MAC midwest archives conference

Moving Archives Forward: The 2022 Midwest Archives Conference Annual Meeting

Session 302 Forward Progress: Diverse Digital Preservation Perspectives Perspectives Perspectives

Forward Progress: Diverse Digital Preservation Perspectives

- Amy Moorman
 Archivist
 - **Avila University**
- Dan Noonan

Digital Preservation Librarian The Ohio State University @dannynoonan1962 "Digital Preservation is about making the best use of your resources to mitigate the most pressing preservation threats and risks."

Guiding Digital Preservation Axiom #9 - Trevor Owens, *The Theory and Craft of Digital Preservation* ©2018









TIMELINE

Avila – 1916

CSJ Federation – 1966

WRSC - 1997

MSARC - 2014



Sisters of St. Joseph of Carondelet, 1956



Peace March at CSJ Federation Assembly, 1986



Coretta Scott King on Avila Campus, 1978





DIGITAL PRESERVATION

"Digital preservation is best thought of as an incremental, ongoing, and ever-shifting set of actions, reactions, workflows, and policies. An iterative approach means that practitioners don't have to start by creating or selecting a comprehensive solution and making hard and fast technology choices to be used for the next 20 years. They can start by taking small steps to prioritize and triage digital collections, while working to build awareness and advocate for resources."

- Schumacher, Jaime, Lynne M. Thomas, Drew Vande Creek, et al, From Theory to Action: "Good Enough" Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions (August 2014) 5.





STARTING POINT

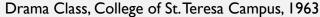
Present Advantages

- Dedicated archives server
- High-Profile digital collections
- Administrator initiatives

Existing Challenges

- No recurring budget
- Executive staff turnover
- Lack of technical support
- Unsecured, disorganized file storage









DEVELOPING A PLAN

Phase I

Surveying existing digital materials

New directory structure

Server security & setup

Updating policies & procedures



Leonard Nimoy visits Avila College, 1978



DISORDER OF HUMAN RECORDS, SLOWLY DECAYING WITH THE CEASELESS MARCH OF TIME ME WITH MY SPREADSHEETS

Source: @archivistmemes

INVENTORIES

ArchivesSpace record

Physic	cal Location	
	Persistent ID	a3f5d594a51f923f24241bab3ef8c49b
	Туре	Physical Location
	Publish?	False
Content		
X ICSJ	00002 CDCM Project C	collection (digital files)/004 - May 2016 Session 1/001 - Curriculum/Articles/Bio Dance pdf
Raw	Formatted	
Scone	and Contents	
Scope	and Contents	\$ eb22e17/leed927701230b49fd5628a14
Scope		eb22e17feed92f701230b49fd5628a14
Scope	Persistent ID	eb22e17feed92f701230b49fd5628a14 Scope and Contents
Scope Sub No	Persistent ID Type Publish?	eb22e17feed92f701230b49fd5628a14 Scope and Contents
	Persistent ID Type Publish?	eb22e17feed927701230b49fd5628a14 Scope and Confents
Sub No	Persistent ID Type Publish? tes	eb22e17feed92f701230b49fd5628a14 Scope and Confents

Quantity Inventory (physical carriers)

Clipboard	- 5	Fo	nt	G .	Alignment	6	Numb	er 5			Styles			Cells
50 * : × \(\sqrt{f_R} \)														
А	В	С	D	E	F	G	н	1	J	K	L	М	N	0
3.00	CD (files)	DVD (files)	HDD (files)	USB drive (files)	Floppy/zip disc (files)	CD (a/v)	DVD (a/v)	VHS (a/v)	Beta (a/v)	Umatic (a/v)	DigVid (a/v)	Cassette (a/v)	1/4" (a/v)	Server (preservation)
UA.009.00001														129 MB
UA.009.00004						3								103 KB
UA.009.00005														97.1 MB
UA.009.00006						3								537 MB
UA.010.00002						1								34.7 MB
UA.010.00003														10.1 GB
UA.011.00001														313 KB
CSJ.00001	14			1	76	105	44	119	31	24	1	1243	11	179 GB
CSJ.00002					10.00									207 GB
WR.00001						2	6	21				31		0
WR.00004				1										4.50 MB
WR.00005	21		1				30	119	12		80	3	2	1.00 TB
WR.00006														590 MB
WR.00009												59		5.84 GB
Total Items	164	21	1	3	112	118	88	288	43	24	81	1361	13	1.86 TI

