

UCHICAGO AND THE DIGITAL ARCHIVE ***A QUEST FOR DIGITAL CURATION WORKFLOWS***

ASHLEY GOSSELAR AND EMILY SCHATZ

DIGITAL PRESERVATION COALITION

DIGITAL PRESERVATION WORKFLOW WEBINAR SERIES

FEBRUARY 6, 2025



How do you develop workflows?

Methods for developing born-digital accessioning and processing workflows that are:

1. Based on best practices
2. Suited to your local context
3. Well-documented

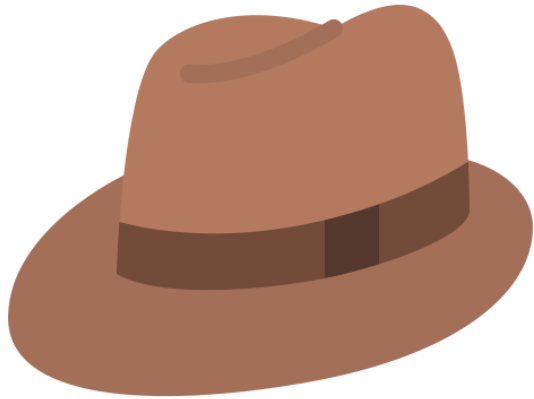




Photo of Regenstein Library by Tom Rossiter

Hanna Holborn Gray Special Collections Research Center

Rare books, archives, and manuscripts;
~85 TB of born-digital content

↑
Head of
Digital Special
Collections



↓
Digital
Accessions
Specialist

Digital Library Development Center

Digital system design, development, and
maintenance



A QUEST FOR ~~THE HOLY GRAIL~~

A FUNCTIONAL DIGITAL CURATION INFRASTRUCTURE



Paramount Pictures



"THERE IS MORE IN THE ~~DIARY~~ THAN JUST THE MAP."



project management plan

Three-Year Master Plan

- Guiding principles
- Work breakdown structure

Asana dashboard

- Tracks tasks (to do, in progress, on hold, complete, canceled)

Governance Document

- Roles and responsibilities across the organization
- Resources required to maintain, stabilize, and grow



UChicago
Library

A TREACHEROUS JOURNEY...



CHALLENGE 1: "ONLY THE PENITENT ~~MAN~~ WILL PASS."

^
archivist



CHALLENGE 1: "ONLY THE PENITENT ~~MAN~~ WILL PASS."

^
archivist

Learn from the past

- Review existing documentation
- Interview colleagues with institutional memory

Learn from others

- Read and watch freely available digital curation resources
- Ask colleagues at other institutions if they will share



CHALLENGE 1: "ONLY THE PENITENT MAN WILL PASS."

archivist

Whiteboard

Digital Media Stream - Transfer Protocols

AG

Share

Yellow = Baseline step. We should always do this.

Orange = A step we might do. There is an additional decision to be made.

Green = Office supplies

Pink = Computer equipment

Blue = Software

The workflow steps are as follows:

- Yellow (Baseline):** Separation sheet; Move media to temporary housing/location; Track location of unaccessioned media; Identify and appraise out any media that does not need to be copied; Quarantine files on transfer media for 30 days from date of creation; Assign each piece of media a unique ID; Create folder structure/landing place for files and accompanying documentation in temporary preservation storage; Capture physical label as metadata; Safely copy all digital files to temporary preservation storage; Virus scan; Verify the integrity of the copied files and generate baseline fixity information; Initial inventory of all files received.
- Orange (Optional):** Photograph media with its labels/metadata; Portable write-blocker; Non-networked computer?; Ethernet cable; Floppy disk drives - 3.5" and 5.25"; CD/DVD drives; Zu Disk Drive; SD Card Reader; Disk Image; Logical copy; Teracopy; Exact Audio Copy; CCA Tools (Bundled with BitCurator).
- Green (Office supplies):** Anti-static bags? For drives?; Barcodes? Sharpies? Envelopes?
- Pink (Computer equipment):** ClamAV (can also be a plugin to BitCurator and Archivematica); Microsoft Defender Antivirus; Sophos Antivirus; HashMyFiles; Karen's Directory Printer; BitCurator; DROID; Archivematica; JHOVE/JHOVE2.
- Blue (Software):** ArchivesSpace; Spreadsheet; BitCurator; ClamAV (also part of BitCurator); Mac Disk Utility; KryoPlus (Floppy Disk); Drivers (CD-ROMs and DVDs) in Indiana workflow; Exact Audio Copy; Guymager (Bundled into BitCurator); FTK; Censilla; Teracopy; Exact Audio Copy; CCA Tools (Bundled with BitCurator).

