1. Quick Guide to Publishing a Digital Collection

This document provides an overview of the University of Illinois at Urbana-Champaign Library’s process for proposing and executing projects to digitize locally held special collections materials and publish them in the Digital Collections.

1. **PLANNING** (2-3 weeks)
   a. **Curator** proposes digitization project using online form. Digital Collections Project Manager (Angela Waarala) receives form and assesses feasibility of project with Curator and staff in Conservation and Metadata Services. Once approved, stakeholders schedule when their staff will be available to do the work necessary to complete the project.

2. **METADATA** (varies based on number of items, at least 1-4 weeks)
   a. **Curator** selects or develops a Metadata Profile for collection with Metadata Specialist (Tricia Lampron).
   b. **Curator** authors collection-level metadata in Medusa digital preservation repository.
   c. **Curator** prepares spreadsheet of object-level metadata for items to be digitized.

3. **CONSERVATION** (varies based on Conservation lab commitments and condition of materials)
   a. Depending on the condition of materials, conservation work may be necessary to prepare them for digitization.

4. **DIGITIZATION** (varies based on Digitization lab commitments and number of items, typically weeks to months)
   a. **Curator** provides inventory spreadsheet and items to be digitized to Digitization Services.
   b. Digitization Services completes digitization, packages files, and deposits them in Medusa. Metadata Specialist ingests descriptive metadata into the Digital Library System.

5. **REVIEW** (varies based on number of items and quality of metadata, typically 1-3 weeks)
   a. Metadata Specialist and Curator review collection and correct errors in metadata and display as needed.

6. **PUBLISH** (1 day)
   a. Metadata Specialist makes collection public in the University of Illinois Library Digital Collections.

For an expanded view of this document, with a deep dive into the technical side of getting things done in the digital library, see the full Managing Digital Collections: A Guide to Local Practice manual.
Managing Digital Collections: A Guide to Local Practice

edited by Kyle Rimkus with contributions from Bethany Anderson, Jesus Espinosa, Tom Habing, Bill Ingram, Patricia Lampron, Elizabeth Lippoldt, Joshua Lynch, Tracy Popp, Christopher Prom, Kyle Rimkus, Madison Sullivan, Angela Waarala, Kristina Williams, and Sarah Yarrito

last edited: October 12, 2017

| 1. Quick Guide to Publishing a Digital Collection | 1 |
| 2. Digital Collections Management Staff and Services | 4 |
| 3. Overview of Services | 5 |
| 3.1. Medusa | 5 |
| 3.2. Digital Collections | 5 |
| 3.3. Technical Architecture | 6 |
| 4. Medusa FAQ | 8 |
| 5. Using Medusa | 11 |
| 5.1. Access Permissions | 11 |
| 5.2. Accessing Files in Medusa | 11 |
| 5.3. Add a New Collection | 14 |
| 5.4. Deposit Content to a Collection | 18 |
| 5.4.1. Create a File Group | 19 |
| 5.5. Considerations for Archival Units | 21 |
| 5.5.1. Relate a Medusa Collection Record to an Archon Collection Record | 21 |
| 5.5.2. University Archives File Package for Electronic Records | 21 |
| 6. Using the Digital Library | 23 |
| 6.1. Structured, Free-form, and External collections | 23 |
| 6.1.1. Packaging for Structured Collections | 23 |
| 6.1.1.1. Single-item Object | 23 |
| 6.1.1.2. Compound Object | 24 |
| 6.1.1.3. Mixed Media | 26 |
| 6.1.2. Packaging for Free-form Collections | 26 |
| 6.1.3. Medusa Metadata versus Digital Library System Metadata | 27 |
| 6.1.4. Index Medusa Collections Metadata in the Digital Library System | 27 |
| 6.2. Common Problems and their Solutions | 28 |
| 6.2.1. Change the Package Profile of a Collection | 28 |
7. The Recommended Project Workflow 30
   7.1. Planning 31
   7.2. Metadata 31
      7.2.1. Select or Create Metadata Profile 31
      7.2.2. Edit Collection-level Metadata 33
      7.2.3. Create Object-level Metadata 33
   7.3. Conservation 36
   7.4. Digitization 36
   7.5. Review 36
      7.5.1. Import Digitized Items into Digital Library 36
      7.5.2. Edit Existing Metadata 41
         7.5.2.1. Edit Metadata in Web Interface 41
         7.5.2.2. Edit Metadata in a Spreadsheet 44
   7.6. Publish 46

8. Medusa Back-end Management 49
   8.1. Ingest Files into Medusa 49
   8.2. Ingest Request Management Overview 49
      8.2.1. University Archives 49
      8.2.2. DS (Digitization Services) 50
      8.2.3. Conservation 50
      8.2.4. Media Preservation Services 50
      8.2.5. Sousa Archives and Center For American Music 50
      8.2.6. Newspapers 51
      8.2.7. Rare Book and Manuscript Library 51
      8.2.8. Other 51
   8.3. Ingest Content 51
   8.4. Add Events 53
   8.5. Approve Ingests 54
      8.5.1. Medusa Notification Emails 55
   8.6. Add items to Digitization Queue as a Medusa Project 56
   8.7. Delete Content from Medusa 61
   8.8. Delete Individual Files or Folders from a File Group 62
   8.9. Restructure Content in Medusa Storage 62
      8.9.1. To edit content on Medusa Production\Primary Storage: 62
      8.9.2. Once the content has been corrected: 63
      8.9.3. Once you delete the files/folders: 63
   8.10. Notes on Deleting or Moving/rename/restructuring Files and Folders 64
   8.11. Useful Tools 64
2. Digital Collections Management Staff and Services

Responsibility for digital collections services, understood in this document to comprise activities related to the Medusa preservation repository and the Digital Collections access service for locally managed digital content, is distributed across a number of departmental units and committees in the library. Chief among these are:

- Content Access Policy and Technology committee and its Digital Production work group; Repositories, Preservation, and Access work group; and Cataloging and Metadata work group.
- Preservation Services unit, where the library has consolidated oversight for the majority of its digital reformatting and digital preservation efforts.
- Content Access Management (a.k.a. Cataloging and Metadata) unit, which manages descriptive metadata on digital projects.
- A variety of people contribute work on digital projects, but the staff roles most frequently referenced in this manual are:
  - Digital collections project manager (Angela Waarala), who plans and manages digitization projects.
  - Metadata specialist (Tricia Lampron), who coordinates metadata effort on digital projects.
  - Digital preservation coordinator (Tracy Popp), who manages Medusa digital preservation repository services.
  - Digital Library Technical Coordinator (Henry Borchers), who supports workflows to ensure files are properly packaged for deposit into their target preservation and access systems.
  - Curator: In this document, the word curator is used as a generic term to designate a collection manager, for example an archivist or rare book librarian, who serves as the primary point of contact for a collecting unit. We recommend that each special collections unit or departmental library with interest in digitization and digital preservation designate a curator to serve as a liaison to the digital projects team.

If you have questions, problems, or suggestions about our digital collections services, send an email to digitalcollections@lists.illinois.edu and someone will contact you. Messages to this email address are monitored by all staff who work in the provision of digital collections services.
3. Overview of Services

This document makes frequent reference to Medusa and its Digital Collections service.

3.1. Medusa

Medusa is the library's digital preservation repository and collection registry. Access to the collection registry is restricted to digital content producers at the University of Illinois Library. The library manages Medusa storage in partnership with the National Center for Supercomputing Applications.

- Live production site: https://medusa.library.illinois.edu/
- Test site (accessible only to library staff): https://medusatest.library.illinois.edu/

3.2. Digital Collections

The digital collections of the University of Illinois at Urbana-Champaign Library provide web access to many of the Library's signature holdings for students, scholars, and the general public. They include digitized books, photographs, newspapers, maps, and other research materials.

The digital collections site features collections powered by a Digital Library System built on top of Medusa services, as well as links out to collections that live in external systems. This clearinghouse of our locally managed digital collections is available at https://digital.library.illinois.edu/collections.

- Live production site: https://digital.library.illinois.edu/
- Test site (accessible only to library staff): https://kumquat-dev.library.illinois.edu/
3.3. Technical Architecture

The Digital Collections site is built on top of Medusa storage. While there is too much going on under the hood to explain here, the diagram below provides a window into our digital library infrastructure. We build and deploy the majority of our locally managed digital collections services using the web programming framework in Ruby-on-Rails, using a number of other freely available open source protocols and software packages.
Medusa sits at the heart of the library’s digital curation efforts as a stable collection registry, as well as a collection management and file storage service. It accommodates two primary types of workflows:

1. Those where internal library users in Digitization Services and Special Collections units deposit content into Medusa and then push this content out for public access using the Digital Library System and its public Digital Collections site.
2. Those where external users deposit content into the Illinois Data Bank or IDEALS institutional repository, and this content is automatically synced to Medusa for safekeeping.
4. Medusa FAQ

What is Medusa?
The Medusa digital preservation repository provides an enduring storage and management environment for the University of Illinois at Urbana-Champaign Library's digital collections. At present, Medusa’s collecting focus is on digitized and "born digital" books, manuscripts, photographs, audiovisual materials, scholarly publications, and research data from the library’s special collections, general collections, and institutional repository. All master files created by the library's digitization units, for example, are by default deposited into Medusa.