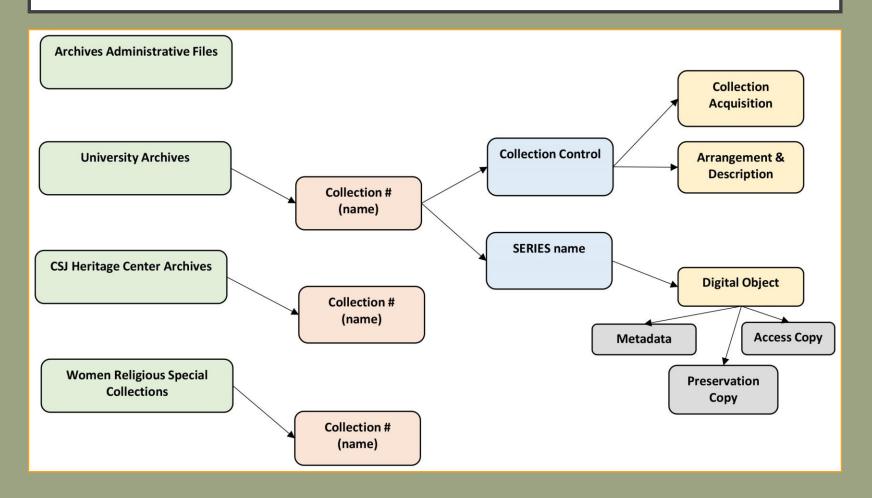


Inventory of Inventories

Pa	Cut Copy ste Forma	D 7	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$		General ▼ \$ ~ % • 60 00 → 00		
	Clipboard	Fg.	Font 😘 Alignme	nt	Number №		
В:	11 *	1 × ✓	f _{sc} Newspapers				
4	А	В	С	D	E		
1	Repository	Collection	File Name		Index Location		
2	CSJ	CDCM	CSJ CDCM Project Collection	Aspace	Online		
3	CSJ	Federation	CSJ.2020.0001 Inventory	DOCX	H Drive - Archives		
4	CSJ	Federation	CSJ oral histories	XLS	Desktop - Processing folder		
5	CSJ	Federation	Records of US Federation	Aspace	Online		
6	N/a	N/a	Digital Asset Inventory	XLS	H Drive - Archives		
7	UA	Alumni	Office of Alumni Relations - Alumni Magazine	Aspace	Online		
8	UA	Lit Magazine	Student Literary Magazine	Aspace	Online		
9	UA	Newspapers	Student Newspapers - The Gleam	Aspace	Online		
10	UA	Newspapers	Student Newspapers - The Teresian	Aspace	Online		
11	UA	Newspapers	Student Newspapers - Avila Collegian	Aspace	Online		
12	UA	Newspapers	Student Newspapers - Avila Asterisk	Aspace	Online		
13	UA	Newspapers	Student Newspapers - Free Fall	Aspace	Online		
14	UA	Newspapers	Student Newspapers - The Avila Phoenix	Aspace	Online		
15	UA	Newspapers	Student Newspapers - Nameless Newspaper	Aspace	Online		
16	UA	Newspapers	Student Newspapers - The Avila Examiner	Aspace	Online		
17	UA	Newspapers	Student Newspapers - The Talon	Aspace	Online		
18	UA	ОМС	Digital Photos Collection	XLS	Desktop - Processing folder		
19	UA	Oral Histories	University oral histories	XLS	Desktop - Processing folder		
20	UA	Registrar	Office of Registrar - Course Catalogs	Aspace	Online		
21	UA	Univ. History	Written Histories	Aspace	Online		
22	UA	Yearbooks	Student Yearbooks (all series)	Aspace	Online		
23	WRSC	Interrupted Lives	Interrupted Lives - Admin Files	Aspace	Online		
24	WRSC	Interrupted Lives	Interrupted Lives - Production materials	Aspace	Online		
25	WRSC	Interrupted Lives	Interrupted Lives - Scripts	Aspace	Online		
26	WRSC	Interrupted Lives	Interrupted Lives - Interview transcripts	Aspace	Online		
27	WRSC	Oral Histories	WRSC Oral history recordings	Aspace	Online		
28			200				
29							



FILE DIRECTORY

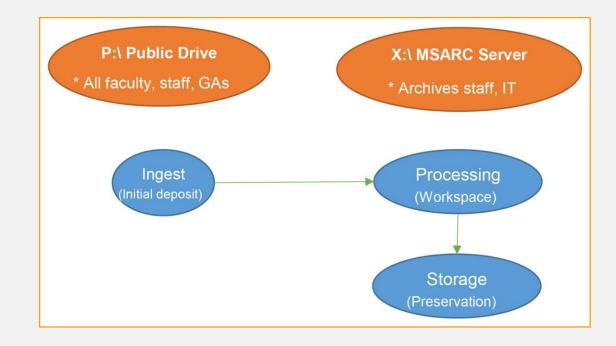






SERVER SECURITY & SETUP





Students fencing outside Blasco Hall





POLICIES & PROCEDURES

Collection policy

- Format neutral
- Appraisal, review & deaccession considerations

Accession/Processing procedures

- Workflows: borndigital transfers, donations, digitized materials
- Email, public drive materials

Metadata templates

- ArchivesSpace fields, particularly location fields for various assets
 - Preservation copy, borndigital, digitized version of analog original, access copy...





DIGITAL POWRR

HTTPS://DIGITALPOWRR.NIU.EDU/

Institute – 2019

"Two-day professional institutes... to build skills for curating and preserving digital collections."

Peer Assessment Program – 2022

"Year-long cohort-based training program... applying digital preservation assessment procedures."



