General Resources for Digitization Workflows

Sample Technical Specifications for Archival Masters

For Text (from CDP Digital Imaging Best Practices)

Text			
	Master	**Access**	**Thumbnail**
File Format	TIFF	JPEG	JPEG
Bit Depth	1 bit bitonal 8 to 16 bit grayscale 48 bit color	8 bit grayscale 24 bit color	8 bit grayscale 24 bit color
Spatial Resolution	Adjust scan resolution to produce a minimum pixel measurement across the long dimension of 6,000 lines for 1 bit files and 4,000 lines for 8 to 16 bit files	150 – 200 PPI	144 PPI
Spatial Dimensions	4000 to 6000 pixels across the long dimension	600 pixels across the long dimension	150 to 200 pixels across the long dimension

For Photographs (from CDP Digital Imaging Best Practices)

Photographs			
	Master	**Access**	**Thumbnail**
File Format	TIFF	JPEG	JPEG
Bit Depth	16 bit grayscale 48 bit color	8 bit grayscale 24 bit color	8 bit grayscale 24 bit color
Spatial Resolution	400 to 800 PPI	150 to 200 PPI	144 PPI
Spatial Dimensions	4000 to 8000 pixels across the long dimension, depending on size of original, excluding mounts and borders	600 pixels across the long dimension	150 to 200 pixels across the long dimension

For Audio (from CDP Digital Audio Best Practices)

Sample Rate	Bit Depth	Pros	Cons
44.1 kHz	16 bit	No file format conversion needed for Audio CD. Maximizes storage space. Appropriate for lower quality source files. Lowest level of processing time. Ubiquitous home audio standard. International standard for Compact Disk (Red Book Standard).	Lowest frequency range acceptable. May not provide sufficient quality for future formats. May have limitations for publication or broadcast, and migration to future digital formats. Limits ability to enhance source file for delivery.
44.1 kHz	24 bit	More accurately reproduces sound of source material. Increased capability to enhance source file for delivery. Increased dynamic range. Acceptable for publication and broadcast. Reflects current professional audio standards.	
96 kHz	24 bit	Standard for DVD/HD Audio. Increased frequency range. More accurately reproduces sound of high frequency, high quality source material, such as musical recordings. Increased potential for enhancement of source file for delivery. More potential for future applications. Potential recommended benchmark for future. Highest recommended current quality. Rapidly growing acceptance. Reflects emerging professional audio standards.	Increased storage space. Increased processing time. No perceptible improvement in sound quality for some source files. Requires conversion to 16 bit and 44.1kHz for delivery on Red Book Audio CD. May require frequency compression for delivery.

For Film and Video (CARLI Digitization Best Practices for Moving Images)

Best practice:

For each program of moving image material, the initial digitization should strive to create an uncompressed, high-quality archival master wherever possible. Uncompressed video requires an enormous amount of storage space, but an uncompressed master is crucial to preserving the integrity of the content over the long term.

- Uncompressed YCbCr or JPEG2000 lossless encoding (codec)
- 640 x 480 resolution (assuming 4:3 original aspect ratio)
- 30 bit sample size
- progressive scanning
- 30 MiB/s data rate
- MXF (.mxf) file format

Acceptable practice:

Archival masters created using lossy compression are not ideal, but may be used when sufficient storage space is unavailable or the material is deemed of less historical importance.

- MPEG-4 AVC (H.264) or DV encoding (codec)
- 640 x 480 resolution (assuming 4:3 original aspect ratio)
- 30 bit sample size
- progressive scanning
- 30 MiB/s data rate
- AVI (.avi) or QuickTime (.mov) file format

File Naming and Directories

File names should be unique and systematic. One method for creating such file names is to think of a file as a set of coordinates.

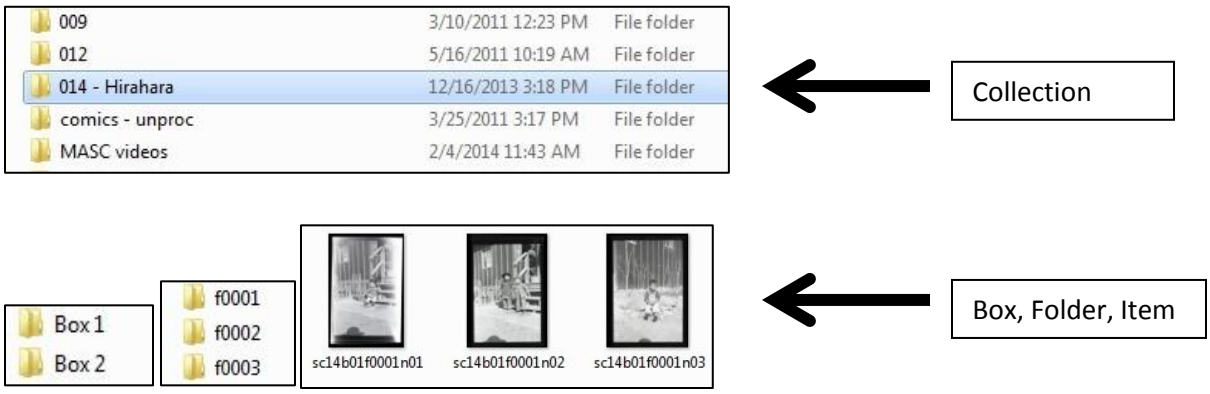
The following item is from Special Collection 14.1, Box 1, Folder 4.