Where is Medusa?
Medusa has a web-accessible collection registry, but it only displays information about collections to authenticated, approved users.
- Live production site: https://medusa.library.illinois.edu/
- Test site (accessible only to library staff): https://medusatest.library.illinois.edu/

What Does Medusa Do?
The Medusa collection registry provides collection curators with a simple, web-accessible service for managing preservation actions over time. It features:
- Forms for editing descriptive, administrative, and rights metadata associated with collections
- Download of files or batches of files to local file servers
- Tracking of preservation events, file provenance, and file statistics
- On-demand verification of file fixity (md5 checksum values) for files or batches of files, and ongoing checksum verification of all files every 90 days
- On-demand extraction of technical metadata (using the File Information Tool Set) for files or batches of files
- A File Format registry with tools for elaborating file format policies and normalization paths

Who has access to Medusa?
Medusa collections are not publicly accessible: Medusa is a back-end system for library collection managers, specifically:
- Local Repository units, defined as any organization that is responsible for general curatorial decisions related to the preservation of, access to, and rights status of collections of digital content in Medusa (for example, an archives, institutional repository, or departmental library). These units may designate trusted staff members to gain access to Medusa.
- Local content Producers, defined as any group that creates digital files to be stored in Medusa on behalf of a repository (for example, digital production units within the Library such as Digitization Services, as well as programs like Media Preservation that manage
digital files created by vendors). These groups should also identify trusted staff members to be trained in using Medusa.

**Where are files in Medusa stored?**
The University Library manages Medusa storage in partnership with the National Center for Supercomputing Applications. Medusa’s storage infrastructure consists of:

- two copies of every file, replicated across two distinct campus nodes
- a third copy of every file backed up off-campus (currently in Amazon Glacier)

**How are collections organized in Medusa?**
The design of Medusa’s interface is modeled as a simple hierarchy of conceptual units.

- Repository -- organization responsible for collections, such as the Rare Book and Manuscript Library ([https://medusatest.library.illinois.edu/repositories/23](https://medusatest.library.illinois.edu/repositories/23))
  - Collection -- a collection of related materials, such as the Collins Collection of Irish Political Cartoons ([https://medusa.library.illinois.edu/collections/53](https://medusa.library.illinois.edu/collections/53))
  - File Group -- a batch of objects or files arranged together

**How do I find collections in Medusa?**
You have two options:

- Browse to your Repository ([https://medusa.library.illinois.edu/repositories](https://medusa.library.illinois.edu/repositories)) to view a list of its associated collections; type within the "Narrow" box a collection title or external ID (for University Archives this is the record series expanded to seven digits without slashes) to view specific collections of files.
- Search all of Medusa for a specific file name in the search box, using the wildcard operator (asterisk) to increase chances of hits. Medusa search queries and triages results by file name, folder name, project items, file group descriptive metadata, collection descriptive metadata, and UUID.

**How do I get files out of Medusa?**
Access to files in Medusa is limited to Repository staff and digital content Producers. Once you have navigated to a File Group, folder, or individual file, use the available download options. After clicking either of the download buttons, Medusa will bundle the files, place them on a webserver for download, and send you an e-mail with the file path location. Keep in mind that the length of time it takes for the e-mail to send depends on the size and number of files.

**How do I deposit files into Medusa?**
Currently, all ingest of content is managed through the Preservation Services unit. When files are ready for ingest, contact the Digital Preservation Coordinator or the Graduate Assistant in Digital Preservation.

**To Note**
1. Medusa is a work in progress. New features are introduced regularly, and suggestions are always welcome.
2. Medusa can be slow at times -- We are still learning how to balance multiple processes at the same time; please be patient but report problems to medusa-feedback@lists.illinois.edu. Issues are often resolved within minutes!
5. Using Medusa

5.1. Access Permissions

- **Curators** should contact the **Digital Preservation Coordinator** for login access. The **Digital Preservation Coordinator** also offers hands-on training to curatorial staff wishing to learn how to use Medusa.
- **Curators** in Medusa are assigned to Repositories representing curatorial units (The Rare Book and Manuscript Library, the Map Library, etc.). **Curators** assigned to one Repository in Medusa cannot edit Medusa records for another Repository.

5.2. Accessing Files in Medusa

Access to files in Medusa is limited to Repository staff (content owners) and digital content Producers. Repository staff may navigate to files by drilling into bit-level “File Groups” and reaching a “download” link. Downloads of large batches of files are best achieved at present by request to the **Digital Preservation Coordinator**, although Medusa staff are working on improving this feature in the Medusa web interface.

To search for collections and files in Medusa,

1. Go to medusa.library.illinois.edu and log into Medusa using your credentials.
2. Click on “Repositories.”
3. Scroll down or use the "Filter results" search box to find your Repository.

4. The Repository page shows all of the collections in the selected repository. From this page you can also see any running processes, current accruals, recent back-ups to Amazon Glacier, and file statistics.

5. To view a Collection, double-click its name. The Collection page displays descriptive metadata about the collection, as well as its constituent File Groups.
6. To files and folders as they are structured on Medusa storage, navigate to a File Group and click “View files.”

7. To download all the folders/files in the folder to which you have navigated, click the download box and select “Download Files and Subdirectories.” If you want to download just the files located in this folder and exclude subfolders, click “Download Files.” Clicking either option will generate a message letting you know that the export is being prepared, and that you will receive an email with a download link.
8. Alternately, by clicking on a single file, you will access an interactive viewer with an option for direct download.

5.3. Add a New Collection

*Note: Prior to adding a new collection verify that it is not already in Medusa. Search for the record series and/or title in Medusa before creating any new collection.*

1. Go to the Medusa collection registry site:
   1. Practice site: https://medusatest.library.illinois.edu/
   2. Production site: https://medusa.library.illinois.edu/
2. Click “Collections” on the menu bar. Login with University NetID and password if prompted
3. On the “Collections” page, click “Add Collection” in the top right.

4. You will then input metadata for the new collection using this page’s tabs. Fill in as many fields as you wish, keeping in mind that it is not necessary to edit them all (only the fields in red below are required). If you aren’t sure how to fill in a certain data entry point, you can refer to previously ingested collections and see how they were completed.
   1. “Descriptive Metadata” tab
      
      **NOTE:** Metadata in the “Descriptive Metadata” tab will be made public in the digital library's list of collections ([https://digital.library.illinois.edu/collections](https://digital.library.illinois.edu/collections)) if set to "open".
      
      i. **“Open collection metadata to public list of collections.”** Check this box if you’d like information about this collection to be made public in the digital library. Typically, a Curator will leave the box unchecked until the collection is ready for publication. To best understand how Medusa descriptive metadata will appear to the public on the Collection listing page cited above, view a public collection record in the Digital Collections site (for example, [https://digital.library.illinois.edu/collections/6ff64b00-072d-0130-c5bb-0019b9e633c5-2](https://digital.library.illinois.edu/collections/6ff64b00-072d-0130-c5bb-0019b9e633c5-2)): 
ii. **Title (required).** Provide the public title that will be used to identify the collection, usually taken from the resource itself (e.g. Portrait of Actors).

iii. **Repository (required).** Select the name of the library or curatorial unit associated with the resource from the drop-down menu (e.g. Sousa Archives). If the resources belong to a repository that is not listed, notify the Digital Preservation Coordinator.

iv. **Description.** Provide an official summary of the collection. If the collection is described in Archon, copy the description field from the Archon record. If set to "open," this description will be made publicly available as a collection "abstract."

v. **Access url.** If the collection exists and is publicly accessible in an external system, provide the access URL for its landing page.

vi. **Physical collection url.** If you have an archival finding aid or other form of collection record, provide a link to it here.

vii. **External ID.** If you have an identifier for this collection used within your collecting unit, for example an archival record series ID, add it here. This field is heavily used by the University Archives.
viii. **Representative Image.** If the collection has already been ingested into Medusa, navigate to it and copy the UUID into this field for the image you would like to represent the collection on its landing page.

ix. **Representative item.** This is an enhanced version of *representative image* for collections of compound objects. Provide the UUID for an object folder for the representative image to display an object title and link to the object itself.

x. **Resource types.** Choose from the list of format types to best indicate what is in the collection.

1. **Cartographic:** Includes maps, atlases, globes, digital maps, and other cartographic items.