- Tool Grid
- Walk the Workflow
- Self-Assessment
- POWRR Plan

NDSA% Levels of Digital Preservation Assessment Tool									
						Level			
This tool can be used to assist you with determining which aspects of digital preservation you have strength in and which you may need to focus future efforts.	Functional Area	ENTER 0, 1, 2	Level 1 (Know your content)	ENTER 0, 1, 2	Level 2 (Protect your content)	ENTER 0, 1, 2	Level 3 (Monitor your content)	ENTER 0, 1, 2	Level 4 (Sustain your content)
To use this tool, you will enter in a 0, 1, or 2 in each box next to a task. The conditional formatting will color code the tasks.		1	Have two complete copies in separate locations	0	Have three complete copies with at least one copy in a separate geographic location	0	Have at least one copy in a geographic location with a different disaster threat than the other copies	0	Have at least three copies in geographic locations, each with a different disaster threat
2 = Achieved	Storage	2	Document all storage media where content is stored	0	Document storage and storage media indicating the resources and dependencies they require to function	1	Have at least one copy on a different storage media type	0	Maximize storage diversification to avoid single points of failure
1 = Work in Progress			Put content into a stable storage			0	Track the obsolescence of storage and media		Have a plan and execute actions to address obsolescence of storage hardware, software, and media
0 = Not started		0	Verify integrity information if it has been provided with the content	0	Verify integrity information when moving or copying content	0	Verify integrity information of content at fixed intervals	0	Verify integrity information in response to specific events or activities
NOTE: It is recognized that some of these tasks may not be 'in scope' for all organizations. For example, an organization may not be the ones in charge of a task and therefore entering a 0, 1, or 2 is not accurate. It is recommended to leave anything that your organization will not do or be able to do blank. This will keep the cell white allowing review to be done on only the colored cells.	Integrity	0	Generate integrity information if not provided with the content	0	Use write-blockers when working with original media	0	Document integrity information verification processes and outcomes	0	Replace or repair corrupted content as necessary
		0	Virus check all content, isolate content for quarantine as needed	0	Back up integrity information and store copy in a separate location from the content		Perform audit of integrity information on demand		
	Control	2	Determine the human and software agents that should be authorized to read, write, move, and delete content	1	Document the human and software agents authorized to read, write, move, and delete content and apply these	0	Maintain logs and identify the human and software agents that performed actions on content	0	Perform periodic review of actions/access logs
	Metadata	2	Create inventory of content, also documenting current storage locations	1	Store enough metadata to know what the content is (this might include some combination of administrative, technical, descriptive, preservation, and structural)	2	Determine what metadata standards to apply	0	Record preservation actions associated with content and when those actions occur
		1	Backup inventory and store at least one copy separately from content			1	Find and fill gaps in your metadata to meet those standards	1	Implement metadata standards chosen
	Content	2	Document file formats and other essential content characteristics including how and when these were identified	1	Verify file formats and other essential content characistics	0	Monitor for obsolescence, and changes in technologies on which content is dependent	0	Perform migrations, normalizations, emulation, and similar activities that ensure content can be accessed
		2		0	Build relationships with content creators to encourage sustainable file choices				

Created October 2019.



FORWARD PROGRESS

PHASE II

- Update ingest process
- Additional backup storage
- Tools to automate processes

