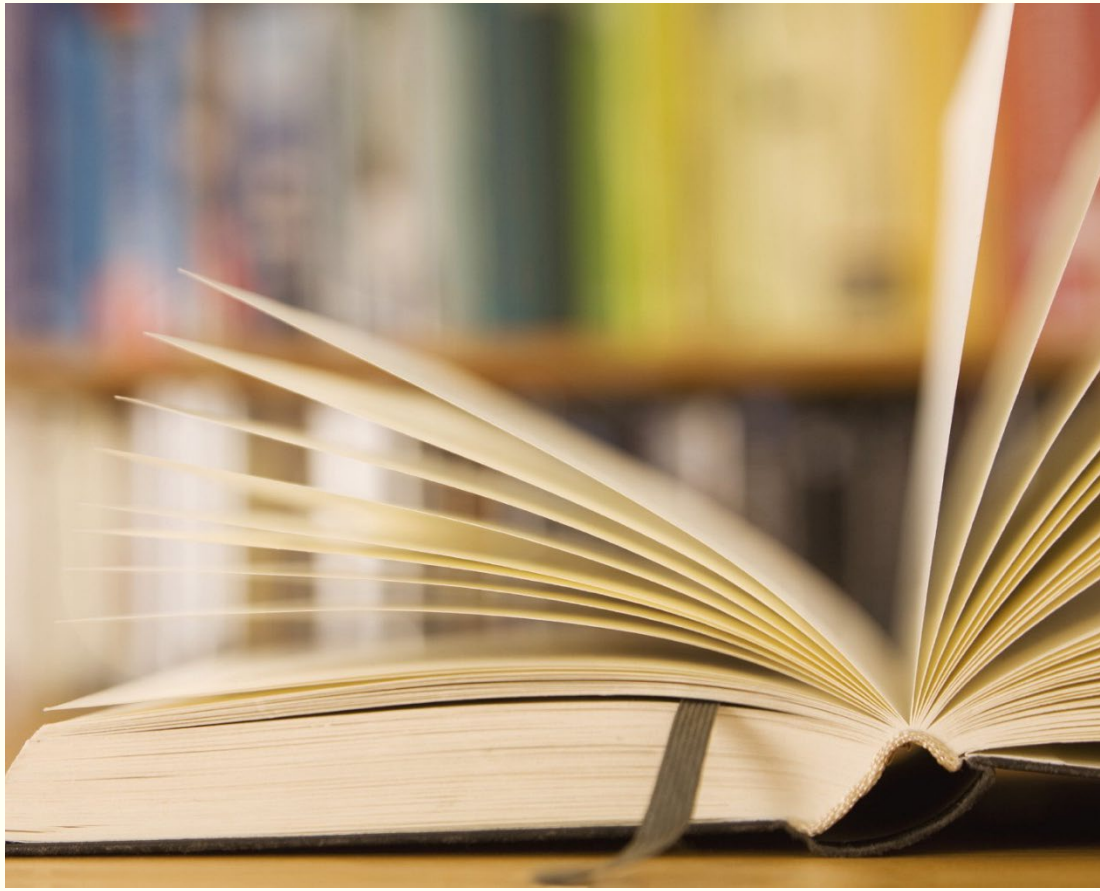
DIGITIZATION, METADATA, & CONTENTdm

Partnering to Build Your Digital Collections



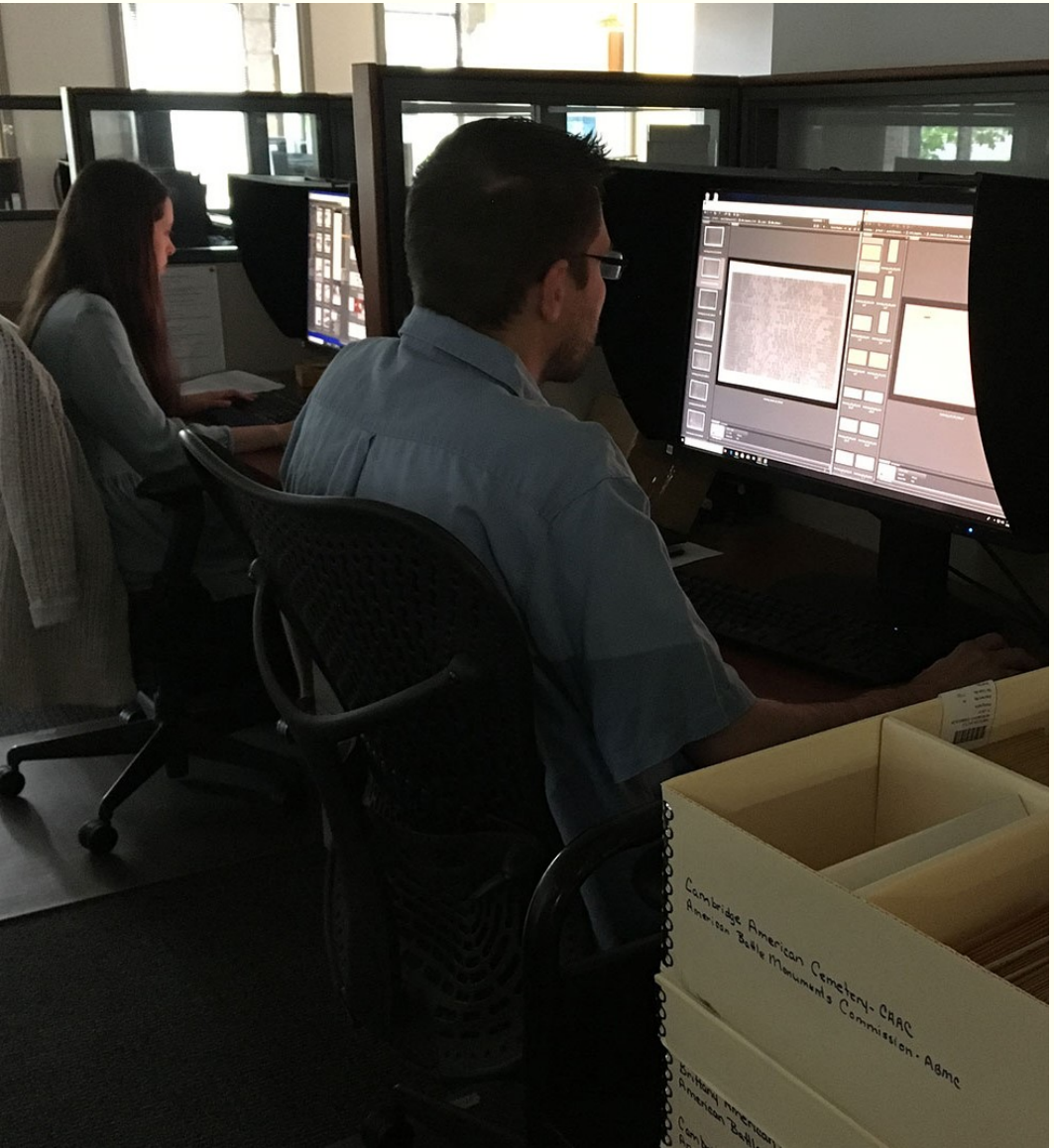
Preparing for a Digitization Project

- How to prepare materials and metadata for digitization
- What makes a collection “camera ready”?
- How does Backstage collect metadata?
- Collection-level metadata
- Final stage of a digital collection: Ingestion of files and metadata into CONTENTdm

Poll: What stage are you at in your digital projects?