2. **Software, multimedia:** Appropriate for any electronic resource without a significant aspect that indicates one of the other "Type of Resource" categories. It includes: software, numeric data, computer-oriented multimedia, and online systems and services.

3. **Mixed material:** Indicates that there are significant materials in two or more forms that are usually related by virtue of their having been accumulated by or about a person or body. Mixed material includes archival fonds and manuscript collections of mixed forms of materials, such as text, photographs, and sound recordings.

4. **Notated music:** Graphic, non-realized representations of musical works, both in printed and digitized manifestations that represent the four components of musical sound: pitch, duration, timbre, and loudness.

5. **Three dimensional object:** Includes man-made objects such as models, sculptures, clothing, and toys, as well as naturally occurring objects such as specimens mounted for viewing.

6. **Sound recording-nonmusical:** used when the sound recording is non-musical in nature.

7. **Sound recording:** Used when a mixture of musical and non-musical sound recordings occurs in a resource or when a user does not want to or cannot make a distinction between musical and non-musical.

8. **Moving image:** Includes motion pictures and video recordings, as well as television programs, digital video, and animated computer graphics, but not slides and transparencies. It does not include moving images that are primarily computer programs, such as computer games or computer-oriented multimedia; these are included in "software, multimedia."

9. **Text:** Resources that are textual in nature.

10. **Still image:** Includes two-dimensional images and slides and transparencies.

11. **Sound recording-musical:** Used when a resource is predominantly a musical sound recording.
xi. **Access systems.** If your content lives in an external system like HathiTrust, the Illinois Digital Newspaper Collections, or something else, indicate it here.

2. “Administrative Metadata” tab

   **NOTE:** These fields will never be made publicly available.

   i. **Private description.** Record information about the resource for use only by the digital content managers at the University of Illinois Library. This field will never be publicly displayed.

   ii. **Contact person email.** Enter the net ID of the person most closely responsible for oversight over this digital collection. When entering a Net ID, there is no need to append "@illinois.edu.”

   iii. **Preservation priority.** Select an ingest priority level for the collection from the drop down menu (e.g. low, medium, or high). If unknown, leave blank. Edit this tab to “ingested” after a collection has been ingested.

   iv. **Notes.** If desired, enter information related to a resource not applicable to the other available fields.

3. “Rights Metadata” tab

   i. **Rights basis:** Information about restrictions imposed on access to a resource (e.g. copyright, statute, or license).

   ii. **Copyright jurisdiction:** The source country of the resource where it is protected by copyright.

   iii. **Copyright statement:** Governs how the materials will be used after access has been provided.

   iv. **Access restrictions:** A statement that indicates rights related to access and use of the digital resource.

   v. **Custom copyright statement.** If applicable, enter a free-form custom copyright statement.

4. “Subcollections” tab

   i. If you would like your collection to display as a “parent” collection which comprises several others as subcollections, select your subcollections from the menu here. This feature has been used extensively by the History, Philosophy, and Newspaper Library, which prefers to display newspaper titles under thematic groupings such as “American Popular Entertainment,” e.g.

      https://digital.library.illinois.edu/collections/f411bc00-fc51-012f-c5ba-0019b9e633c5-f.

5. When you are finished creating metadata, click the “create” button on the left to save and create the new collection. You may navigate to your collection in Medusa and edit its metadata at any time.

5.4. Deposit Content to a Collection

   **NOTE:** This section is most relevant to digital preservation technical staff.
5.4.1. Create a File Group

Medusa keeps track of batches of files related to collections. These are referred to in the system as File Groups. The File Groups editing interface allows for extensive information about files to be provided, but most users will only need to indicate the basic information below:

1. In a collection record, select "Add File Group" under the "File Manager" tab.

2. Descriptive Metadata tab
   1. **Title:** Provide a descriptive name as a reference for this batch of files. This will display alongside your file group, so we recommend something simple and descriptive. Some library departments have elected to develop a standard title for file groups. Other suggestions include:
      1. October 2016 Accrual
      2. 2012 TIFF batch
      3. 2009 preservation batch
   2. **Description:** Rarely necessary.
   3. **Access URL:** If known, otherwise leave blank.
   4. **Resource types:** Select whether materials are text, moving image, etc

2. "Administrative Metadata" tab
   1. **External ID:** not currently used (generally documented at collection level)
   2. **Contact Person Email:** not currently used (already documented at collection level)
   3. **Producer:** Choose the production unit from the drop-down list. If your unit is not represented in this list, please contact the Digital Preservation Coordinator.
   4. **Cfs root:** Shows where on the file storage the content is located; will automatically be updated with an auto-ingest.
   5. **Private description:** If necessary, usually left blank.
   6. **Provenance note:** Usually left blank. Enter information related to the origin of the resource and the method by which the resource achieved digital form.
   7. **Acquisition method:** Usually left blank.
8. **Storage level:** Always start with external file group to document file location information. When you auto-ingest, a bit-level file group containing the viewable files will automatically be created in the collection.

9. **External file location:** Give the path to the networked file share where the files are located that you would like ingested into Medusa. **NOTE:** This information is most useful when identifying new external content to be queued for ingest into Medusa, but isn’t used much now that Medusa is an active service.
   1. Drill down until you reach the folder that contains the files.
   2. If you are using Windows, copy the file path from the Computer navigation bar, which would look something like this: Y:\Processed Preservation Files Ready for Medusa Ingest\Legacy Content\Digital Rare Book Collection.
   3. Replace the drive letter ("Y" in the above example) with the server address, such as \storage.library.illinois.edu\DCC. Omit the drive letter as it will likely refer to a location that is unique to your workstation. **Notes:** Archives location: \storage.library.illinois.edu\UA\UA\.
   4. The complete path to the files should look like this: \storage.library.illinois.edu\DCC\ Processed Preservation Files Ready for Medusa Ingest\Legacy Content\Digital Rare Book Collection.

10. **Staged file location:** Give the path to the staging area location of the files being ingested. These files must be within one of the designated staging areas. Ex.: \storage.library.illinois.edu\UA\MedusaStaging\0801058

11. **File format:** Leave blank

12. **Package profile:** Leave blank

13. **Total file size (GB):** right click topmost folder located in the external store and select properties. Enter the reported size displayed next to the "Size" label.

14. **Total files:** right click topmost folder located in the external store and select properties. Enter the reported number of files displayed next to the "Contains" label.

3. "Rights Metadata" tab
   If unknown, do not fill out information in this tab.
   1. **Rights basis:** Information about restrictions imposed on access to a resource (e.g. copyright, statute, or license).
   2. **Copyright jurisdiction:** The source country of the resource where it is protected by copyright.
   3. **Copyright statement:** Governs how the materials will be used after access has been provided.
   4. **Access restrictions:** A statement that indicates rights related to access and use of the digital resource.

4. When all the information has been entered, select "Create" to save your changes in Medusa.
5.5. Considerations for Archival Units

5.5.1. Relate a Medusa Collection Record to an Archon Collection Record

Specific details on collection management ought to be worked out between repositories and Preservation staff, but in a nutshell:

- In Archon, create a Digital Object record linked to a Collection record for Collection X.
- In Medusa, create a Collection record for the digital content related to Collection X.
- In the Medusa Collection record, paste the Archon series ID (omitting slashes and adding padded zeroes) into the “external ID” field.
- In the Archon Digital Object record’s “ID” field, add the Medusa collection’s “UUID”

5.5.2. University Archives File Package for Electronic Records

Bullets below represent folder names.

- rgsgrrss
  - preservation
  - access
    - online
    - nearline
  - pdi
    - manifests
    - docs

Definitions

- **rgsgrrss.** Nine-digit (numeric) collection identifier from Archon e.g. 1519030 for 15/19/30. Ties collection level descriptive metadata to the archival packet. Supplement with “a” for ALA e.g. 4503044a. For example:

![Folder Structure Example](image)

- preservation. Contains preservation master files. In the case of born digital records, holds disk image or whatever was turned over to the archives. In case of disk image WILL NOT include any original files extracted from the image.
- access. Access folders hold copies of any files put into access system, after they have been weeded and arranged. If derivatives, migrated copies have been generated, derivatives will be given the suffix “~d3riv” before extension and stored in the same folder as the original. Example: Filename_asdf.tif -> Filename_asdf~d3riv.jpg All
migrations will be documented in a processing note on the archon digital object record, at the most specificity possible.

- online. copies of any files put into an access system available via the open web.
- nearline. files available only upon request. Provided to users via zip packet and or on media.