PHASE III

Integrated system

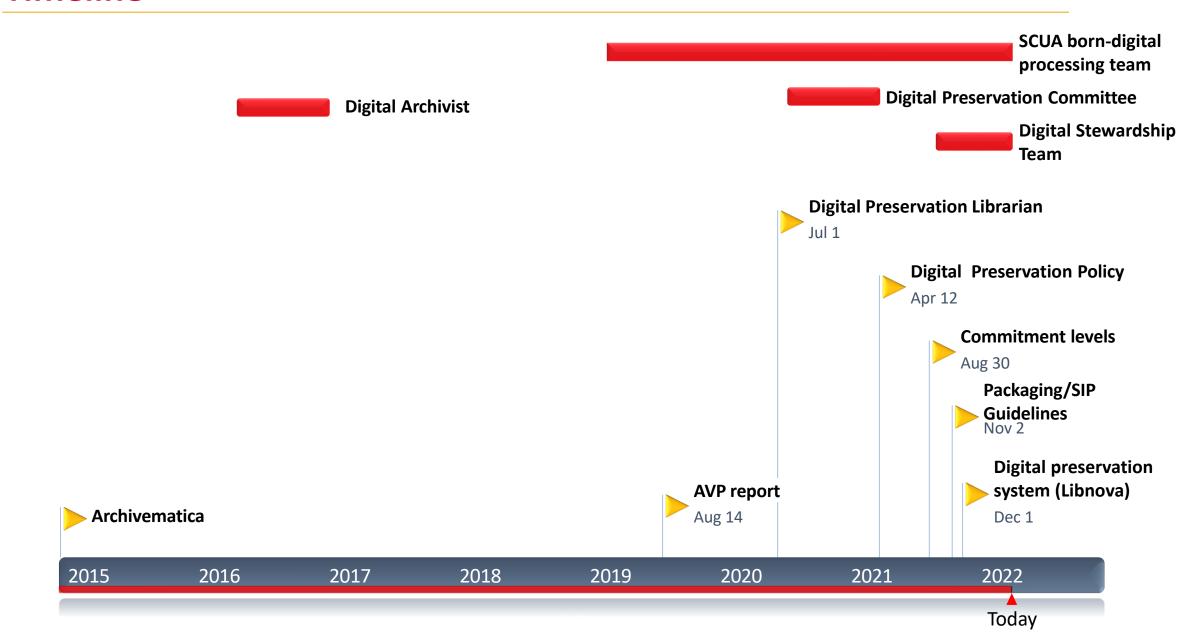




Avila Spirit Squad, 1976



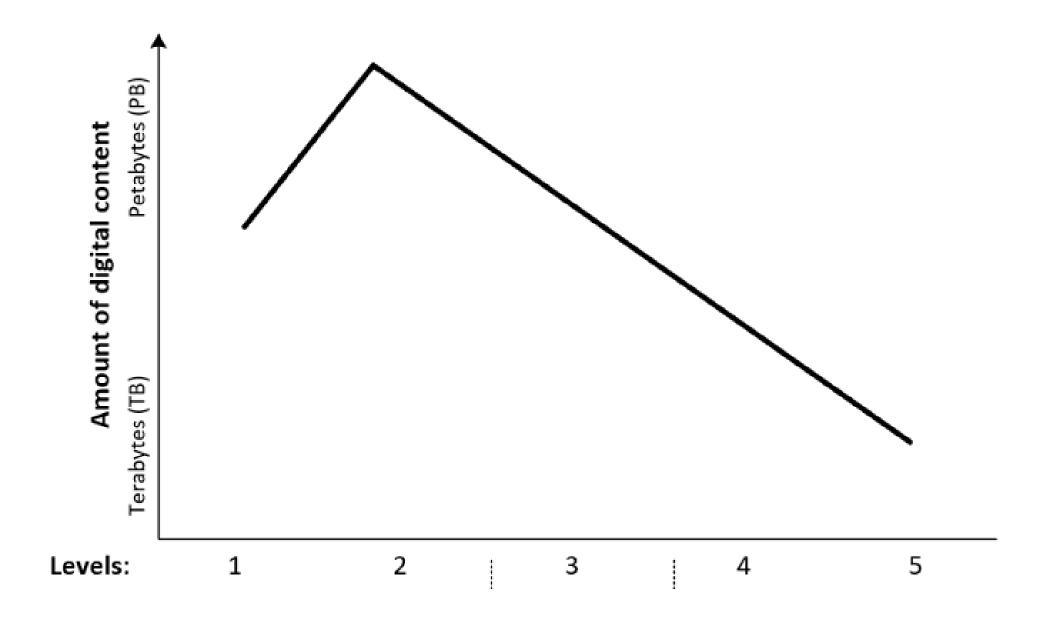
Timeline





Functional Area	Level								
r uncuonal Area	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					

Preservation Services/Actions	Level 1	Level 2	Level 3	Level 4	Level 5	
Preservation Metadata	Minimal/None	Minimal	Full	Full	Full	
Preservation Formats	Maybe	Maybe	Yes	Yes	Yes	
Preservation Copies (multiple locations)	B + 1X + Offline	B + 1-2X + Offline	B + >=4X + Offline	A+B + >=4X + Offline	A+B + >=4X + Offline	
Preservation Approach	Bit	Bit	Full	Full	Full	
Open/Closed	Open/Limited	Open/Limited	Open/Limited	Closed	Closed	
Confidential	No	No	No	Yes	Yes	
Rare/Unique	No	No	Yes	Yes	Yes	
Independence	Low	Low	High	High	High	
Preservation Action Level	Low	Low	Medium	High	High	
Born Digital (BD)/Digitized	Both/Often PDF	More Digitized	More BD	More BD	More BD	
Regulated	No	No	No	No	Yes MIT	





Preservation 1 (22/43)

- Tech
 - Files on platform w/ characterization & local storage with fixity monitoring
- Actions
 - Preservation files in 2nd location w/ fixity monitoring
- Business records
 - Review cycle
 - Storage environments
 - Formats, access, & preservation files; files to support accessibility
 - Inventory Control
- Package contents
 - Checksums
 - Manifest
 - Metadata (descriptive, provenance, preservation, rights)

IOWA STATE UNIVERSITY

Preservation 2 (25/53) Actions ▶ metadata files in second location w/ fixing monitoring, on-demand access audit **Business records** Procedures for packaging Package contents Metadata (technical, OCR, structural, transcripts, VTT) IOWA STATE UNIVERSITY

Preservation 3 (44/114)

- Tech
 - ▶ Libnova/S3, Glacier or Azure bucket
- Actions
 - Archival packaging with METS & LD or RDF & LD, preservation files in 2nd location w/ fixity monitoring, platform w/ ongoing access & actions logging, schedules for log review, dependencies monitoring, fixed interval fixity monitoring, migration actions, files & metadata in 3rd location
- Business records
 - Dependencies
 - File obsolescence and tech migration plan
 - Archival packaging with METS and LD or RDF & LD
- Package contents
 - Documented behavior
 - Files with forensics reports