HOW TO PREPARE MATERIALS AND METADATA FOR DIGITIZATION

Determine how you want the final product to appear for the end user



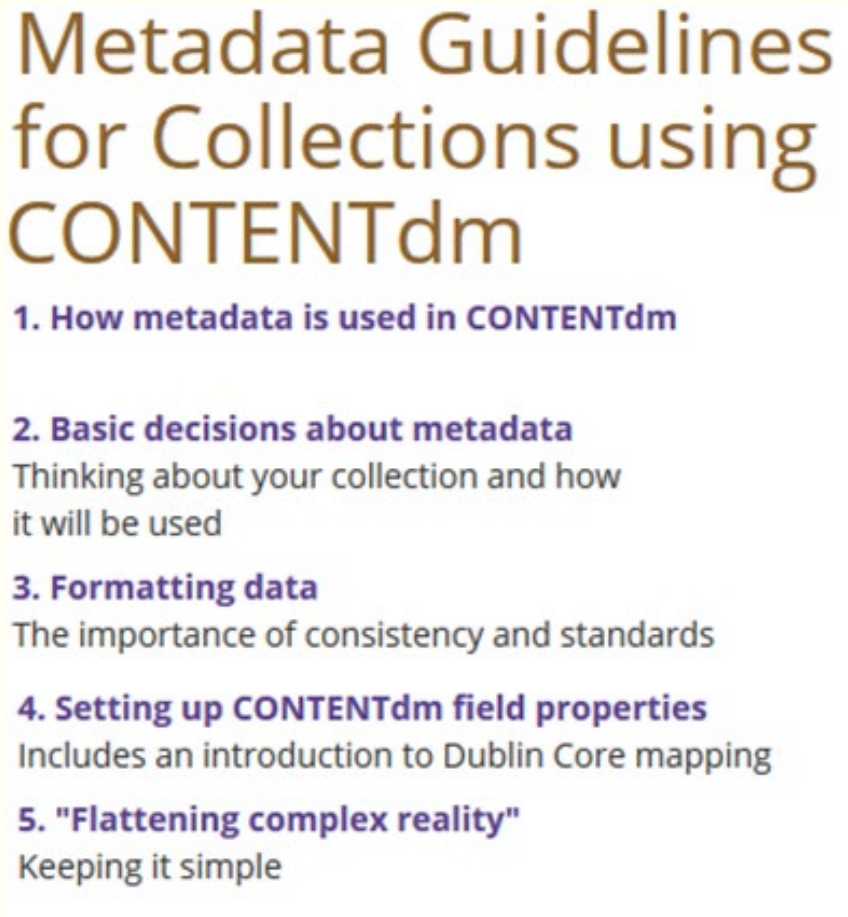
The first step for any digitization project is to ask:

How will the collection be used?

What is the audience level?

How much detail does the metadata relating to the collection need to include?

Organizing the Collection and Metadata



CONTENTdm organizes files as objects in the collection build.

- The key for objects is to determine the level needed to distinguish one object from another.*
- An object may be a bound volume or issue.
- Manuscript collections and loose pages may be loaded as folder objects, organized by factors such as date, subject, or topic.
- Each object will contain individual files/images.
- What metadata do your objects need?
- How does the metadata you plan to gather connect to the objects?

