

# Digital Preservation Webinar Series: Identify



# Introductions

- **Sarah Wegley**  
Digital Metadata and Web Services  
Governors State University  
University Park, IL
  - **Julie Mosbo**  
Preservation Librarian  
Texas A&M University  
College Station, TX
  - **Paul Go**  
Systems Librarian/Library Technology Manager  
Illinois Institute of Technology  
Chicago, IL
- Library of Congress  
Digital Preservation Outreach and Education (DPOE) Program



# DPOE's Mission

“The mission of the Digital Preservation Outreach and Education (DPOE) program of the Library of Congress is to **foster** national outreach and education to **encourage individuals and organizations to actively preserve their digital content**, building on a collaborative network of instructors, contributors, and institutional partners.”

# Modules

**Identify - what digital content do you have?**

**Select - what portion of that content will be preserved?**

Store - what issues are there for long term storage?

Protect - what steps are needed to protect your digital content?

Manage - what provisions are needed for long-term management?

Provide - what considerations are there for long-term access?

# Objectives

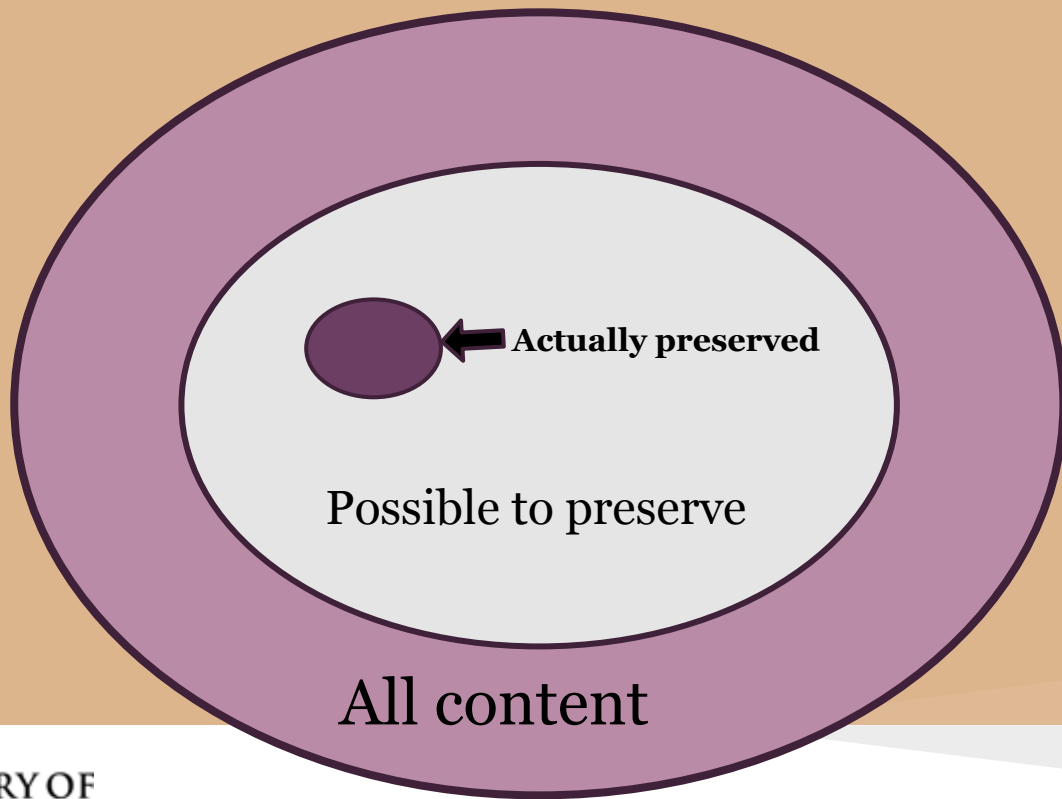
- Learn why an inventory is necessary.
- Learn how to do an inventory for digital preservation.
- Review sample inventories.
- Discuss inventory considerations.
- Prepare for the next stage in the digital preservation process - SELECT.

*Why do we  
identify  
content?*



*Do you have an inventory?*

# *How will an inventory help?*



# *Key Takeaway*

An explicit inventory  
is the best way  
to identify content.