IOWA STATE UNIVERSITY

Preservation 4 (48/132)

- ► Tech
 - ► LTO, APTrust or MetaArchive
- Actions
 - Files & metadata in 4th location
- Business records
 - Risk management procedures
 - ► Emulation environment
- Package contents
 - Donor records
 - Documentation of entire preservation environment in system-independent storage

IOWA STATE UNIVERSITY







- Carbon footprint
- Maturity models
 - Digital Maturity (Rossman, 2018): Strategy, Leadership, Business model,
 Operations, People, Culture, Governance, Technology
 - Digital Preservation Capability (Dollar, Ashley & Misic, 2014)
 - Rapid Assessment Module v2 (DPC, 2021): Organizational: Viability,
 Policy/Strategy, Legal, IT, Continuous Improvement, Community

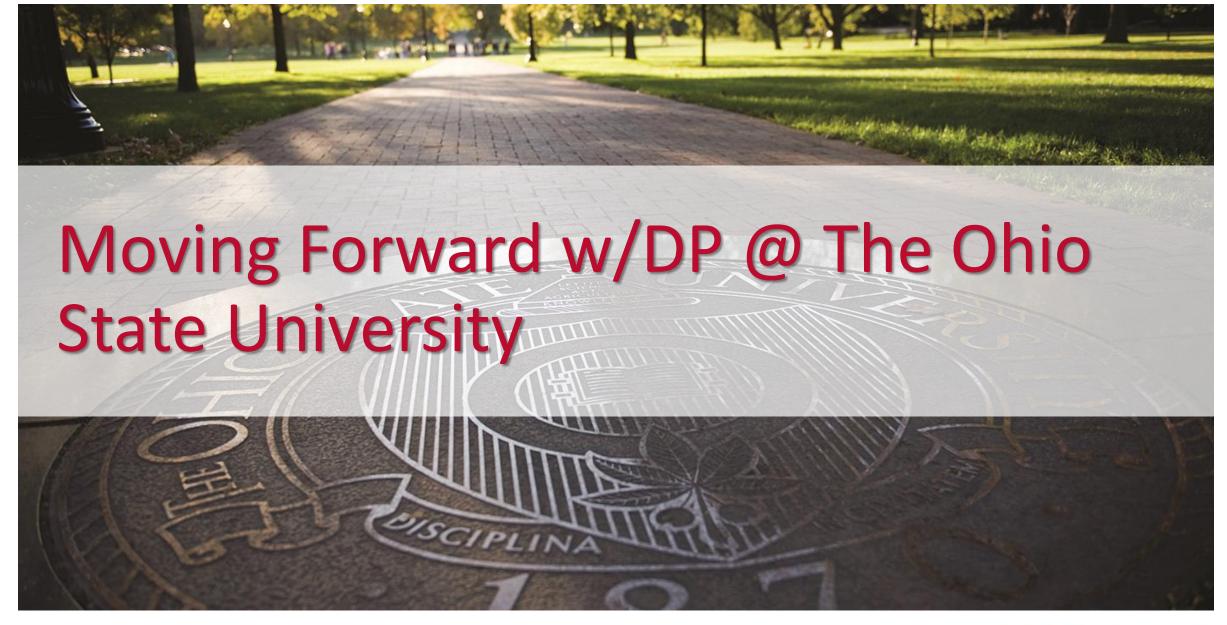
DPCMM Components	Current Capability	Year 1	Year 2	Year 3	Year 4	Year 5	Difficulty
INFRASTRUCTURE							
Policy	1	2	⇔	⇒	3	⇒	IOW
Strategy	0	2	⇒	⇒	3	⇒	MEDIJM
Governance	0	1	2	3	⇒	⇒	MEDIJM
Collaboration	1	2	⇒	3	⇔	⇒	МГВІЛМ
Technical Expertise	1	⇒	⇒	2	⇒	3	MFD1.JM
Open Standard Technology Neutral File Formats	1	⇔	2	⇒	3	⇒	WFBIJM
Designated Community	1	⇨	2	⇒ Doll	3 ar, Ashley	⇒⁄ and Misi	мегылы с, 2014















Our Journey @ Ohio State...



- Policy to Ethos
- Environmental Scan
- Workflows, Workflows Everywhere
- Priority, what is a priority?
- But what about born digital?
- Collaboration, Documentation & Transparency





Why Ethos?