* University of Washington, https://www.lib.washington.edu/cams/mig/advice

Determine what metadata you would like shown with the files

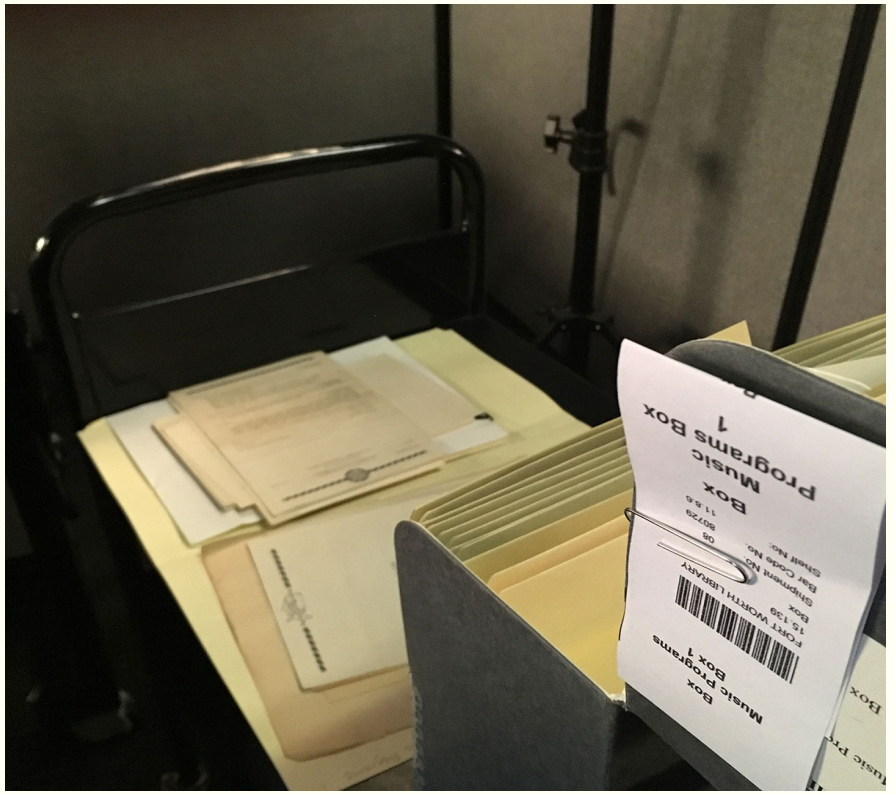
Based on the University of Washington's cataloging and metadata guidelines, you should ask yourself while preparing metadata:

- What kind of information do you need to describe the items or collection?
- What do the end users need to know about the items and the collection? (for example: where did the material come from, what is the collection's significance, who created the items in the collection)
- How much detail is needed for the collection?

The University of Washington library recommends following standards to collect and ingest metadata into a build. This will help the metadata be consistent, which in turn will increase searchability across multiple collections and within the collection.

Preparing the Physical Material for Digitization

Organize the material based on how you would like to see the material structured in the final build



Process the collection to ensure it is camera ready

WHAT MAKES A COLLECTION “CAMERA READY”?

The digitization equipment and methodologies that Backstage has in place have been chosen to minimize the risk of damaging materials during the digitization process. In order to achieve that goal, it is our recommendation that the materials be prepared by the institution to qualify them as camera ready.

Collections that do not qualify as camera ready may still be processed for digitization when the institution grants permission for Backstage or the vendor to perform basic material preparation. Material preparation may incur costs if performed by the digitization vendor.

Recommended Material Preparation Steps

Enclosures

- Items have been removed from their enclosures
- Staples and other fasteners have been removed

Items arrive flat and unfolded

- Newspapers have been unfolded
- Creases have been flattened
- Dog-ears have been unfolded

Misc Objects Removed

- Bookmarks, researcher notes, flags, sticky notes, and other ephemera have been removed unless the material is meant to be digitized

Bindings can lay flat at 180-degrees

- Covers and pages turn without breaking
- Fragile and broken bindings have been removed
- Tight bindings have been removed or strings have been cut