- **pdi.** 'preservation description information' holds technical, structural, and administrative metadata or documents that cannot be accommodated in the digital object fields of archon, or that are specific to individual files
  - manifests. contains manifests for files in preservation and access folders, such as those generated by Karen’s Directory Printer
  - docs. holds any information pertinent to understanding the records such as accession emails, deed of gift, processing plans, etc.
6. Using the Digital Library

6.1. Structured, Free-form, and External collections

The Digital Library System offers three presentation modes for its digital collections:

1. **Structured.** Collections of materials of a similar type, with comprehensive item-level metadata. For example, a collection of digitized photographs of Abraham Lincoln (https://digital.library.illinois.edu/collections/6ff23a90-95b4-0131-1105-0050569601ca-f) or a set of digitized bound volumes of Sanborn Fire Insurance Maps (https://digital.library.illinois.edu/collections/6ff64b00-072d-0130-c5bb-0019b9e633c5-2). These collections are most often produced via digitization in the Library's Digitization Services group.

2. **Free-form.** Collections of born digital electronic records, often of varied content, frequently with little item-level description, presented via a folder/file browse interface rather than an object list. For example, born digital electronic records that the University Archives acquired from the prominent scientist Carl Woese (https://digital.library.illinois.edu/collections/cd535a80-4852-0133-a7fd-0050569601ca-3/tree).

3. **External.** Digital collections of library content that live in other access platforms. These are displayed alongside digital collections managed in the digital library system on our digital collections page (https://digital.library.illinois.edu/collections). Examples include collections of digitized newspapers from our Illinois Digital Newspaper Collections (https://digital.library.illinois.edu/collections/d83bd4b0-b7bb-0133-1d02-0050569601ca-4) and digitized monographs in the HathiTrust Digital Library (https://digital.library.illinois.edu/collections/c49d3ab0-01b9-0130-c5bb-0019b9e633c5-d). Items in external collections are currently not searchable in the Digital Collections platform, but we are exploring this as a possibility.

6.1.1. Packaging for Structured Collections

Structured collections most often result from digitization work done by **Digitization Services**, and are packaged by staff in that group with the assistance of the **Digital Library Technical Coordinator**.

6.1.1.1. Single-item Object

Digital image collections where every item is a discrete image (that is, no multi-page items) utilize the **single-item object** profile. A representative collection is the Collins Collection of Irish Political Cartoons.

- In the Digital Collections: https://digital.library.illinois.edu/collections/810c6d90-e3fb-012f-c5b6-0019b9e633c5-e
- In Medusa: https://medusa.library.illinois.edu/bit_level_file_groups/2466
The single-item package is structured with two folders named “access” and “preservation”. The “preservation” folder contains master files in TIFF format. The “access” folder contains derivatives in JPEG2000 format optimized for quick display, including pan and zoom, and the ability to quickly deliver images in other formats like JPEG and PDF in the Digital Library System. All “preservation” and “access” files must be named exactly the same, except for their “tiff” and “jp2” extensions.

![access]

- 046192752003_064_Box2_Invitation.jp2
- 21853171452003_031_Box9_HomeRule.jp2

![preservation]

- 046192752003_064_Box2_Invitation.tif
- 21853171452003_031_Box9_HomeRule.tif

*NOTE:* file names in “preservation” and “access” folders, with the exception of extensions, must match.

6.1.1.2. Compound Object

Digital image collections where some or all items are composed of multiple images, for example a digitized bound volume with many pages, utilize the compound object profile. A representative collection is the Mapping History at the University of Illinois collection.

- In the Digital Collections: https://digital.library.illinois.edu/collections/2a9ff50-a613-0134-2339-0050569601ca-4
- In Medusa: https://medusa.library.illinois.edu/bit_level_file_groups/2732

The compound-object package is structured with one folder for each object (each folder must have a unique name), which contains folders named “access” and “preservation”. The “preservation” folder contains master files in TIFF format. The “access” folder contains derivatives in JPEG2000 format. All “preservation” and “access” files must be named exactly the same, except for their “tiff” and “jp2” extensions. File naming conventions will be decided by project planning groups, but in most cases we do not require files to follow a specific naming pattern.
NOTE: Folder titles must be unique to the collection; in addition, file names in “preservation” and “access” folders, with the exception of extensions, must match.

In addition to preservation and access folders, compound objects may feature folders to contain a supplementary PDF. This has been utilized in the James Edward Myers Sheet Music Collection (https://digital.library.illinois.edu/collections/692ae4c0-c09b-0134-2371-0050569601ca-6) of digitized sheet music, many of whose pages are complemented by supplementary PDF lyric sheets. In this case, the compound object folder looks the same as above, but there is an additional “supplementary” folder for the PDF. The PDF is not required to have the same name as the other files.
6.1.1.3. Mixed Media

While the Digital Library System is optimized for structured collections in the single item or compound item package, it can accommodate collections of mixed media. The **mixed media** profile and its usage in the system are still under development. However, it has been used for the Sousa Archives Music Instrument Digital Image and 3D model collection (https://digital.library.illinois.edu/collections/fe824900-6c8d-0134-1e34-0050569601ca-b).

Due to the complexity and experimental nature of this package format, we are omitting documentation on it at this point. An overview of the package structure is however available in our wiki: [https://wiki.illinois.edu/wiki/display/LibraryDigitalPreservation/Mixed-Media+Object+package](https://wiki.illinois.edu/wiki/display/LibraryDigitalPreservation/Mixed-Media+Object+package).

6.1.2. Packaging for Free-form Collections

The **free-form** profile offers archivists the ability to provide users with an online experience akin to browsing boxes and folders of content. As such, these collections feature nested folders holding multiple different file types. All the **curator** needs to do is point the digital library to a folder of free-form content in Medusa, and it can be made available for perusal.

A representative collection is the Carl Woese Papers.
- In the Digital Collections: [https://digital.library.illinois.edu/collections/cd535a80-4852-0133-a7fd-0050569601ca-3](https://digital.library.illinois.edu/collections/cd535a80-4852-0133-a7fd-0050569601ca-3), specifically its “tree browse” view at [https://digital.library.illinois.edu/collections/cd535a80-4852-0133-a7fd-0050569601ca-3/tree](https://digital.library.illinois.edu/collections/cd535a80-4852-0133-a7fd-0050569601ca-3/tree).
• In Medusa: https://medusa.library.illinois.edu/bit_level_file_groups/2105

Detailed instructions on setting up a free-form collection will be found below.

6.1.3. Medusa Metadata versus Digital Library System Metadata
Medusa houses collection-level metadata; the Digital Library System houses object- and item-level metadata.

6.1.4. Index Medusa Collections Metadata in the Digital Library System
If you edit collection-level metadata in Medusa and want to see the changes reflected in the Digital Library, follow the steps below.

1. Login to Medusa (medusa.library.illinois.edu), navigate to the Collection record, and edit information in the “Descriptive Metadata” tab (for a field-by-field overview, see 5.3. Add a New Collection in this manual).
2. Login to the Digital Library (http://digital.library.illinois.edu/admin). Keep the Medusa Collection record series open.
a. Click on “Collections” at the top left

b. Next, you may index metadata from Medusa by clicking “Index Collections Metadata from Medusa” at the top right. To check progress of the index, go to Server > Tasks. After a few minutes, you should be able to filter your search to find the newly added collection.

6.2. Common Problems and their Solutions

6.2.1. Change the Package Profile of a Collection

If you chose the wrong package profile for a collection in the Digital Library System and the collection has zero items imported, you can edit the profile in the collection record without a problem. However, if you have already imported files and need to point at a different folder with files packaged according to a different profile, things are a little more complicated. Follow the steps below to solve this problem.

1. Navigate to your collection from https://digital.library.illinois.edu/admin/collections
2. Open your collection in the Digital Library System’s administrative interface, and in Medusa. For example:
   a. https://digital.library.illinois.edu/admin/collections/8d420900-c5b7-0134-237b-0050569601ca-2
   b. https://medusa.library.illinois.edu/collections/1172
3. If descriptive metadata has been added to your objects, discuss this with the Metadata Specialist before proceeding and be sure, from your objects page (for example, https://digital.library.illinois.edu/admin/collections/8d420900-c5b7-0134-237b-0050569601ca-2/items) to export a metadata backup by clicking “Metadata” > “Export as TSV.”

**NOTE:** This situation is to be avoided at all costs; please do not add descriptive metadata to something if you know the package profile is wrong.
4. If the collection is public, notify the collection curator that this collection will be temporarily down for maintenance.

5. In the Medusa record (for example, https://medusa.library.illinois.edu/collections/1172), uncheck the box for “Open collection metadata to public list of collections.” (If this box was not checked, disregard this step).

6. Re-index Medusa collection metadata in the Digital Library system by navigating to https://digital.library.illinois.edu/admin/collections and clicking “Index Collections Metadata from Medusa.” Once this completes, the collection will not be available in the public interface.

7. Purge your Digital Library objects by navigating to the objects page (for example, https://digital.library.illinois.edu/admin/collections/8d420900-c5b7-0134-237b-0050569601ca-2/items) and clicking “Purge items.” Accept the warning that all items will be removed from the Digital Library System. Keep in mind that the files will not be deleted from Medusa; records of them will simply be purged from the Digital Library System.

