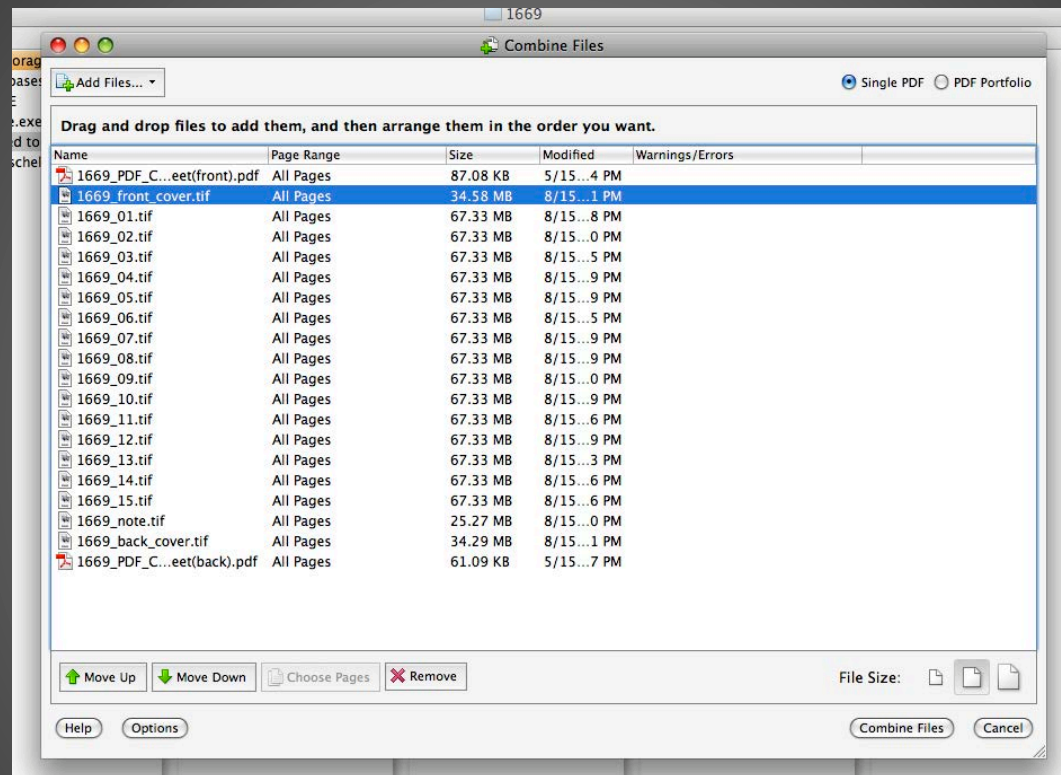
Or you may choose File→Create PDF/or Combine→Merge/Combine Files into a Single PDF



A Start Menu may pop up. You may choose “Combine Files into PDF”