Recommended Material Preparation Steps

Items can be flattened

Items can be flattened with glass in a book cradle or on a copyboard stand

Metal bindings have been removed

Books bound with metal studs/grommets have been disbound

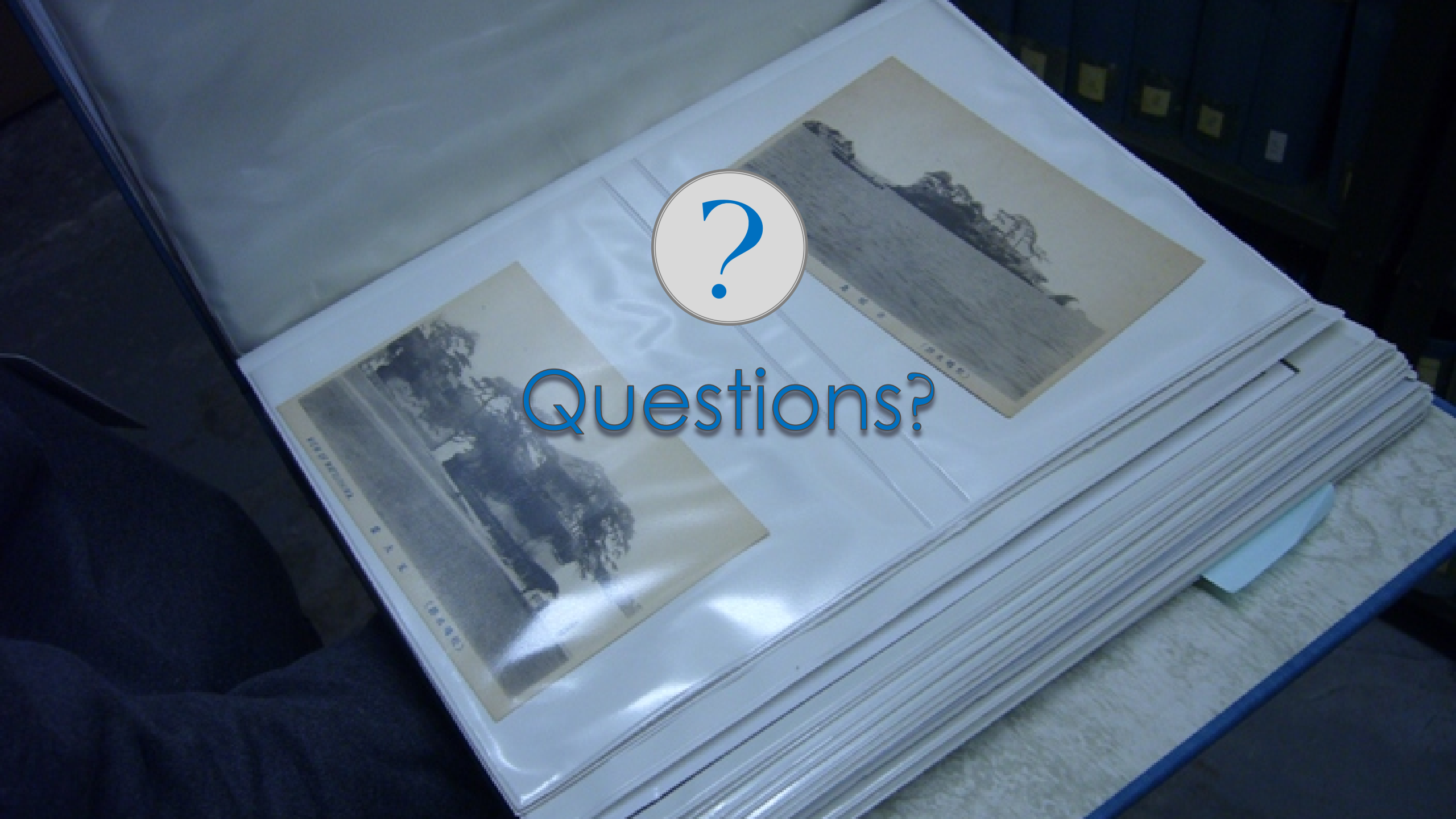
Duplicates Removed

Collections have been weeded for duplicates or flagged to alert the digitization vendor not to digitize duplicates