8. Return to your collection page (for example, https://digital.library.illinois.edu/admin/collections/8d420900-c5b7-0134-237b-0050569601ca-2) and correct the following fields with their new values:
   a. Package profile
   b. Medusa File Group
   c. Medusa CFS Directory
   d. Representative item
   e. Representative image

9. Click on “Objects.” The page should show zero objects.

10. Click “Import.” Use the default option for an initial import. For most structured collections, do not check the box “extract creation date from embedded file metadata.” The import step will initiate an object import based on the Medusa folder and package profile referenced in the collection record. View progress of the import at “Server” > “Tasks” (https://digital.library.illinois.edu/admin/tasks).

11. Before making public, make sure the Metadata Specialist approves of whatever state the descriptive metadata is in.

12. When ready to make the collection public, go to the Medusa record (for example, https://medusa.library.illinois.edu/collections/1172), and check the box for “Open collection metadata to public list of collections.”

13. Re-index Medusa collection metadata in the Digital Library system by navigating to https://digital.library.illinois.edu/admin/collections and clicking “Index Collections Metadata from Medusa.” Once this completes, the collection will become available again in the public interface.
7. The Recommended Project Workflow

Poorly structured projects, or worse, projects whose teams change the order of key steps mid-stream, tend to foment angst and confusion for stakeholders. In our experience, the steps below have proven best when it comes to managing simple, successful digitization efforts across multiple units.

1. **PLANNING** (2-3 weeks)
   a. **Curator** proposes digitization project using [online form.](#) **Digital Collections Project Manager** (Angela Waarala) receives form and assesses feasibility of project with **Curator** and staff in **Conservation** and **Metadata Services**. Once approved, stakeholders will schedule when their staff will be available to do the work necessary to complete the project.

2. **METADATA** (varies based on number of items, at least 1-4 weeks)
   a. **Curator** selects or develops a Metadata Profile for collection with **Metadata Specialist** (Tricia Lampron).
   b. **Curator** authors collection-level metadata in Medusa digital preservation repository.
   c. **Curator** prepares spreadsheet of object-level metadata for items to be digitized.

3. **CONSERVATION** (varies based on Conservation lab commitments and condition of materials)
   a. Depending on the condition of materials, conservation work may be necessary to prepare them for digitization.

4. **DIGITIZATION** (varies based on Digitization lab commitments and number of items, typically weeks to months)
   a. **Curator** provides inventory spreadsheet and items to be digitized to **Digitization Services**.
   b. **Digitization Services** completes digitization, packages files, and deposits them in Medusa. **Metadata Specialist** ingests descriptive metadata into the Digital Library System.

5. **REVIEW** (varies based on number of items and quality of metadata, typically 1-3 weeks)
   a. **Metadata Specialist** and **Curator** review collection and correct errors in metadata and display as needed.

6. **PUBLISH** (1 day)
   a. **Metadata Specialist** makes collection public in the [University of Illinois Library Digital Collections.](#)

The most common variant on the workflow above is using step 2b to create item records in order to hasten getting them into the digitization queue. Typically, the curator intends to go back and enrich their descriptive metadata after digitization. This is perfectly acceptable, but can become chaotic if the curator attempts to enhance metadata while step 4, digitization, is occurring. This is because, at
present, the system and tools Digitization Services uses to monitor digitization queues is different from the system used to manage descriptive metadata. While we hope to improve this in the future so that people can work on metadata while digitization is occurring, for now these are much better understood as separate steps with no overlap.

In cases of variation from the workflow described above, be sure to communicate with the Digital Collections Project Manager to make sure all people involved in group effort from curatorial units, Preservation, Conservation, and Metadata are aware of potential deviations from the plan.

7.1. Planning

Digitization projects require coordination across multiple library units. For this reason, planning is key. The planning process begins when a Curator proposes digitization project using online form. The Digital Collections Project Manager receives the form and assesses project feasibility with the Curator and staff in Conservation and Metadata Services. Once approved, stakeholders will schedule when their staff will be available to do the work necessary to complete the project.

7.2. Metadata

The steps below are primarily for collections of archival materials selected for digitization. Metadata considerations for monographic items with records in Voyager are slightly different. (More information to come; for the time-being, please refer questions to the Digital Collections Project Manager).

7.2.1. Select or Create Metadata Profile

When planning a project, a Curator will, with the assistance of a Metadata Specialist, choose or develop a Metadata Profile for the collection to be digitized. The profile determines available fields for metadata description, and you may think of it as a template for descriptive metadata for each item in your collection.

1. To view existing Metadata Profiles or create new ones, go the the DLS Metadata Profiles page at https://digital.library.illinois.edu/admin/metadata-profiles. Login with University NetID and password if prompted.
2. Optional: To create a new profile:
   a. Click “Add Profile”
      i. In the resulting dialog box, type the name of the Metadata Profile. Standard practice is to name the profile after the collection. Then click “Create Profile.”

      ![Add Profile dialog box](https://digital.library.illinois.edu/admin/metadata-profiles)

      ii. You may now find your newly created metadata profile in the list of alphabetically-arranged Metadata Profiles.

      ![Metadata Profiles list](https://digital.library.illinois.edu/admin/metadata-profiles)

   b. Optional: After creating a new profile, you will need to customize it; otherwise, it will simply contain default elements.
      i. First, select the name of your collection in the alphabetically-arranged list on the metadata profiles page
         [https://digital.library.illinois.edu/admin/metadata-profiles](https://digital.library.illinois.edu/admin/metadata-profiles)
      ii. Use the interface pictured below in order to choose the fields for your profile. The fields you choose will be available for every item in your
collection once you are ready to enter item level metadata later.

![Sample Collection Elements](image)

Explanation of Metadata Profile editing interface:
1. Edit: allows you to change the name of the profile and choose the default element the collection will sort on.
2. Clone: copies profile
3. Delete: deletes profile
4. Add Element: brings up a dialog that allows you to add a new metadata element to the profile
5. Delete Checked: allows for the deletion of any selected element(s) in the profile. To select an element, use the check box in the far left of the row.
6. You may also edit elements in the profile by clicking the “Edit” button to the left of the element’s name. This will bring up a similar dialog as “Add Element.”

7.2.2. Edit Collection-level Metadata

Collection-level metadata is housed in the Medusa collection registry; object- or item-level metadata is managed in the Digital Library system. Edit collection-level metadata using the instructions earlier in this document on editing collection records in Medusa (6.3. Add a new collection).

After editing your collection-level metadata, make sure to index this metadata in the Digital Library System (...7.1.4. Indexing Medusa Collections Metadata in the...) so it appears there.

7.2.3. Create Object-level Metadata

Prior to digitization, the Curator provides a spreadsheet of metadata describing the items that constitute the digital collection. This serves two functions:

1. It provides staff in Digitization Services with identifying information for items to be digitized, ensuring the proper organization of files during and after digitization.
2. It provides a solid basis of descriptive metadata for ingest into the Digital Library System.

To create this spreadsheet, the Curator goes to the collection in DLS under "Metadata" and exports a blank tab-separated-values (TSV) file.

1. Log in with your NetID and AD password at https://digital.library.illinois.edu/admin
   a. Click “COLLECTIONS” in the dashboard.
   b. In the “Filter” box, type the name of your collection. When your collection appears, click it.
   c. Click “Objects” in the top right corner
   d. Click the “Metadata” button. From the drop down menu, click “Export as TSV”
e. Choose “All Items”, then Click “Export”

![Export TSV](image)

f. You now have an empty spreadsheet file containing descriptive metadata values for your collection. This can be edited in Excel, but we recommend editing it in Google Sheets instead, especially if the collection features special characters or diacritics (for example é, ü, ñ), as Excel does not manage these properly.

g. In your spreadsheet, you will notice a number of additional column names for administrative metadata (“uuid”, “parentId”, etc.). Many of these can be disregarded at this step. At this stage, you can even delete unnecessary columns to unclutter your view of metadata to be edited.

h. Add the column headings “Unique Identifier” and “Notes”. These will be provided to staff in **Digitization Services** who digitize the items.

i. Use “Notes” to provide care and handling instructions to the person who will digitize the item.

ii. **Important**: Use “Unique identifier” to provide a unique folder name in which the images will be stored. Creating this now will help technical staff join descriptive metadata to digitized items later in the process. For this step, most **curators** provide a pre-existing identifier for the item. Consult the **Digital Collections Project Manager** for guidance.

i. Fill out your metadata in consultation with the project’s **Metadata Specialist**.
j. When complete, the **Curator** provides the metadata spreadsheet to **Digitization Services**.

*NOTE:* For the digitization of items that have barcodes and Voyager records, much of this metadata can be scanned in in an automated fashion rather than rekeyed. Consult the **Digital Collections Project Manager** for more details.