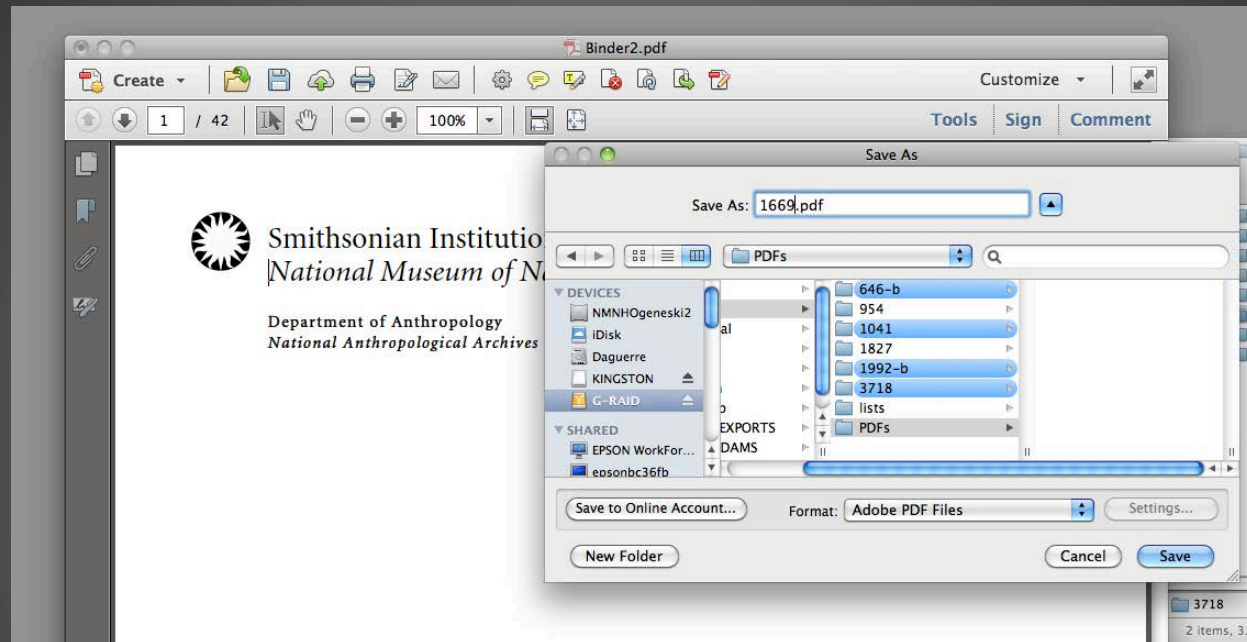
# How to create access - derivatives



In the Combine Files window that opens select “Add Files” or drag files and PDF cover sheets into the window. Make sure Single PDF is checked and make sure the file size icon is on the middle setting. Add the individual images to be combined, or a whole folder containing them all.



# How to create access - derivatives

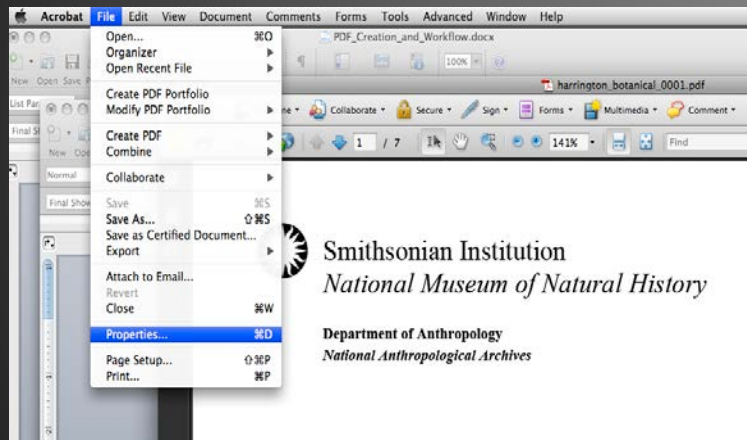


Click Combine Files. Depending on the number of images, it may take a few or several minutes to finish; you can watch the progress at the bottom of the window. Once the PDF is created, it will appear in a new window, starting with the cover sheet. From here, go to File→Save.

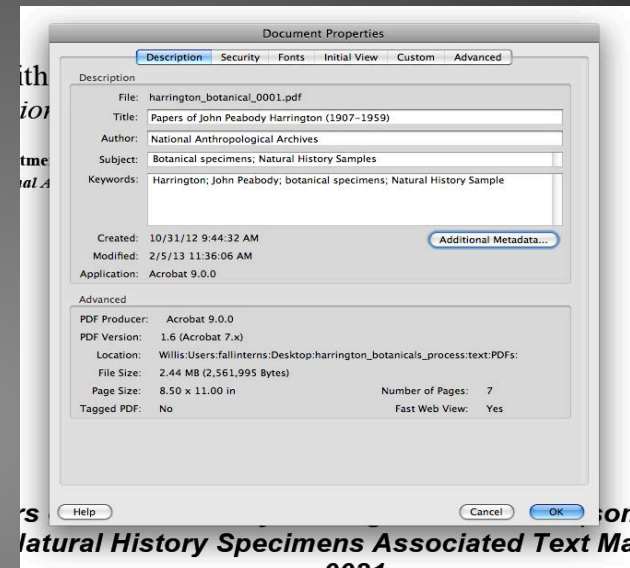


# How to create access – derivatives

Once the file has been saved, go to File→Properties. A “Document Properties” window will pop up