Microfilm Type and Condition

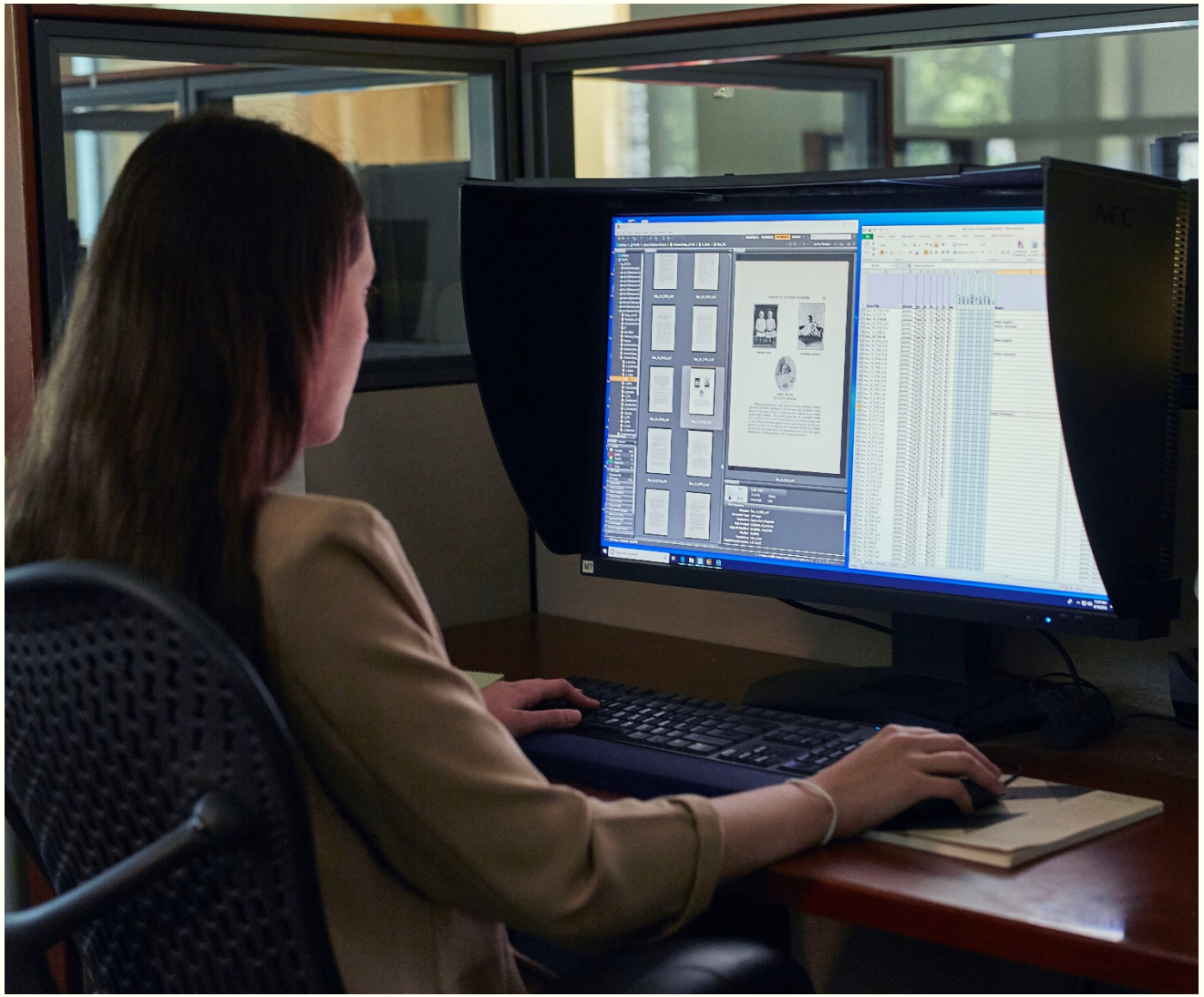
Microfilm should be polyester-based and of good condition.

If the film does not qualify and is of acetate-based film, Backstage offers microfilm reformatting services which may include duplication to 35mm or 16mm polyester film.


?
Questions?

HOW DOES BACKSTAGE COLLECT METADATA?

During the digitization process, Backstage can collect metadata using different methods to rename and organize the digital files into objects for ingestion into CONTENTdm.



Sources of metadata (on the item level)

Client-provided metadata

Metadata can be provided by the client in a universal file format, such as an Excel spreadsheet or Google sheet, and can be a great way to name the files with specific file names

Structural-level Metadata

Metadata that is considered structural-level can contain folder numbers, box numbers, volume numbers/titles, and issue dates.