### 7.3. Conservation

In many cases, conservation actions must be taken on items to stabilize them prior to digitization. For this reason, it is recommended to include a conservator for project planning meetings in order to assess the condition of materials prior to digitization, and to adequately schedule time for conservation actions in concert with other project stakeholders. The Conservation Lab, like the Digitization Services studio, provides services to many units in the library, and both have multiple commitments and workloads to balance. For this reason, please understand that projects generally need to be scheduled in advance to be successful.

### 7.4. Digitization

The **Curator** or curatorial unit staff drop off physical items at DS. Physical objects must have the following:

1. Identifying record series, item titles, boxes, folders, item numbers written on them
2. Barcodes at the object level

Next, **DS Staff** digitize items, keeping files matched to the correct metadata rows as provided, following the steps below.

1. Items are digitized
2. Files are processed for correct formatting
3. Quality control is performed on files
4. Inventory spreadsheet is updated by DS staff as necessary
5. Files are prepared for ingest into preservation and access repositories
6. Original items are picked up by **Curator**

### 7.5. Review

At this stage, metadata has been created for items in a spreadsheet, and the items have been digitized. Now it is time to import image files and descriptive metadata into the digital library, review the results, and make improvements.

#### 7.5.1. Import Digitized Items into Digital Library

The initial import of metadata and image content will be performed by technical staff, most likely the **Metadata Specialist**. The steps are outlined below.
1. Before import of Medusa images occurs, we need to let the Digital Library System know where in Medusa the files are located. To do this, you (in most cases, the project’s Metadata Specialist) navigate to the administrative Collections page, locate the collection to be published, and click the “Edit” button.

2. You’ll also want to have the Medusa collection record open in another tab (this is most easily accessed by going to View > Medusa). Consider keeping 1) the Digital Library collection record and 2) the Medusa collection record open in separate windows, stacked or side-by-side.
3. Make sure that the appropriate Metadata Profile and Medusa Package Profile are selected for the collection’s files.

4. In the Medusa record, navigate to File Group that contains the content to be imported. Click on the “Metadata” button.
5. Copy the UUID number from the Administrative Metadata section (e.g., 26bb0180-71f6-0134-1e7e-0050569601ca-5).

6. Paste the UUID number in the “Medusa File Group UUID” field in the Digital Library record.

7. If you are using the free-form profile, you may wish to point to a File Group folder below the root level. This occurs most frequently with the University Archives, who often point to a
folder at the level of "[root]>access>online." Using the steps outlined above, paste the Medusa UUID for the intended access folder into the Digital Library System field "Medusa CFS Directory UUID."

8. To import files from Medusa, click on the “Objects” at the top right. Then click on “Import” at the top right of the following page.

b. In the pop-up box, make sure “Create” is selected. In most cases, you will want to leave the two boxes below unselected (extract embedded file metadata into DLS metadata; extract creation data from embedded file metadata), as these are used primarily for free-form collections using a profile meant to extract embedded metadata from files (more information to come).
2. Importing files from Medusa to the DLS now begins. In the DLS, CURATOR may now check Ingest progress at https://digital.library.illinois.edu/admin/tasks
3. Once ingest into the DLS is complete, collection objects will show up on Digital Library collections page: https://digital.library.illinois.edu/admin/collections/26bb0180-71f6-0134-1e7e-0050569601ca-5/items
4. Prior to proceeding with editing metadata, the Metadata Specialist combines administrative metadata from the initial import with descriptive metadata created earlier in the process. (This process is undocumented in the manual at present).

7.5.2. Edit Existing Metadata

In the collection in DLS under “metadata,” a Curator may edit descriptive metadata for collection items in one of three ways: using the native web-based metadata tabular editor, using the item editing interface (not yet documented), or exporting the metadata into a spreadsheet for editing and reingest.

7.5.2.1. Edit Metadata in Web Interface

1. Log in with your NetID and password at http://digital.library.Illinois.edu/admin
2. Click “COLLECTIONS” in the dashboard
2. In the “Filter” box, type the name of the collection to be edited. When the collection appears, click the collection title
3. Click “Objects” in the top right corner

![Image 1](https://example.com/image1.png)

4. (Optional) To edit a specific item or set of items, search for them using the facets to the left, or using the ‘filter’ box

![Image 2](https://example.com/image2.png)

5. Once the items have been filtered (or to edit all items in the collection if no filters were chosen), Click the Metadata tab and select “Edit Result Set” in the drop down menu

![Image 3](https://example.com/image3.png)
6. To create or edit metadata, click the desired cell, then start typing.

Note on Multi-item Objects: Multi-item objects contain object level and item level metadata. The object level is represented with an icon or the designated representative thumbnail, and the item level is represented by a thumbnail next to an arrow, as seen below:

7. When metadata has been edited, the box will turn red. Once all edits have been made, click “Save Changes”
8. Once “Save Changes” has been clicked, the metadata will be available for viewing in the digital library (https://digital.library.Illinois.edu)

7.5.2.2. Edit Metadata in a Spreadsheet

1. Log in with your NetID and AD password at https://digital.library.Illinois.edu/admin
   a. Click “COLLECTIONS” in the dashboard

![Dashboard](image1)

   b. In the “Filter” box, type the name of your collection. When your collection appears, click it

![Collections](image2)

   c. Click “Objects” in the top right corner

![Objects](image3)
d. Click the “Metadata” button. From the drop down menu, click “Export as TSV…”

e. Choose “All Items”, then Click “Export”

f. Edit the metadata in a spreadsheet editor, preferably Google Sheets, as this is better at preserving diacritics than Excel.

g. Upload the edited TSV using the same “Metadata” button dialogue used to export the TSV above
7.6. Publish

When a collection is ready to be made public in the Digital Collections, take the following steps:

1. Navigate to the collection record in the Digital Library System’s administrative interface and click “Edit”.
2. Select the appropriate rights statement.
3. Under “Allowed Roles” at the bottom of the form, check boxes to restrict access as needed for the collection
   a. For open access collections, do not check anything
   b. For collections restricted to campus-only use, check “UIUC Campus”. This will provide access to users within the UIUC campus IP range.
c. Once you are ready to make the files public, check the “Published” box at the top left and save the form.

4. In another browser window or tab, go to Medusa, navigate to the collection record, click Edit, check the box “Open metadata to public list of collections.”
5. Return to the Digital Library System’s administrative interface, go to the “Collections” page, and click “Index Collections Metadata from Medusa.” This may take several minutes. If you would like to follow its progress, do so under Server > Tasks in the upper menu.

6. Verify that the collection is accessible in the public interface.
8. Medusa Back-end Management

The instructions below are intended primarily for staff in the Digital Preservation lab. Most curators will not need to do the tasks below, but may find reading the section useful in order to understand the work of the Digital Preservation Graduate Assistant (DP GA).

8.1. Ingest Files into Medusa

Once a file group is created the digital content can be transferred to, or ingested, into Medusa. Most ingest processes are handled through the Preservation Unit unless a particular Producer or curatorial unit has been given permission to ingest content. When files are ready for ingest, the Digital Preservation Coordinator or the DP GA will initiate deposit.

Each Repository has a somewhat different workflow for ingesting materials into Medusa. It’s up to the DP GA to maintain good communication with the various departments about how they would like their Medusa collections to look during and after the ingest process. Each Medusa ingest will have a repository, collection and file group associated with it. The DP GA will need to utilize the search feature in Medusa to best understand which file groups will go with which collections and repositories. Most file groups will already have collections and repositories in place to be ingested. Overall, the basic workflow structure is:

Stage --> Ingest --> Simple Assessment --> Amazon Glacier.

The University Archives, Sousa Archives, and Digitization Services (DS) are the departments that most heavily use Medusa’s ingest features, although others are always welcome to learn and benefit from them.

8.2. Ingest Request Management Overview

Below is a brief guide for managing ingest requests from various library units. Presently, each unit has a different way of notifying the Digital Preservation GA (DP GA) that there is content to ingest and how they manage tasks.

8.2.1. University Archives

\storage.library.illinois.edu\UA\MedusaStaging

**Contact:** Bethany Anderson (bgandrsn@illinois.edu)

**Note:** Archives maintains a workflow spreadsheet: [https://docs.google.com/spreadsheets/d/1j-ivFQlA57Vh61e3evfhCGk9lgbORiQZN5czVFo9zgM/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1j-ivFQlA57Vh61e3evfhCGk9lgbORiQZN5czVFo9zgM/edit?usp=sharing) Ask Bethany to add you as an editor (you may have to use a Google account for this). All UA ingests are located and tracked here. There are tabs for three types of tasks: “Staging” has new content for which you must make a new collection in Medusa, “Accruals to Medusa” has
content being ingested into already existing collections and “Replacements” is for content which you are deleting in Medusa and replacing. Each task also has a “Notes” field with specific instructions for that ingest. You should have a meeting to go over the spreadsheet in more detail with Bethany and the GA in Archives.