CHALLENGE 2: "ONLY IN THE FOOTSTEPS OF ~~GOD~~ WILL HE PROCEED."

^
best practices



CHALLENGE 2: "ONLY IN THE FOOTSTEPS OF ~~GOD~~ WILL HE PROCEED."

best practices



OSSArcFlow
Guide to Documenting Born-Digital Archival Workflows

AUTHORS
Alexandra Chassanoff and Colin Post

EDITORS
Katherine Skinner (Lead Editor), Jessica Farrell, Brandon Locke, Caitlin Perry (Copyeditor), Kari Smith, Hannah Wang

CONTRIBUTORS
Christopher A. Lee, Sam Meister, Jessica Meyerson, Andrew Rabkin, Yinglong Zhang

DESIGNER
Hannah Ballard

EDUCOPIA INSTITUTE
Community Cultivators

Pre-ingest preparation

- Appraisal
- Deed of Gift



Ingest

- Safely transfer to secure server

Accession + Process

- Arrangement
- Description
- Restriction review
- Migration
- AIP and DIP



CHALLENGE 2: "ONLY IN THE FOOTSTEPS OF ~~GOD~~ WILL HE PROCEED."



best practices



Describing Workflows

Phase	Conditions	Step	Description	Software	Hardware
Pre-accessioning		Action carried out	[Staff role] does this action, described in more detail		This hardware used
Accessioning	If this is the case, then...	Action carried out	[Staff role] does this action, described in more detail	This software used	

Phase	Conditions	Step	Person Responsible	Description	Office or archival supplies	Software	Hardware	Result / Deliverable?
Pre-Ingest	If digital media is found in an analog collection	Items prepped for transfer to DAS/backlog. If needed, items put in new housing. Separation sheet is filled out.	Accessioning student / archivist / coordinator	Optical media without cases placed in archival plastic cases. Separation sheet recording collection name and accession number, box/folder in which digital media was found, number and type of removable digital media. One copy of separation sheet is left in the original physical folder/box location. Another copy of the separation sheet and the digital media are placed in new folder(s) for transfer to DAS/backlog. New folder(s) labeled with the collection name, accession number (if applicable), and originating box number.	Archival folders, CD cases	Microsoft Word	Computer	All items safely stored. Filled out separation sheet in originating box/folder of analog collection, and a filled out separation sheet placed alongside digital media in new folder for transfer to DAS/backlog.