Item-level Metadata

This metadata can be more complex and extensive, and can be used if the collection is organized at the item level and not at the structural level.

For example, in a correspondence collection, each letter might be an object, rather than at the folder or box-level.

Client File Naming Sheet Example

Box Number	Folder Number	Title	File Name
1	1	Correspondence 1940-1945	ACC1978_1940-1945
1	2	Correspondence 1945-1955	ACC1979_1945-1955
1	3	Correspondence 1955-1960	ACC1980_1955-1960
2	1	Correspondence 1960-1965	ACC1981_1960-1965
2	2	Correspondence 1965-1980	ACC1982_1965-1980

COLLECTION-LEVEL METADATA

Collection-level metadata for CONTENTdm builds

This metadata will remain static throughout the collection and is typically provided by the institution to the vendor:

- While determining the collection-level metadata that will be used, Backstage provides our clients with a working metadata document that allows the institutions to plan out the metadata that will be ingested into the build.
- In addition to static metadata, the template serves as a schema guide for relative metadata determined from folder and file objects (i.e., naming, technical metadata)

Backstage is capable of assisting with the formatting and identifying of collection level metadata.

Backstage Collection-level Metadata Template Example

Your Field Name	DC Map	Date Type	Provided by	Examples *(from various collections)*	**Your Metadata**	Notes	Searchable?	Hidden?	Required?	Vocab?
Title	Title	Text	client	*BrownAndWhite_19051107*		The Title field name can be anything, but "Title" DC map is compulsory. (This is the title assigned to each compound object.)	Yes	No	Yes	No
Subject	Subject	Text	client			Subject and keywords.	Yes	No	No	No
Description	Description	Text	client	*An archive of the Montclarion student newspaper from Montclair State University in New Jersey.*		Examples include, but are not limited to, an abstract, table of contents, reference to a graphical representation of content, or free-text account of the content.	Yes	No	No	No
Creator	Creator	Text	client				No	No	No	No
Publisher	Publisher	Text	client	*Lehigh University*			No	No	No	No
Contributors	Contributors	Text	client				No	No	No	No
Date	Date	Date	BSLW	*1905-11-07*		If Data Type "Date" instead of "Text" then adhere to ISO 8601 convention.	No	No	No	No
Type	Date	Text	client	*text*			No	No	No	No
Format	Format	Text	client	*image/tiff*			No	No	No	No
Identifier	Identifier	Text	client	*7019854*			No	No	No	No
Source	Source	Text	client	*newspaper*			No	No	No	No
Language	Language	Text	client	*eng*			No	No	No	No
Relation	Relation	Text	client				No	No	No	No
Coverage	Coverage	Text	client	*United States; Pennsylvania; Lehigh County; Bethlehem*			No	No	No	No
Rights	Rights	Text	client	*Public domain*		Intellectual Property Rights (IPR), Copyright, and various Property Rights.	No	No	No	No
Full Text	None	Full Text Search	BSLW	*[OCR results or transcript contents]*			No	No	No	No

FINAL STAGE OF A DIGITAL COLLECTION: INGESTION OF FILES AND METADATA INTO CONTENTdm

After the files have been organized into their corresponding objects, the files and metadata will be ingested into the CONTENTdm build.