This Photo by Unknown Author is licensed under CC BY

- Policy: a definite course of action adopted and pursued by an institution (e.g. government, ruler, political party) for the sake of prudency, expediency, facility, etc.
 - https://www.dictionary.com/browse/policy
- Ethos: the character or fundamental values of a person, people, culture, or movement [or organizational program].
 - https://en.wiktionary.org/wiki/ethos
- It is not just one person/one program; it is a organizational collaboration





Digital Preservation @ Ohio State

- 2012 2013: DigCCur Project to develop Digital Preservation Policy Framework
- 2013 2016: Task forces and workgroups
- 2015: Fedora/Sufia
- 2016: Web archiving
- 2017: DPL
- 2018-2019: Fedora/Hyrax upgrade
- 2019-2020: Enviro scan/birth of Ethos
- 2020 current: DP&A
- 2022: BDR

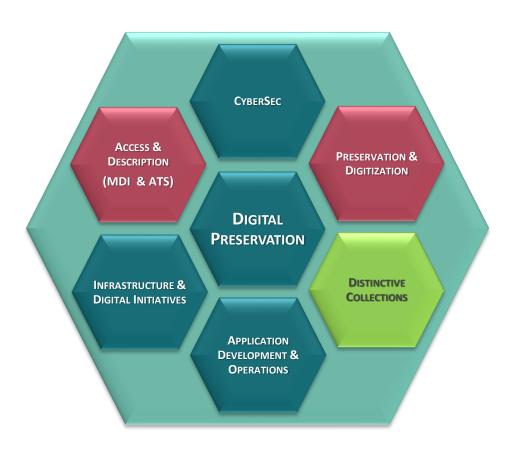


This Photo by Unknown Author is licensed under CC BY-SA-NC





Digital Preservation @ Ohio State

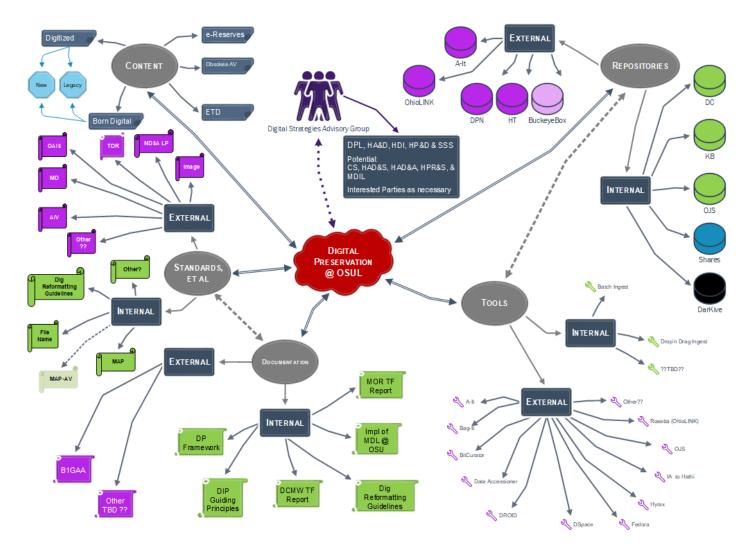


- Articulate the requirements for creating and acquiring preservable digital content...
- Develop guidelines, protocols, and organizational practices when needed...
- Develop and oversee workflows and activity to migrate digital content...
- Collaborate with content owners, subject experts, metadata experts, and others to ensure that appropriate workflows, tools, and infrastructure...
- Contribute to a continuous process of assessment and QC.





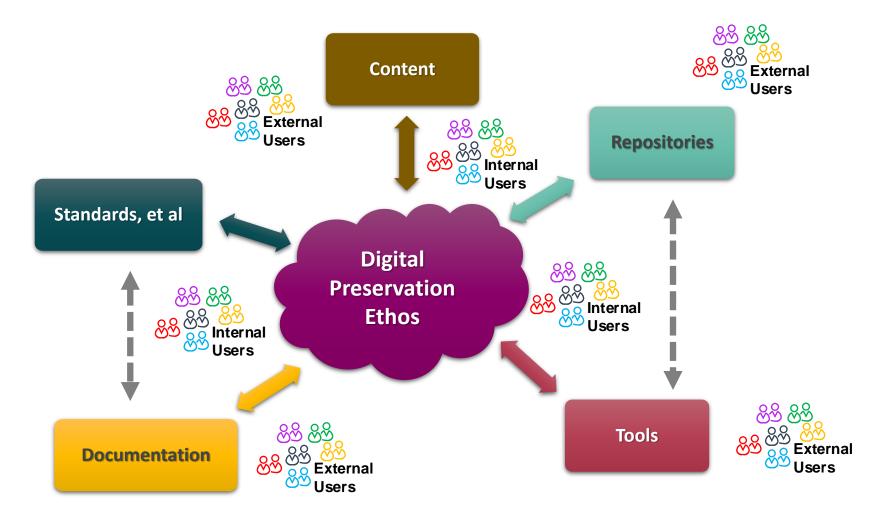
Digital Preservation @ Ohio State







Digital Preservation @ Ohio State









BTAA Review

- Preservation Policies
- Guidelines
- Mission & Vision
- Tools & Standards
- Hardware & Software