The file name for this is:

sc14_1b01f004n01

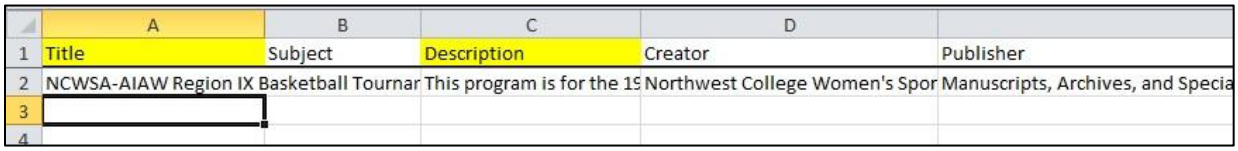
- **Special Collection 14.1**
- **Box 1**
- **Folder 4**
- **Item number 1**

Archival masters should be arranged according to their physical originals.
Example:



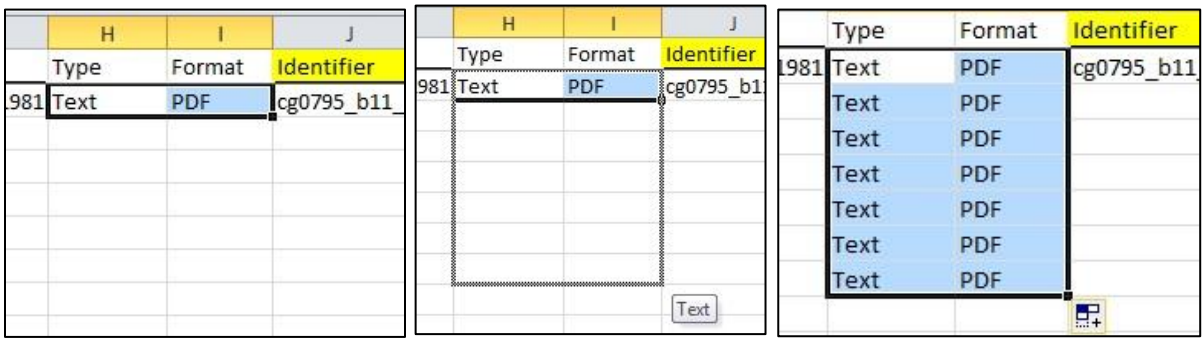
Creating Metadata

While there are many ways to create metadata, one relatively easy method is by using a spreadsheet. Start by opening a new blank spreadsheet and entering metadata fields in the first row. Below, is an example of a spreadsheet set up for Dublin Core.



	A	B	C	D	
1	Title	Subject	Description	Creator	Publisher
2	NCWSA-AIAW Region IX Basketball Tournar		This program is for the 19	Northwest College Women's Spor	Manuscripts, Archives, and Specia
3					
4					

If there are fields that can be filled in with a sequence or the same value, use the autofill option to streamline the process.



	H	I	J
	Type	Format	Identifier
1981	Text	PDF	cg0795_b11_

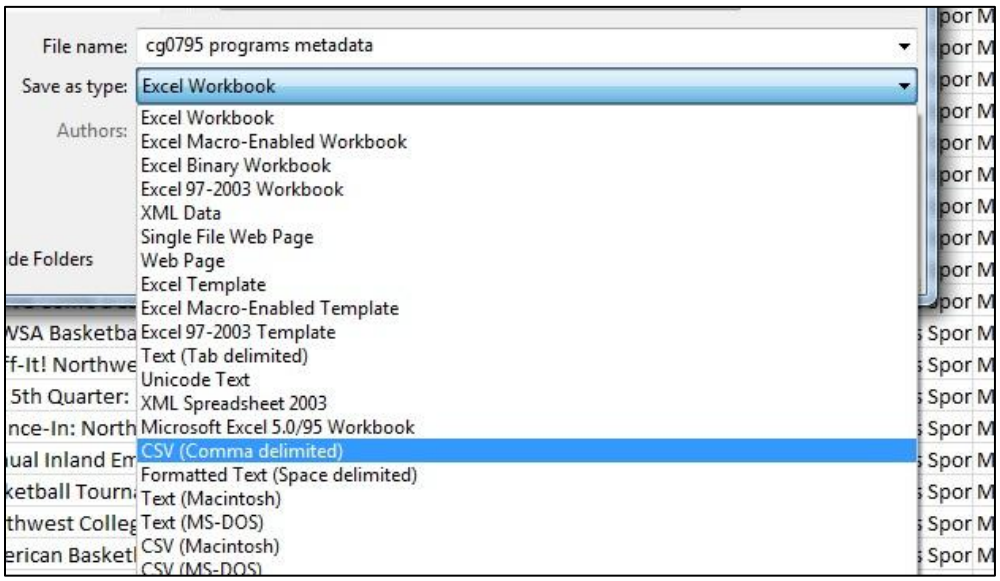
	H	I	J
	Type	Format	Identifier
1981	Text	PDF	cg0795_b1