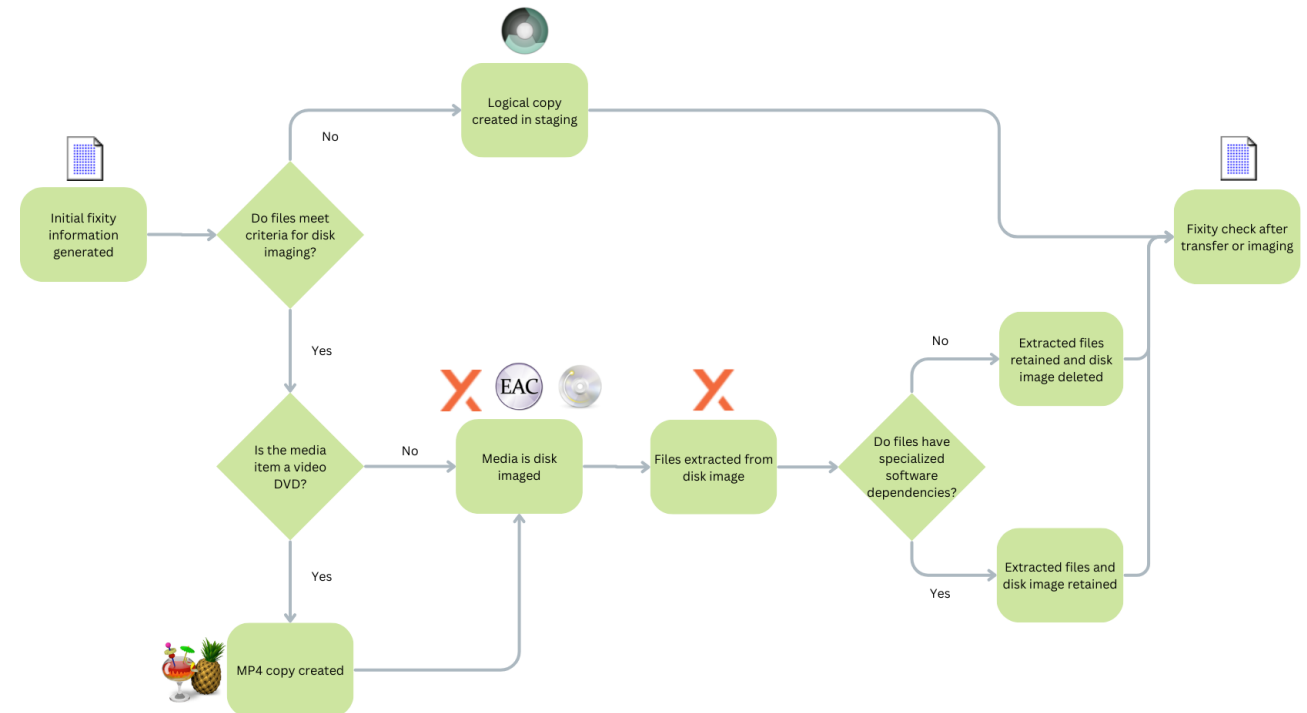
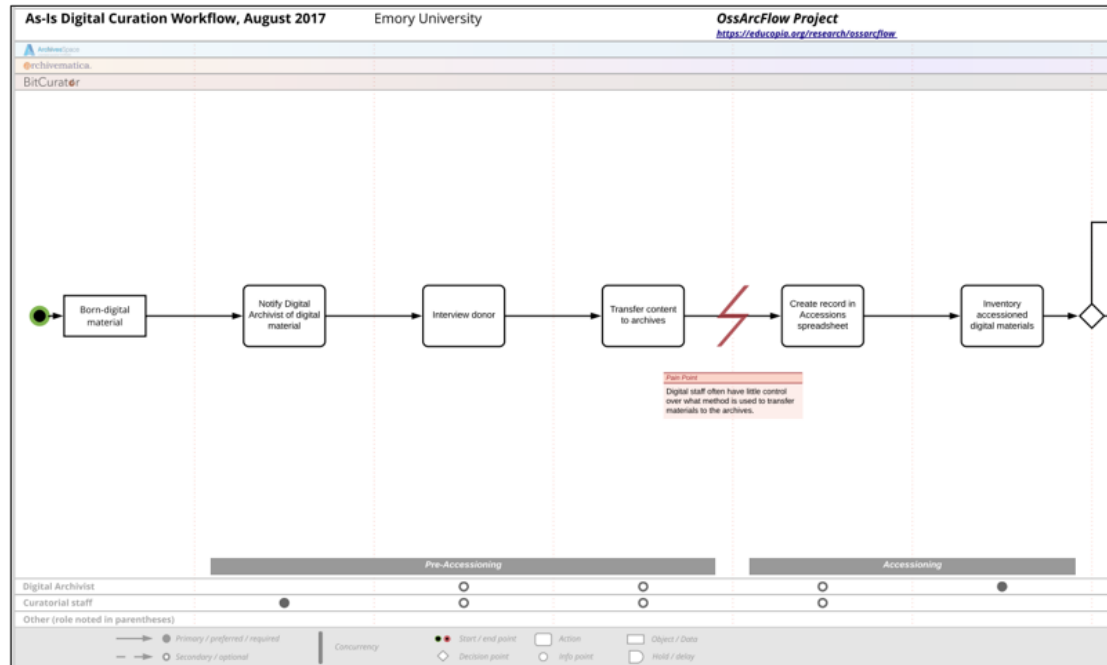
PROCEED IN THE FOOTSTEPS OF THE ~~WORD~~



Guide to Documenting Born-Digital Workflows



Representing Workflows Visually





UChicago
Library

STAKEHOLDER REVIEW



My God, I must be blind.