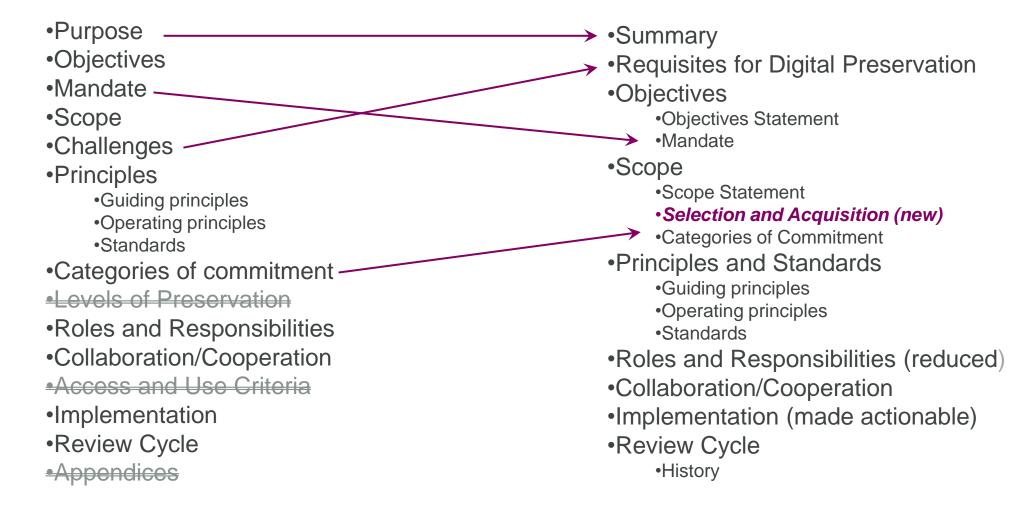
Shared Elements

- Operating Principles: 10 institutions
- Purpose: 9 institutions
- Scope: 8 institutions
- Roles: 8 institutions
- Objectives: 7 institutions
- Mandate: 7 institutions

Historical perspective: "Digital Preservation Policy Framework: A Case Study," *EDUCAUSE Review online*. 49 no. 4 (July 2014) (http://er.educause.edu/articles/2014/7/digital-preservation-policy-framework-a-case-study)











Digital Preservation & Access Work Group

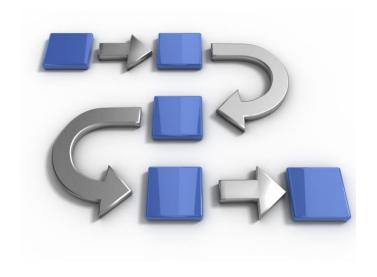


- Grassroots
- Cross-functional
- Consistent Strategies
- Knowledge Management
- Transparency
- Areas of focus
- Task #1: Identify Workflows





Differentiating the Processes



What are the current digitization, born digital acquisition, preservation and access processes in the organization?

- High Level POV
- The activities, functions, and/or hand-offs

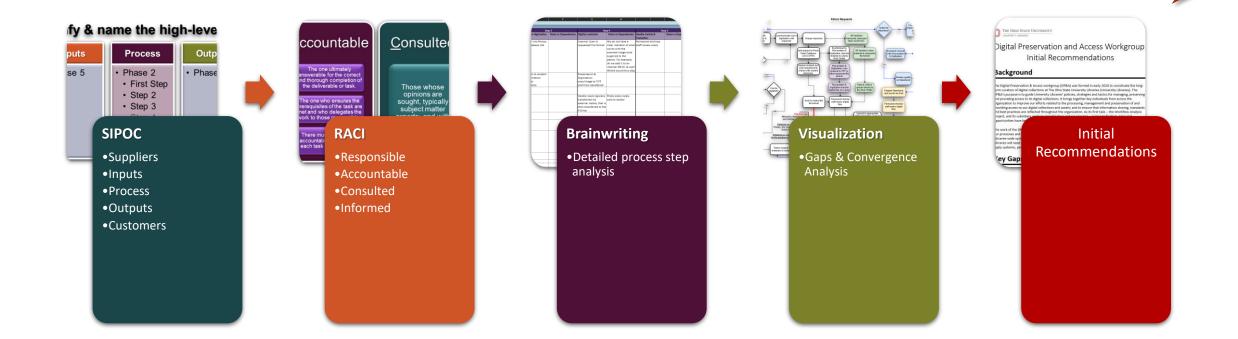






Where We've Been...Workflow Mapping Analysis

2020 2021







What have we shed light on?