# Creating the Inventory Sheet

- Category
- Title and Description
- Dates
- Location
- Extent
- Format

# Sample Inventory Template

<b>Category</b>	<b>Title and Description</b>	<b>Date</b>	<b>Location</b>	<b>Extent</b>	<b>Format</b>

# GSU Library

<b>Category</b>	<b>Title and Description</b>	<b>Date</b>	<b>Location</b>	<b>Extent</b>	<b>Format</b>
University Archives Digitization Project	The Creative Woman - a feminist journal created at GSU	1977-1992 2012 digitized externally	Internet Archive, GSU Digital Collections - CARLI, Archive DVD	48 journal issues	PDF
University Archives Digitization Project	Campus Construction - photos and letters from the architect concerning the project's progress.	1972-1974	GSU Digital Collections - CARLI, Archive DVD	106 photos, 27 letters	PDF, jpg
Born Digital	GSU View - university newsletter	2007-current	GSUVault, Archive hard drive	76 documents as of 11/2013	PDF versions of the original webpages

# Extent - example from GSU Library

	A	B	C	D	E
1	<b>Size of collections in CARLI Digital Collections as of January 8, 2014</b>				
2	<b>Collection Alias</b>	<b>Size (K=Kilobytes, M=Megabytes, G=Gigabytes)</b>			
3					
4	gsu_const	777M			
5	gsu_ethno	202M			
6	gsu_histdoc	208M			
7	gsu_pres	85M			
8	gsu_sculpt	129M			
9	gsu_tcw	155M			
10	gsu_wee	1.4G			
11					
12					

CARLI provides statistics for Digital Collections at: <http://collections.carli.illinois.edu/carlistats/>

# Texas A&M Library

<b>Category</b>	<b>Title and Description</b>	<b>Date</b>	<b>Location</b>	<b>Extent</b>	<b>Format</b>
HathiTrust/ Online Journal Systems (OJS)	OMG Oil Mill Gazetteer	1908-20 13	IRODS, DiSC_Upload	119 Vols.	TIFF PDF
Cushing Memorial Library & Archives	Cervanets	1605-20 09	IRODS, DiSC_Upload DiSC_Work	163 Titles, 489 individual pages	TIFF JPEG
OAKTrust Institutional Repository	INA Wachsmann	1970s-19 80s	IRODS, DiSC_Upload DiSC_Work	14 binders 6,479 slides	TIFF JPEG 2000

# Texas A&M Library

Size of Collections in IRODS as of February 21, 2014					
Collection Alias	Size (K=Kilobytes, M=Megabytes, G=Gigabytes)				
IRODS_Wachsmann_binder 1	4.26 G				
IRODS_OMG	638.21 G				
IRODS_TAMU University Press Books	456 M				



The inventory's **content** is more important than style and format.

*Keep this in mind throughout the project.*

# Inventory Considerations

- Scalable
- Current
- Available
- Documented
- Serviceable



# Scalable

- *A living document*
- Add rows *and columns* as new digital content is discovered and new information about that content is pertinent.
- For example, adding a new column to track content that is required to be preserved at a certain level if that becomes required.

# Current

- Old inventories are unhelpful.
- Keeping the inventory current helps make sure it is useful.
- Revisiting the inventory on a regular basis can ensure this.

# Available

- Inventory should be on a shared directory or cloud service (such as Google Docs).
- Never keep as a single copy on one person's computer.

# Documented

- A documented inventory is able to be referenced later.
- A *living document* should reflect current inventory.
- A documented inventory can be shown to administrators and used for budgeting.

# Serviceable

- Make sure the document is useful both as an inventory and a starting off point for other projects.
- An inventory can lead to other projects.

# *Key Takeaway*

Following the SCADS model will help you identify your scads of digital objects.



# Sources of digital content

- Scanned photos or documents
- Digital cameras
- Hard drives
- Thumb drives/flash drives
- Floppy disks/zip drives/other media
- CDs/DVDs/video tapes
- Shared network drives
- Digital Collections - collaborative sites hosted on external servers
- Institutional repositories
- Records management tools

# Planning Your Inventory Process

- Who will be participating?
- What content and level of detail will you be documenting?
- Where will you store the inventory?
- When is it due?
- How will you know when you are done?



# Factors for choosing the level of detail

- Resources available
- Extent of content
- Content's nature and location
- Timeframe for completion

# *Key Takeaway*

The purpose of the inventory is to assist in later modules.

Aim for breadth of content

over depth of detail.



# Identification Results

Identify potential digital content you may need to preserve.

Treat the inventory as a management tool that grows as your preservation program grows.

Identify gaps in the digital content to inform your acquisition process.

Use the inventory as a basis for the next step in the digital preservation process: **SELECT**

# Questions?

- **To learn more about the DPOE curriculum**  
<http://www.digitalpreservation.gov/education>
- **To contact the Library of Congress about digital preservation**  
<http://www.digitalpreservation.gov/contact>