Assessing Workflows

Functional Area	Level			
	Level 1 (Know your content)	Level 2 (Protect your content)	Level 3 (Monitor your content)	Level 4 (Sustain your content)
Storage	Have two complete copies in separate locations Document all storage media where content is stored Put content into stable storage	Have three complete copies with at least one copy in a separate geographic location Document storage and storage media indicating the resources and dependencies they require to function	Have at least one copy in a geographic location with a different disaster threat than the other copies Have at least one copy on a different storage media type Track the obsolescence of storage and media	Have at least three copies in geographic locations, each with a different disaster threat Maximize storage diversification to avoid single points of failure Have a plan and execute actions to address obsolescence of storage hardware, software, and media
Integrity	Verify integrity information if it has been provided with the content Generate integrity information if not provided with the content Virus check all content; isolate content for quarantine as needed	Verify integrity information when moving or copying content Use write-blockers when working with original media Back up integrity information and store copy in a separate location from the content	Verify integrity information of content at fixed intervals Document integrity information verification processes and outcomes Perform audit of integrity information on demand	Verify integrity information in response to specific events or activities Replace or repair corrupted content as necessary
Control	Determine the human and software agents that should be authorized to read, write, move, and delete content	Document the human and software agents authorized to read, write, move, and delete content and apply these	Maintain logs and identify the human and software agents that performed actions on content	Perform periodic review of actions/access logs
Metadata	Create inventory of content, also documenting current storage locations Backup inventory and store at least one copy separately from content	Store enough metadata to know what the content is (this might include some combination of administrative, technical, descriptive, preservation, and structural)	Determine what metadata standards to apply Find and fill gaps in your metadata to meet those standards	Record preservation actions associated with content and when those actions occur Implement metadata standards chosen
Content	Document file formats and other essential content characteristics including how and when these were identified	Verify file formats and other essential content characteristics Build relationships with content creators to encourage sustainable file choices	Monitor for obsolescence, and changes in technologies on which content is dependent	Perform migrations, normalizations, emulation, and similar activities that ensure content can be accessed