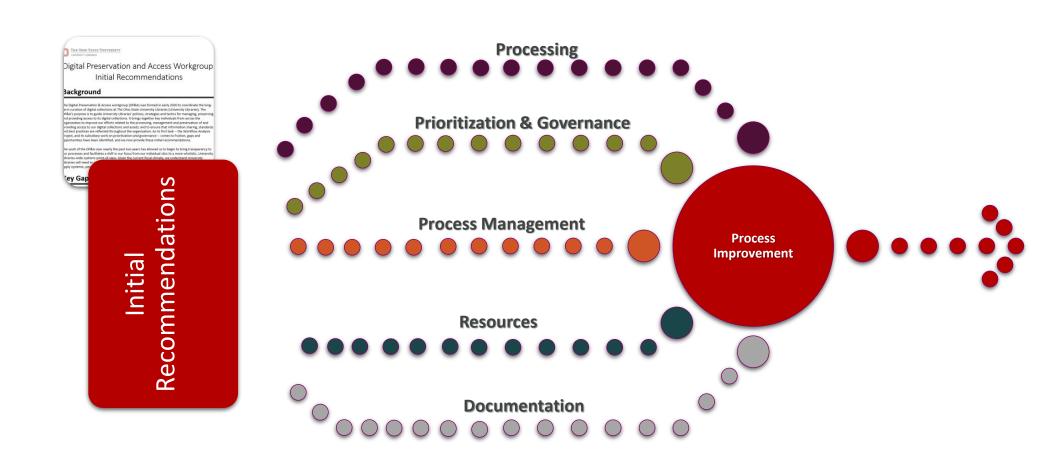
licensed under CC BY-SA-NO

- Created opportunity to contemplate processes for the first time
 - Making Tacit & Implicit Knowledge Explicit
- Shift from "Silo Thinking" to encouraging "Systems Thinking"
 - Making silos transparent & permeable
- Consolidating documentation; making it visible
- Surfaced documentable gaps that we kind of knew were there
- Total Cost of Stewardship





What's next?







Recommendations: Prioritization & Governance



https://line.17qq.com/articles/hchgplhdv.html

- Change Management Process
 - ✓ Problem Statement
 - ✓ Current State → Future State
 - ✓ Defining the Change
 - ✓ Benefits
 - Process
 - ✓ Definition of "Project"
 - Prioritization Factors
 - Measuring Progress





Prioritization Factors



https://line.17qq.com/articles/hchgplhdv.html

Pre-conditions

- Has the collection been accessioned?
- Does this collection meet copyright moratorium exemption?

Alignment

- Priorities of University Libraries
- Who is requesting the work and why?
- EDIA: Amplify nationally and globally marginalized voices by processing born digital or digitizing collections that represent historically underrepresented groups.

Opportunity

- Disaster mitigation
- Fragility and obsolescence of materials:
- Does this work lead to future efficiencies?
- Historic use patterns
- Critical mass
- Distinctiveness

Readiness

- Project size and scope
- Complexity
- · Level of Processing
- Metadata (existing)
- Workplan readiness

Dependencies (negatively scored)

- Metadata (required):
- Exigent circumstances
- Accessibility
- Content restrictions





Next Steps: Immediate Impact



- Complete prioritization sub-workgroup rubric and governance process...
- Inventory of known projects and activities...
- Documentation repo...
- Kaizen event(s)...





Next Steps: Long-term Impact

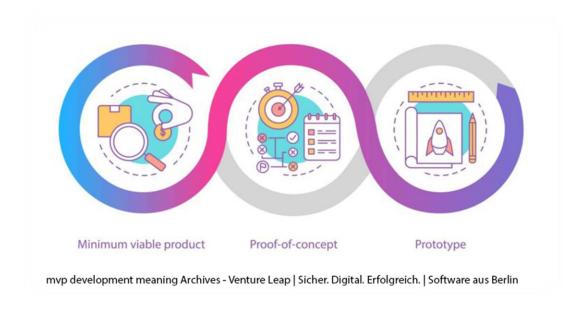


- Process re-engineering, which is dependent upon aforementioned Kaizen event(s) outcome(s)...
- Expertise and skills development...
- Development, articulation and implementation of standardized best and good-enough practices...
- Digital asset management dashboard...





MVP for New Digital Repo



- Use case for minimal viable product (MVP)
- Lean Start Up
 - Build ⇔ Measure ⇔ Learn





What it is and is not...



This Photo by Unknown Author is licensed under CC BY-SA-NC

- Dim archive
- Proposed content
 - Minimally described by design
 - KB legacy preservation files
 - Temporally restricted files
- Mediated access
 - Not publicly accessible

- Analogous to the Book Depository
 - More intellectual control
 - DROID & Bag manifests
 - DLP Report
- Fedora ID and finding aid linkage





The on-going journey...

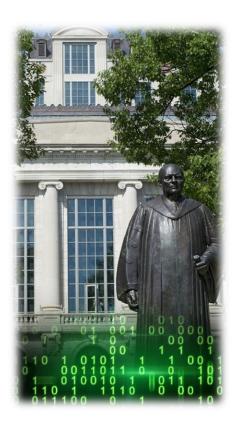


- Continued Collaboration
- Continued Experimentation
- Further Documentation
- Transparency





Contact Info



Dan Noonan

Associate Professor
Digital Preservation Librarian
University Libraries | Information Technology
The Ohio State University
614.247.2425 Office

noonan.37@osu.edu | go.osu.edu/Noonan

@dannynoonan1962

Pronouns: he/him/his





Discussion/Comments/Questions



This Photo by Unknown Author is licensed under CC BY-NC-ND

Contact Information

IOWA STATE UNIVERSITY

- Daniel Hartwig
 - Head, Special Collections & University Archives
 - <u>dhartwig@iastate.edu</u>



- Amy Moorman
 - Archivist
 - Amy.Moorman@avila.edu



- Dan Noonan
 - Associate Professor, Digital Preservation Librarian
 - noonan.37@osu.edu
 - 614.247.2425
 - @dannynoonan1962