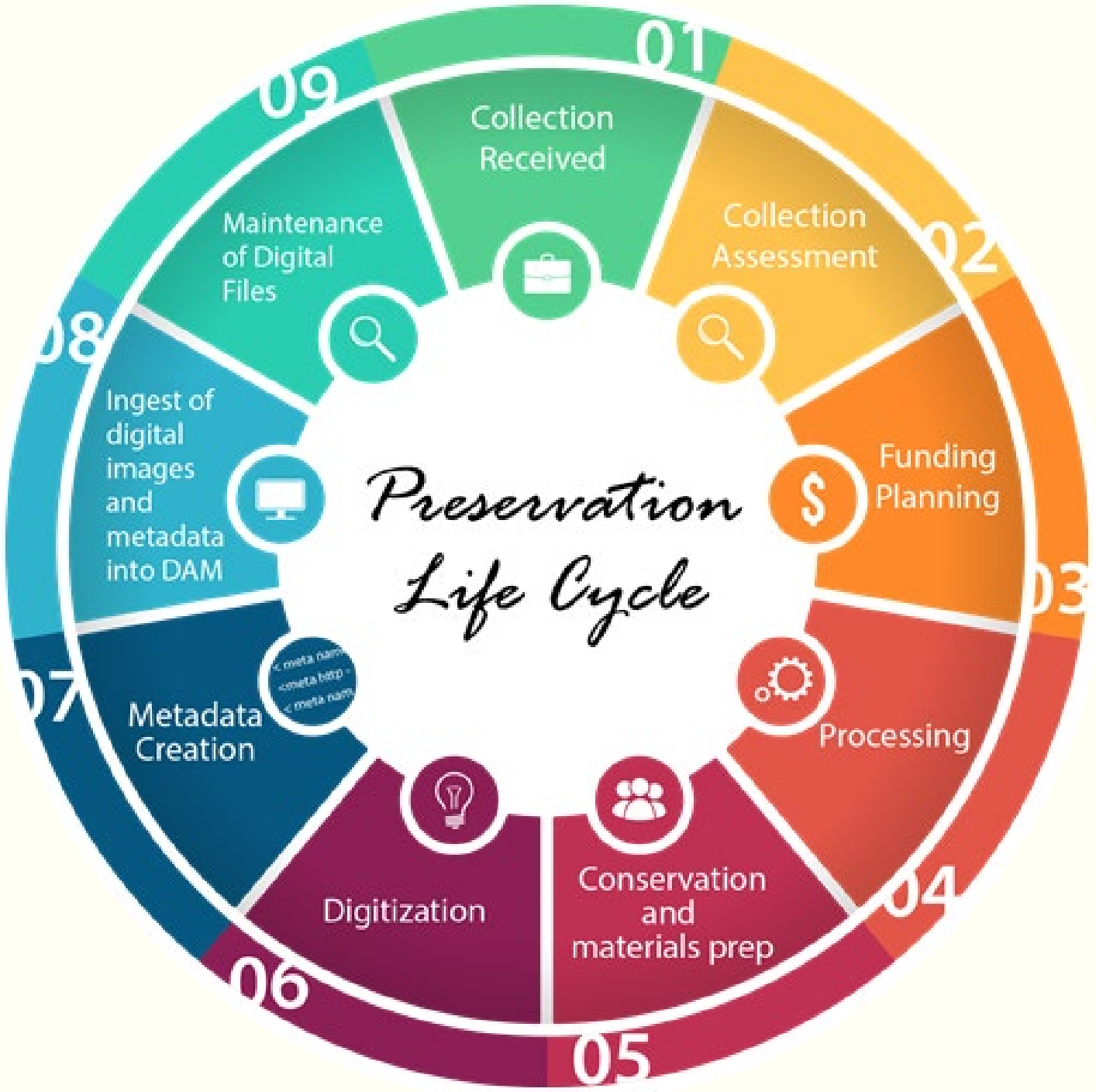
You will receive a CONTENTdm file set as a back-up, which includes:

- JPEG-2000 access images
- JPEG thumbnail images
- Uncorrected OCR
- PDFs, bound per object

We also recommend creating derivatives that can be accessed outside your build.

Recommended files:

- Archival-quality master TIFFs
- PDFs, bound per object
- JPEG access images for each TIFF


Preservation
Life Cycle
01
Collection Received
02
Collection Assessment
03
Funding Planning
04
Processing
05
Conservation and materials prep
06
Digitization
07
Metadata Creation
08
Ingest of digital images and metadata into DAM
09
Maintenance of Digital Files

Questions or Comments?

Thank you for joining us today!

Backstage
LIBRARY WORKS

1.800.288.1265
info@bslw.com
www.bslw.com
Provo, Utah and Bethlehem, Pennsylvania