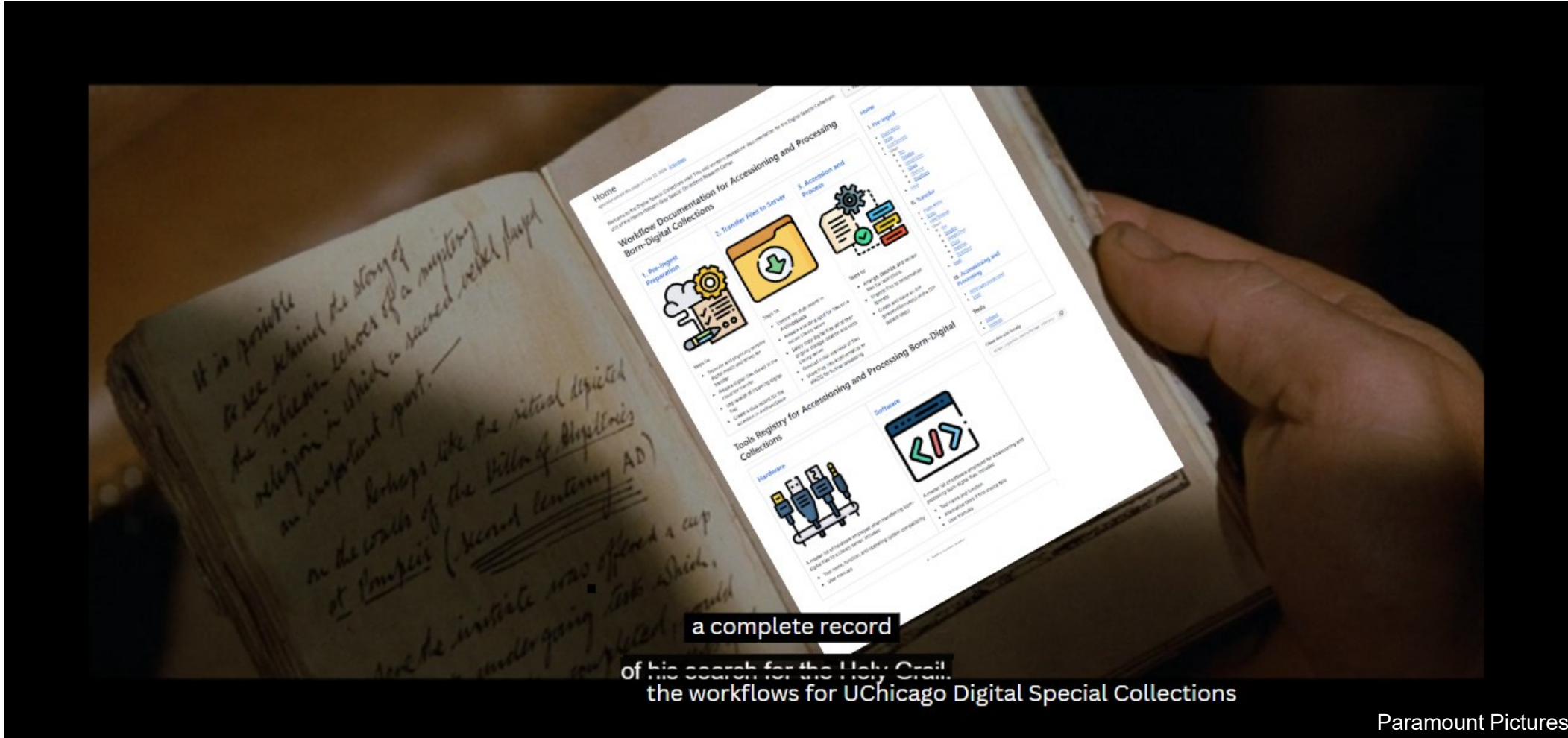
NDSA Levels of Digital Preservation Matrix

1. **Availability** for long-term use
 - a. Threat areas: pre-repository care, selection policies, rights management
2. **Identity**, or reference information, allows digital objects to be discovered and retrieved
 - a. Threat areas: pre-repository care, generation & maintenance of descriptive & structural metadata
3. **Persistence** of the bit sequences, making a digital object usable and retrievable
 - a. Threat areas: physical media management, media refreshment policy, hardware migration policy, & data security policy
4. **Renderability** of a digital object, which can be used in a way that preserves its significant characteristics
 - a. Threat areas: format management workflows & policies, preservation strategies, & repository knowledge of stakeholder community
5. **Understandability** of digital content to members of the designated community
 - a. Threat areas: repository knowledge of the designated community, metadata capture, & retention policies
6. **Authenticity** of a digital object - that it is what it purports to be
 - a. Threat areas: metadata collection & management practices, security procedures, & workflow documentation procedures & policies

In what ways does your repository ensure....	Notes on...	What policies are currently in place that address...	What policies are needed to address...
Availability			
Identity			
Persistence			
Renderability			
Understandability			
Authenticity			

Simple Property-Oriented Threat model for Risk Assessment (SPOT)

The Grail Diary Manual



Paramount Pictures

GitHub Wiki


Home

agosselar edited this page on Nov 22, 2024 · 9 revisions

Welcome to the Digital Special Collections wiki! This wiki contains procedural documentation for the Digital Special Collections unit of the Hanna Holbom Gray Special Collections Research Center.