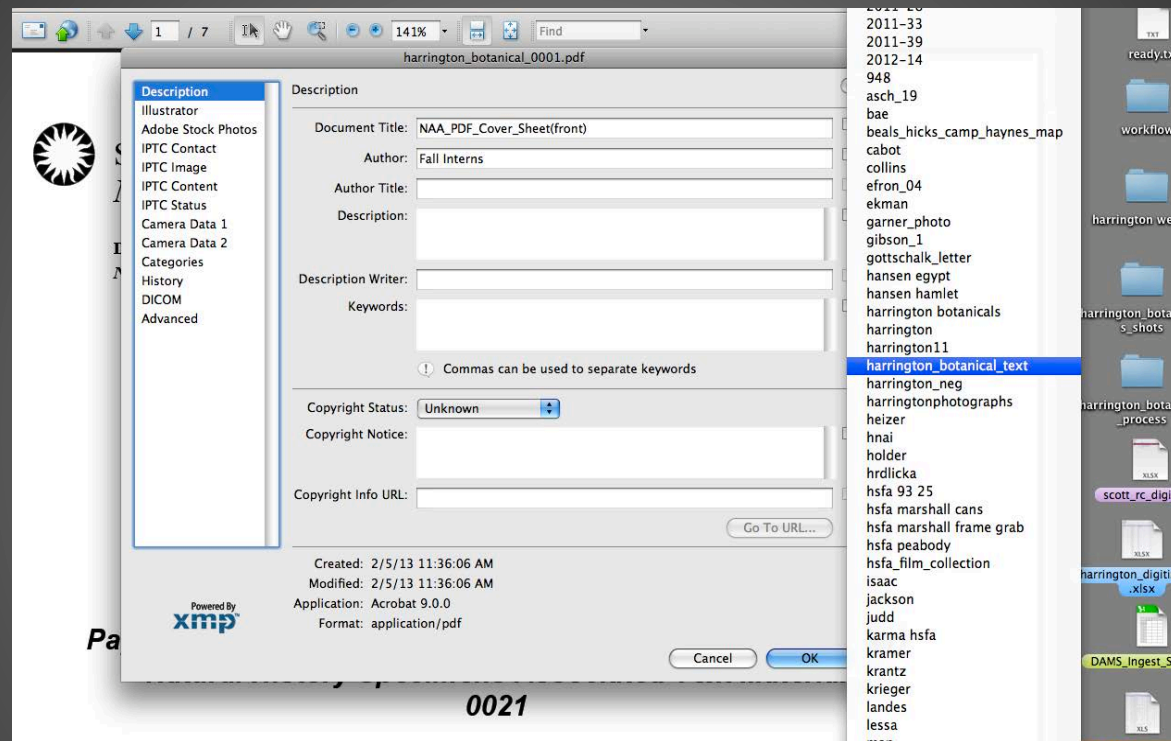
(If the metadata template has been previously saved/exported, it will be available in Acrobat.)



Click “Additional Metadata”



# How to create access - derivatives



A File Info window will pop up (it will look similar to the one in Photoshop). Click the arrow to the right of “Description.” A list of metadata templates will appear. Choose the appropriate template. The fields will populate. Click Ok. Save PDF before closing.



Smithsonian  
*National Museum of Natural History*

ANY QUESTIONS?!

**THANK YOU!**

naultj@si.edu

*Convening Great Lakes Culture Keepers: A Regional Institute for Tribal Librarians, Archivists, and Museum Curators ,  
April 26-29, 2015 at the Mille Lacs Indian Museum and Trading Post*