8.2.2. DS (Digitization Services)

\storage.library.illinois.edu\DCC\MedusaStaging\n
Contact: Angela Waarala (waarala2@illinois.edu)

Note: DS has started managing their own ingests. Any ingests assigned to the DP GA will be assigned as JIRA tasks (see the DS Digitization Backlog here: https://bugs.library.illinois.edu/secu...view=20029&projectKey=DCC&view=planning.nodetail). The JIRA tasks will contain specific instructions for each ingest. Occasionally the DP GA may be contacted more informally, via email or chat for priority content that needs to be ingested within the day. For more on DS’s workflow see: https://uofi.app.box.com/file/35887488634

8.2.3. Conservation

\storage.library.illinois.edu\ConsUnit\MedusaStaging

Contact: Cher Scheider (schnedr@illinois.edu)

Note: We have recently developed a workflow for Conservation. Due to the large amount of content Conservation staff are moving finished files into the staging area and arranging it by month. There will be a folder for each collection and within that folder will be folders with month and year. What works best is to ingest the previous month’s folders at the beginning of each month. For example, in the first few days in September, ingest all August2017 folders.

8.2.4. Media Preservation Services

\storage.library.illinois.edu\AVPres\MedusaStaging\n
Contact: Joshua Harris (jsharrs@illinois.edu)

Note: Currently, AVPres content is ingested in large batches. The DP GA will get an email detailing specific instructions for the content being ingested. The DP GA may also meet informally to go over any questions or issues. We have found these to be helpful especially with collections that have multiple file groups for different media formats.

8.2.5. Sousa Archives and Center For American Music

\storage.library.illinois.edu\SousaArchives

Contact: Scott Schwartz (schwrtzs@illinois.edu)

Note: As of Summer 2017 talks have started on formalizing the Sousa ingest process. Talks will continue with the Digital Preservation Coordinator and DP GA regarding how notification will be sent when there is content ready for Medusa, as well as continuing discussions on processing prior
to ingest. Note that **Media Preservation Services** ingests content into Medusa on behalf of Sousa for collections they digitize.

### 8.2.6. Newspapers

Currently, there is not a formalized processes for ingesting digitized newspaper content. Previous assistants worked on an ad hoc basis ingesting content from the Illinois Digitized Newspaper Project. In Fall of 2017 the Digital Preservation Coordinator, Digital Reformatting Coordinator (Will Schlaack), and **DP GA** will assess upcoming ingest needs and develop a formalized strategy.

### 8.2.7. Rare Book and Manuscript Library

\storage.library.illinois.edu\RBML

Most Rare Book and Manuscript Library ingest are managed in an automated fashion through Digitization Services. The file share above is utilized for ingests of materials that fall outside of these automated workflows.

### 8.2.8. Other

\storage.library.illinois.edu\Medusa_Staging_2\n
The **DP GA** may be contacted directly about ingesting contents from our purchased data program or for other units/projects such as the digital libraries. These will most likely be assigned via JIRA with meetings or emails as needed for longer term projects.

**Quick guide to communication on ingest requests:**

- **UA** — Google Spreadsheet & email
- **DCC** — Jira Tasks and email occasionally
- **Conservation** — check staging area each month
- **AVPres** - Primarily email with meetings as needed
- **Sousa** — Primarily email
- **Other** — most likely Jira or email

### 8.3. Ingest Content

1. During this process, **make sure files have been copied from the UA server location** (\storage.library.illinois.edu\UA\UA) to the **UA Medusa Staging area** (\storage.library.illinois.edu\UA\MedusaStaging) **in order to complete an automatic ingest**. Entering in file path information can be found in instructions about adding files to Medusa above.
2. Once the files are staged you have two options depending on the collection:
   1. If this is a collection that was just created, after staging the content and creating a file group, click on the gear icon and select “Create bit-level file group.” This will create the file group into which your files will be ingested.
b. **or** if this is an accrual to an already existing bit-level file group and collection, click on the “edit” icon and under “Administrative Metadata” scroll to “Related File Groups” and select the bit-level store you will be ingesting to. This will connect the file group you just created to the bit-level file group you are ingesting the files into.

3. Next, navigate to the bit-level file group (if you are ingesting files into a subfolder inside the main file group, navigate to the exact folder/level that you want to ingest files into. If the files you are ingesting follow the same folder structure as the file group you can simply ingest at the top most level), click on the gear icon and select “Add Files.”

4. A window will pop up displaying six staging areas. Click on the appropriate one (UA) and you will see a list of the folders and files located at UA/MedusaStaging. From here you can select the folder you are ingesting or navigate down to the appropriate folder/files you are
ingesting and select it/them and click ingest. This will start the auto ingest process.

5. If you are working with University Archives materials make sure to update the Google doc as you go along under the “Completed” tab (input date of ingest, and file deletion, whenever that happens).

8.4. Add Events

1. In some cases you may need to document your steps as “Events” within Medusa. When you auto-ingest, the fact that you did so will create an automatic event after the collection has been backed up in Amazon. Every other event in the process will need to be initiated by you. You will need to click on “Actions” and “Add” an event.
2. Some important events you will want to document in the preservation process would be: when you have initially created an external file group, when you've deleted staged files, when you've added an accrual, etc.

8.5. Approve Ingests

Once you finish the Add Files process, you will then need to approve ingest from the Medusa Dashboard.

1. In the Medusa Dashboard click on “Accruals.” This page shows you all the current jobs ingesting to Medusa or backing up to Amazon. The initial “Add Files” action in the file group page should have added the ingest onto this Accruals page.

2. On this page the most relevant thing to note is the “Status” field. Some statuses you will see include:
   a. Checking for existing files/Checking sync – These are visible when the “Add Files” action has occurred very recently. Once checking sync is done you will be able to approve the ingest. No Action Is Required while at checking sync stage. Basically, “Checking for existing files” checks the staging area for the files being ingested and “Checking sync” checks that they synced properly onto the Medusa file server so that they can be ingested. These statuses are usually very brief, though if you just moved files within the specific staging area it might take longer.
b. Awaiting approval – First level of approval. This can be approved by the person who added the files.

c. Awaiting Admin Approval – second level of approval. This must be approved by a Medusa admin, typically the DP GA.

d. Copying files – Files are being copied onto the condo.

e. Amazon backup – This is the final stage of the ingest/backup process. The content has been ingested and is backing up to Amazon Glacier. When the back-up is finished you will get a confirmation email (explained below) and the files will no longer be on the Accruals page list.

8.5.1. Medusa Notification Emails

Digital preservation staff managing ingests will receive a number of email notifications as files go through the ingest process. If everything goes correctly, the DP GA should receive four emails total. The emails are related to the accruals tab on the main dashboard. Examples of these emails can be found here: https://uofi.app.box.com/file/198818630996

1. **Subject: Medusa accrual pending** - This email is triggered when you select Add Files in the file group page and select content to be ingested. The email describes in detail the content being ingested, the staging area and what the file structure will look like after ingest. It also shows you if there are any md5 conflicts (which you will want to look out for). In essence this email informs you that this ingest may now be approved at the accruals tab. The person who added the files should get this email.

2. **Subject: [medusa-admin] Medusa accrual requested** - This email is triggered when you submit the first level of approval for the ingest. It contains similar information to the first email. All Medusa admins get this email. This email notifies you that the ingest has gone through the first approval level in the accruals tab.

3. **Subject: Medusa accrual completed** - This email notifies you that the accrual was completed. It shows you what the accrual directory is (ex. 872/2948, 872 being the collection number and 2948 being the file group number). It also includes the name of the collection and repository that were ingested into.

4. **Subject: Amazon backup progress** - This email notifies you that the back-up to Amazon is complete. It shows you the directory as well as the collection and file group id numbers. (Note: Because this email does not include the collection or repository name, it is important to check the collection and file group numbers against those in the 3rd email to double check that both steps are complete.

*Note:* Emails three and four will arrive at the same time (once the entire process is complete) even though the accrual will finish before the Amazon back up. In the accruals tab the status will change from “copying files” to “amazon back up” but you will not get email 3 until the back-up is complete.
8.6. Add items to Digitization Queue as a Medusa Project

Medusa “Projects” are used to create digitization queues. They are listed at https://medusa.library.illinois.edu/projects and are used primarily by staff in Digitization Services. Every project is linked to a Medusa Collection.