Workflow Documentation for Accessioning and Processing Born-Digital Collections


1. Pre-Ingest Preparation



Steps to:

- Separate and physically prepare digital media and drives for transfer
- Prepare digital files stored in the cloud for transfer
- Log receipt of incoming digital files
- Create a stub record for the accession in ArchivesSpace


2. Transfer Files to Server



Steps to:

- Update the stub record in ArchivesSpace
- Prepare a landing spot for files on a secure Library server
- Safely copy digital files off of their original storage location and onto Library server
- Conduct initial appraisal of files
- Move files into Archivematica or ePADD for further processing

3. Accession and Process

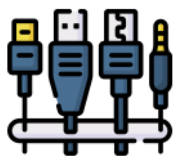


Steps to:

- Arrange, describe, and review files for restrictions
- Migrate files to preservation formats
- Create and store an AIP (preservation copy) and a DIP (access copy)

Tools Registry for Accessioning and Processing Born-Digital Collections


Hardware



A master list of hardware employed when transferring born-digital files to a Library server. Includes:

- Tool name, function, and operating system compatibility
- User manuals

Software



A master list of software employed for accessioning and processing born-digital files. Includes:

- Tool name and function
- Alternative tools if first choice fails
- User manuals

Digital Media

Emily Scharzt edited this page 2 weeks ago · 55 revisions

Overview and Visualization

Full visualization

I. Pre-Ingest



Separate and physically prepare items for transfer

- Before materials are brought to a designated drop-off area for the Digital Accessions Specialist, they are prepared by the Accessions Assistants or another member of Archives staff. Detailed steps for separating, preparing, and transferring digital media to the Digital Accessions Specialist are on Box: <https://uchicago.app.box.com/file/1566764130866>.
- Items such as videotapes, audiocassette tapes, reel-to-reel tapes, vinyl albums, and films should stay with the physical collection. If these are mistakenly put in a drop-off area, they should be returned to the Archives and Manuscripts Accessions Supervisor.

Log receipt of items

- Add an entry to the SCRC accession log in order to claim or note an accession number for the accession.
 - Entries for the digital media or born-digital portions of hybrid accessions should be recorded in the row immediately below the entry for the analog portion of accession. The following fields should be used in the accession log:
 - Accession number. (Entries for the digital media or born-digital portions of hybrid accessions should use the same accession number as the analog portion of the accession. These accessions no longer have -dig or -digital appended to the accession numbers.)
 - Date received (if logging digital media separated from an analog collection was received).
 - Date accessioned
 - FY accessioned
 - Staff member assigned
 - Type of collection
 - Personal papers / records / collection
 - Location (either "In Digital Accessioning" or "LDR")
 - Restrictions

Create a stub accession record in ArchivesSpace

- Create a stub accession record in ArchivesSpace. The stub record should include the following:
 - Collection title
 - Accession number
 - Accession date
 - Inventory list of media items
 - Location of physical media
 - Conditions governing rights and access
 - Provenance information that is known at the time materials are received

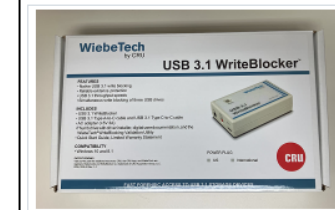
Weed items that are out of scope

- Working with the University Archivist and the Head of Digital Special Collections, appraise out any items that are

WiebeTech USB 3.1 WriteBlocker

agosselar edited this page on Dec 17, 2024 · 8 revisions

WiebeTech USB 3.1 WriteBlocker



This WriteBlocker is for devices that need a [USB A](#) or [USB C](#) connection. This includes:

- Flash drives
- Blu-ray players
- Floppy disk drives
- Zip disk drives

ArchivesSpace

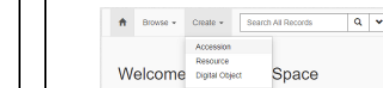
agosselar edited this page on Dec 20, 2024 · 19 revisions

Overview

Accession Records

Creating Accession Records

- To start a new accession, click on the Create drop down menu in the top left corner of the main screen and select Accession.