Text

	Type	Format	Identifier
1981	Text	PDF	cg0795_b11
	Text	PDF	
	Text	PDF	
	Text	PDF	
	Text	PDF	
	Text	PDF	
	Text	PDF	

If multiple people are working on the project, you may want to designate which fields are to be filled out by whom. In the above case, the fields highlighted in yellow are to be filled out by the digitizer.

If you are uploading to a content management system, you may need to convert the spreadsheet to a specific type of text file (e.g., comma separated values, tab delimited). Many spreadsheets allow for this conversion under the “save as” command.



Sharing Platforms

Content management systems have become commonplace in most university library systems. Below are some factors to take into consideration when selecting and implementing a content management system.

From "Content Management Systems in Libraries: Opportunities and Lessons Learned," by Jonathan Blackburn, Eli Neiburger, Karen Coombs

6 lessons learned (by Blackburn):

1. Start with a content management plan: who does what, when, and how often (and how are they accountable)
2. Get staff input: find out what your content creators want
3. Secure support from administration: if they're not behind it, it will never happen
4. Choose right tool(s) for the job: if it doesn't meet the organization needs, don't use it, no matter how "cool" it is
5. Be flexible and embrace workarounds
6. Outsource when possible

Trade-offs between CMSs observed by Karen Coombs:

- flexibility vs simplicity
- customization vs staff resources
- staff skill set vs ease of use
- empowerment vs responsibility
- support vs functionality
- one tool vs many tools

From "Digital content management: the search for a content management system," by Yan Han, Library Hi Tech, volume 22, number 4 (2004).

Han discusses the selection of a CMS based on 5 categories:

- Preservation issues
- Metadata issues
- Access issues
- System features
- Other noticeable issues

Other CMS Resources

Bradford, Eden Lee. *Content Management Systems in Libraries: Case Studies*, Lanham, MD: Scarecrow Press, 2008.

Connell, Ruth Sara. "Content Management Systems: Trends in Academic Libraries," Information Technology and Libraries, volume 32, number 2 (June 2013).

Handling Guidelines

Library of Congress Collections Care
http://www.loc.gov/preservation/care

Northeast Document Conservation Center (NEDCC) Preservation Leaflets (for paper and photos)
http://www.nedcc.org/free-resources/preservation-leaflets/overview

Magnetic Tape Storage and Handling: A Guide for Libraries and Archives, by John Van Bogart
http://www.clir.org/pubs/abstract//reports/pub54

National Film Preservation Foundation Film Preservation Guide
http://www.filmpreservation.org/preservation-basics/the-film-preservation-guide-download

Specifications and Standards

FADGI Guidelines (multiple formats)
http://www.digitizationguidelines.gov/guidelines/

BCR's CDP Digital Imaging Best Practices (images and text)
http://mwdl.org/docs/digital-imaging-bp_2.0.pdf

BCR's CDP Digital Audio Best Practices
http://www.mndigital.org/digitizing/standards/audio.pdf

CARLI Digitization Best Practices for Moving Images
http://www.carli.illinois.edu/sites/files/digital_collections/documentation/guidelines_for_video.pdf

ALCTS Minimum Digitization Capture Recommendations (multiple formats)
http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations#photographic_processes

Metadata Standards

General Resources

Standards at the Library of Congress
http://www.loc.gov/standards/

NISO's Understanding Metadata
http://www.niso.org/publications/press/UnderstandingMetadata.pdf

FADGI Homepage (for metadata guildelines and more)
http://www.digitizationguidelines.gov/

Technical Metadata

University of Illinois Best Practices for Technical Metadata
http://www.library.illinois.edu/dcc/bestpractices/chapter_10_technicalmetadata.html#10.1MinimumRequirements

Metadata Schemas

PREMIS (preservation metadata)
http://www.loc.gov/standards/premis/

METS (Technical and Descriptive)
http://www.loc.gov/standards/mets/

Dublin Core
http://dublincore.org/

CDP Dublin Core Best Practices
http://www.mndigital.org/digitizing/standards/metadata.pdf

Vocabularies

Art and Architecture Thesaurus
http://www.getty.edu/research/tools/vocabularies/aat/

ISO 639-2 Language Codes
http://www.loc.gov/standards/iso639-2/php/code_list.php

Library of Congress Authorities
http://authorities.loc.gov/

Thesaurus of Geographic Names
https://www.getty.edu/research/tools/vocabularies/tgn/