7. Locate your Collection in Medusa by either browsing Collections or Repositories. To search for the collection through “Collections”:
   a. Go to the Medusa website: https://medusa.library.illinois.edu/
   b. Click “Collections”. Login with University NetID and password if prompted
   c. Search for the collection created in Step 1 above by typing the Collection’s name into the filter box.
d. Alternatively, you may find your collection by going to repository

i. search for your repository in the filter box

ii. Click the title of your repository

iii. Click on the title of the collection created in step 1

 e. Click "Add Project"

f. Enter metadata for the project. This metadata describes as a whole this particular batch of items about to be digitized. The following fields are required:

i. Title

ii. Manager email
iii. Owner email  
iv. Start date  
v. Status

vi. **Note:** This screen also features the fields “ingest folder” and “destination folder UUID” related to the automated ingest feature; these are not documented in this manual at present (coming soon).

g. To save and create the project, click “Create” at the top left.

![New Project Screen](image)

h. On clicking “Create,” you will be redirected to a page with an interface for the project:
8. Ideally, a **Curator** has already provided a metadata spreadsheet to the **Metadata Specialist** for use in this step. **Digitization Services** will upload this spreadsheet, possibly after making modifications for administrative purposes, into Medusa to establish a digitization queue.

A comprehensive overview of Project metadata is described below, primarily for the benefit of Digitization Services staff. Technical staff may use any spreadsheet tool for editing the template provided above but Google Sheets is recommended. It allows for smoother collaboration, oversight, and tends to format more cleanly and easily than Excel when converting to CSV. To note:

a. The inventory is used during inventory to identify objects and match them to the correct bibid or identifier

b. The lab cannot digitize and name content with correct file names or ingest it into a repository unless the inventory is clear and correct

c. To see all available fields, please consult the [Medusa Database Inventory Template](#).

The template has columns suitable for a wide range of cataloged and uncataloged materials. Use all columns relevant to your project. The following columns are for items in the Voyager Catalog:

i. Barcode (may be used either for items in or not in the Voyager Catalog)

ii. Call Number (required)

iii. Title (required)

iv. Author (required)

v. Imprint (required)

vi. BibID (required). Used for file naming

vii. OCLC number

viii. Item number

d. The following fields are for archives items not in the Voyager Catalog:

i. Unique Identifier (required). Used for file naming

ii. Record Series ID (required)

iii. Archival Management System URL (required). Taken from Archon

iv. Series

v. Sub-series
vi. Box
vii. Folder
viii. Item title
ix. Creator (required)
x. Date (required). A significant date associated with object
xi. Rights Information
xii. Local title
xiii. File count
xiv. Notes
xv. Batch
xvi. Barcode (may be used either for items in or not in the Voyager Catalog)
e. Note that the rest of the columns are for administrative metadata entered by DS staff
f. Format the inventory spreadsheet barcodes column by doing the following:
   i. Select the barcodes column
   ii. Go to Format Cells
   iii. Go to the Number tab
   iv. Select the Custom category
   v. Select “0” code to format the barcodes correctly for upload
g. Review the directions for saving a spreadsheet in the correct format. For Windows use “Unicode Text.” See the figure below for reference.
9. In Medusa, DS Staff uploads inventory CSV spreadsheet to project.
   a. Login to Medusa
   b. Go to the green “Projects” tab.
   c. Use the “Filter” textbox to locate your respective project and click on it.
d. Click on the upper right hand corner to the “Upload” button [ ]

![Medusa Interface](image)

Add Item  Export  Metadata

e. Check to make sure CSV file meets the specifications described on the upload page. Select your file with the “Browse...” button and click “Upload”

![Medusa Interface](image)

f. You will receive an email confirming your upload including the number of items that successfully uploaded

g. It is wise to confirm your upload on your project page to make sure all of your items appear. To do so, locate your collection on the project page and compare the items there to those on the original CSV you uploaded.

h. Digitization occurs.

10. In Medusa, a **DS STAFF MEMBER** ingests content files.

### 8.7. Delete Content from Medusa

Recently, an option for deleting content through the Medusa platform has been implemented. However, deleting content is generally not done without a compelling reason such as replacing restructured or erroneous content. If you need to delete an entire file group:

1. Click on the Bit-level file group,
2. Next, click on the “edit” icon.
3. When you are on the Edit File Group page, click on the trash icon to delete. A pop-up will ask if you are sure you wish to delete the file group.
4. Click “OK to proceed.”
5. The next page will ask you to document the reason for deletion. Put in any relevant information for why this content is being deleted.
6. Click “Create.”

This process will notify Medusa Managers (currently Preservation Librarian, Digital Preservation Coordinator, Medusa Research Programmer) that content has been requested for deletion. One of them must approve it in order for the content to be deleted. When the deletion is approved by a Medusa Manager with “Medusa File Group deletion approve” as the subject and more information about the deletion in the body of the email. Currently deletions are stored for 30 days in the event that the deletion must be undone.

8.8. Delete Individual Files or Folders from a File Group

There is no way in the Medusa web application to delete files or folders within a file group. To do this, you have to log in to the preservation storage directly and delete any files or folders manually. This option is to be used only in rare cases of need.

8.9. Restructure Content in Medusa Storage

You may need to modify files or folder structures once content has been ingested into Medusa due to content restructuring or error. If the DP GA receives a request to significantly modify content which has already been ingested, accessing the content directly on Production storage may be required. If there is an alternate option to directly accessing Production storage, such as downloading the file group from the collection registry and modifying it on your workstation, and the option does not take an inordinate amount of time to complete, use that option.

It is HIGHLY recommend to create a copy of the folder containing erroneous content from Medusa Production to a working storage area and then replacing the folder in production. This provides you with a working copy in the event that you inadvertently delete content or make a mistake. If you are working directly on production there are little to no safeguards against mistakes. If you erroneously delete data while working on production storage it may not be easily recoverable – if at all.

8.9.1. To edit content on Medusa Production/Primary Storage:

1. Log into storage.library.illinois.edu via Remote Desktop; use your AD credentials for authorization
2. Click YES on the security certificate warning
3. Click on the File Explorer icon on the taskbar.
4. Verify that a mapped drive named library\adsnfs.adstor.illinois.edu\ is indeed mapped. If not, double-click on the NFS-Mount batch file on the computer desktop.
5. Click on the library\adsnfs.adstor.illinois.edu\ mapped drive.
6. Double click on the Medusa directory.
7. Double-click on the Production directory.
8. Double-click on the root directory.
9. Look for the collection number (e.g. 416), double-click on the folder.
10. Double-click on the appropriate file group (e.g. 730)
11. Drill down as you need to for editing.
12. If you are doing more than deleting/replacing a couple files MAKE A WORKING COPY of the content rather than working with content in place. There are no accidental safeguards when working directly on the production server. If you delete content inadvertently, it's not easily recoverable – if at all.
   e.g. copy Y:\medusa\production\root\416\730\4120221 to \working-1\PresUnit\BornDigitalPres\MedusaFileEditing
13. Use the BornDigitalPres\MedusaFileEditing as a working area.
14. Edit as needed.

8.9.2. Once the content has been corrected:
   1. Verify that the restructured/replaced content is indeed correct.
   2. Delete the source files/folder from Production.
   3. Copy the corrected content to the correct production location.
   4. Delete the working copy content from the working storage area

8.9.3. Once you delete the files/folders:
   1. Proceed to the collection page in the Medusa repository and go into the bit level file group from which you deleted content.
   2. Once you are on the File Group page, click on the gear icon and select Run Simple Assessment. Running a Simple Assessment checks the content for this file group in the NCSA ADS storage against the content displayed in Medusa and ensures that Medusa accurately displays what is in the Condo. In other words, it will make Medusa show what is currently in the NCSA ADS storage (i.e. the files you deleted will no longer show in Medusa).
   3. Manually create an event indicating that content has been replaced\deleted manually from the file group (note 9/7/17: there currently isn't an option for this event; I've requested that Howard add one – tmp)
8.10. Notes on Deleting or Moving/renameing/restructuring Files and Folders

Whenever moving files, make sure you use the file copying program TeraCopy to do so. If you do not seem to have TeraCopy installed on your computer please contact the Digital Preservation Coordinator.

You may encounter permissions issues and will not be able to move or delete files/folders on the NCSA ADS. If this happens, contact the Medusa Research Programmer (Howard Ding) via a JIRA ticket or email and let him know exactly what you need deleted and or moved or renamed.

8.11. Useful Tools

Teracopy
https://wiki.cites.illinois.edu/wiki/display/LibraryDigitalPreservation/Teracopy

Treesize Pro
https://wiki.cites.illinois.edu/wiki/display/LibraryDigitalPreservation/Treesize

OTRS
http://www.library.illinois.edu/it/helpdesk/service/OTRS

Jira
Use this platform to add a description of tweaks that need to be added to the Medusa interface.
You must have permission to access Jira. If you do not please contact Tracy Popp (tpopp2@illinois.edu) and she will contact the Jira administrator:
https://bugs.library.illinois.edu/browse/MED/?selectedTab=com.atlassian.jira.jira-projects-plugin:issues-panel