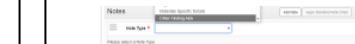
- Refer to the instructions on the workflow pages to determine which fields should be filled out and refer to the [Description Guidelines](#) for more specifics about how each field should be filled out.
 - Please note: ArchivesSpace will not allow you to save the record until you have at least entered an identifier (the accession number).
- Remember to save the accession record by clicking on the blue save button when you are done; ArchivesSpace does not autosave records.

Update Accession Record Status

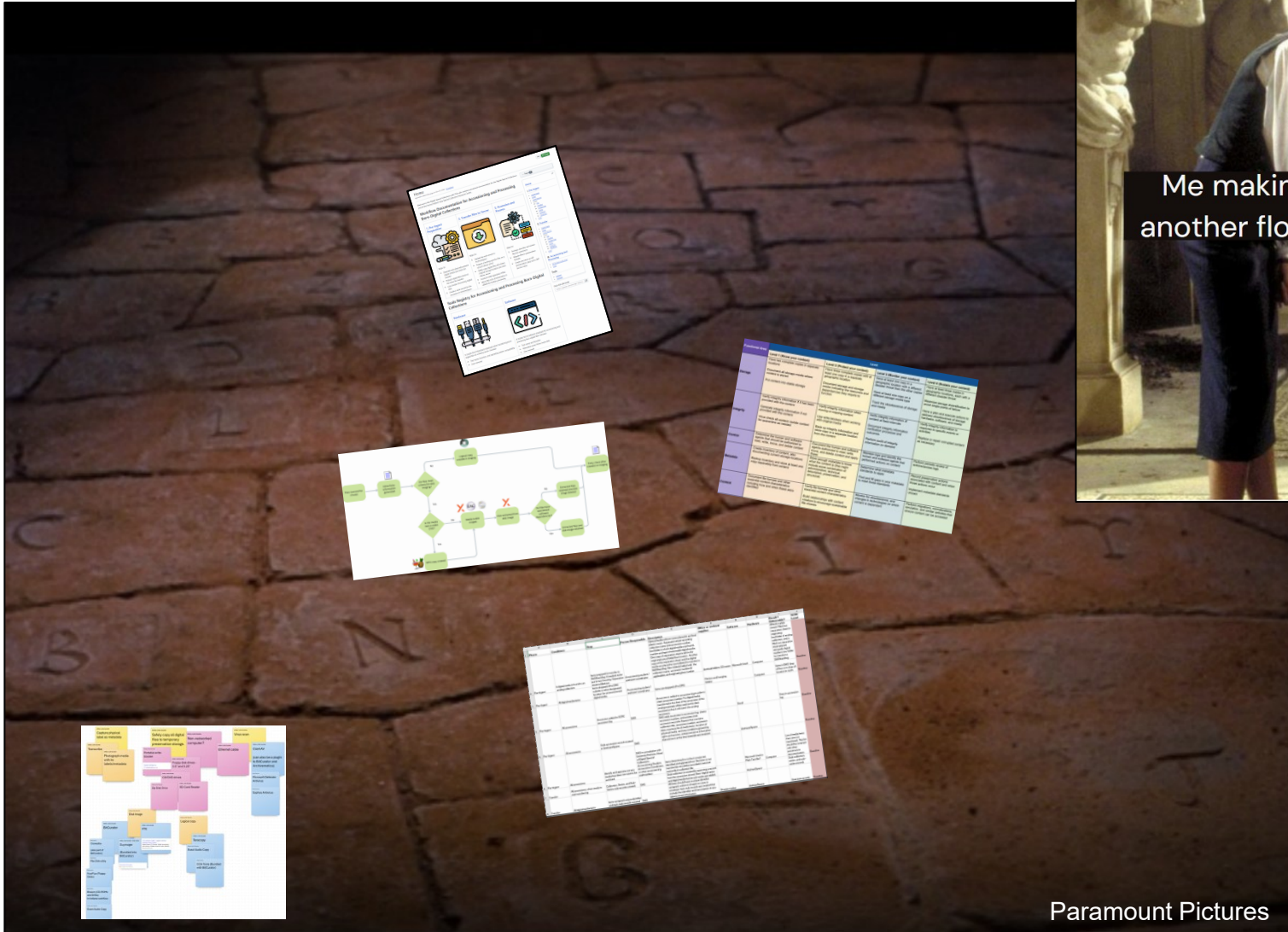
- Open up the Accession you want to update.
- Click on the blue Edit button in the top left.
- Scroll down to the Collection Management section.
- If you are starting a new accession record you will need to click on the Add Collection Management Fields button.



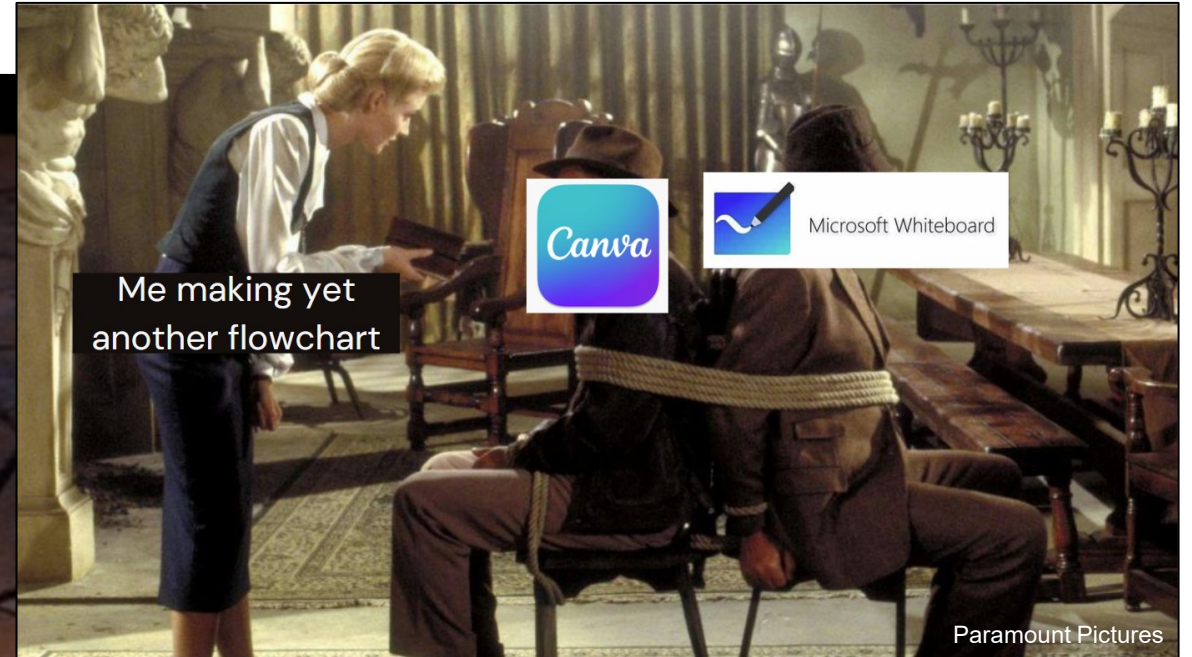
- In the drop-down menu next to Processing Status select the correct status for the accession. Refer to the [Description Guidelines](#) for which status is appropriate.



- Remember to click the blue Save Accession button at the bottom of the page when you have finished editing the record.



Paramount Pictures



**CHALLENGE 3: "ONLY IN THE LEAP FROM ~~THE LION'S HEAD~~
WILL HE PROVE ~~HIS~~ WORTH."**

^
the workflow's

^
theory to practice



Paramount Pictures

Lessons Learned

**“Remember what the
Grail Tablet said.”**

Learn from others.

Tap into community.

Do not reinvent the wheel.

**“But in the Latin
alphabet, ‘Jehovah’
begins with an ‘I.’”**

Get comfortable with
iteration.

“He chose...poorly.”

Choose tools that are
appropriate for your needs
and resources.



Lessons Learned

“Indiana, let it go.”

Don't let perfection get in the way. “Good enough” is okay.



THANK YOU.

Ashley Gosselar
ashleylocke@uchicago.edu

Emily Schartz
ejschartz@uchicago.edu

Resources list

<http://bit.ly/40OQAf4>

