

It's Not Impossible: Digital Preservation Programs, and How You Can Make One Too

Larissa Kraye

Leon S. McGoogan Health Sciences Library



University of Nebraska
Medical Center

Hi, I'm Larissa Kraye the Digital Archivist at the McGoogan Health Sciences Library at the University of Nebraska Medical Center. My poster, It's Not Impossible: Digital Preservation Programs and how you can make one too highlights the library's recent assessment of our digital preservation environment and development of an action plan to implement a program that meets national standards.



Digital Preservation

- Not just file back up!
- Digital Preservation uses policies, strategies, and actions to ensure long-term access to digital content by managing the challenges of file corruption, media failure, and technology changes.



Our institutional intellectual property is at risk of being lost. Hardware and software become obsolete, storage media can fail, and files can become corrupt. Digital preservation is more than just a back-up, it is an on-going process that needs to be addressed programmatically. A robust program will have strong organizational support through solid policies, a dedicated budget, and knowledgeable and engaged staff. The organization must also meet the technological challenges of long-term access through varied storage, complete metadata, and intellectual control of their digital content.

ISO 14721:2012 – Open Archival Information System

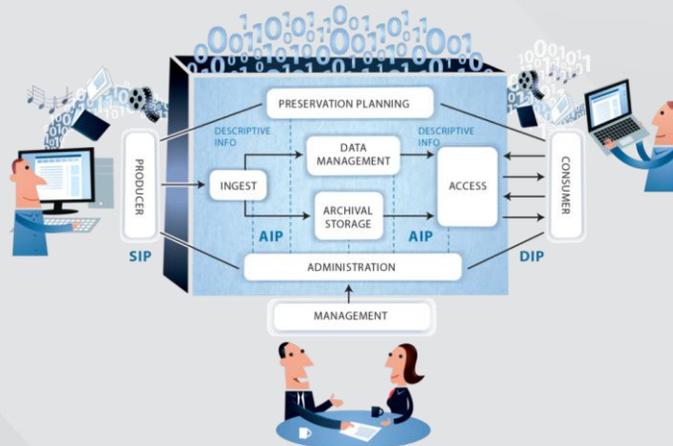


Image by: Jørgen Stamp, CC BY 2.5 DK, via [Wikimedia Commons](#)

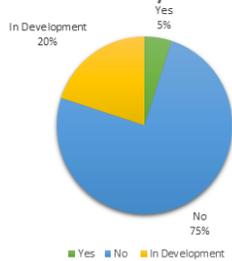
Digitalbevaring.dk

Assessment tools and best practices are based on the ISO standard Open Archival Information System (OAIS) reference model. The OAIS model identifies and describes the functions needed to provide long-term preservation and access of materials. The model includes both organizational actions through administration and preservation planning as well as technical actions with data management and archival storage.

Survey of Academic Health Sciences Libraries



Do you currently have a Digital Preservation Policy?



Do you currently use any digital preservation systems?



In January 2021, a survey was sent to the Association of Academic Health Sciences Libraries listserv regarding current digital preservation environments. We wished to get a feel for the landscape among similar institutions and found that while this was an area most institutions were aware needed attention, most had yet to devote resources to it. Of the 20 responses, the majority indicated that they did not have any digital preservation policy, procedures or workflows. We expect that the work we have done can be used to strengthen other libraries' programs or at least give them a place to start.



Assessment Tools

- Digital Preservation Coalition – Rapid Assessment Model (DPC RAM)
 - Comprehensive maturity model that benchmarks both organizational technology capacities
- National Digital Stewardship Alliance’s (NDSA) Levels of Digital Preservation
 - Rubric to benchmark technical requirements
 - Used by most preservation tools and software

Our first step was to assess our current environment. While there are a variety of assessment tools available, the library utilized two popular tools to complete the work . We chose to use the organizational infrastructure section of the DPC RAM while using the NDSA’s Level of Preservation for technical assessment. We choose this hybrid approach because most preservation tools and software utilize the NDSA’s rubric to align their functions to best practice so we felt using this rubric would allows us to more easily evaluate what tools we needed to acquire.

DPC RAM – Organization Infrastructure

Green – Complete
Yellow – In Progress
Red – Not Started



Organizational Capabilities	Levels			
	1 - Awareness	2 - Basic	3 - Managed	4 - Optimized
Organizational Viability: Governance, organizational structure, staffing and resourcing of digital preservation activities.	The organization is aware of the need to support digital preservation activities.	There is senior management buy-in and support.	There is commitment from senior management.	Benefits of digital preservation are recognized, characterized, and embedded throughout the organization.
		Staff have assigned responsibilities and the time to undertake them.	Staff have the skills they need to carry out digital preservation activities and access to relevant expertise where required.	A cross departmental digital preservation management board has been established.
		A budget for digital preservation has been allocated (may be time-banded).	A dedicated core budget for digital preservation has been allocated.	One or more staff are considered to be experts in their field.
		Staff development requirements have been identified.	Budget, staff roles, and development needs are regularly assessed.	Budget, staff roles, and development needs are proactively assessed in anticipation of future changes.
			Metrics and reports can be generated about the digital archive to help inform resourcing, planning, and management.	Metrics and reports about the digital archive are considered with perspective of future needs to proactively inform resourcing, planning, and management.
Policy and Strategy: Policies, strategies, and procedures which govern the operation and management of the digital archive.	The organization is aware of the need to develop a policy framework and may have some relevant policies, but no digital preservation policy or strategy exists.	A high-level digital preservation policy or strategy exists.	The digital preservation policy/strategy is aligned with other organizational policies and is reviewed according to an agreed schedule.	A full suite of policies, strategies, and procedures relating to the preservation of digital content is in place.
		Other policies relating to digital preservation may exist but when are gaps in coverage.	A suite of documented processes and procedures for managing content within the digital archives exists.	Policy and strategy is fully implemented and staff actively engage with it.
		Some procedures for managing digital content are in place and may be documented.	Responsibility for digital preservation is clearly owned.	Policy, strategy, and procedure is proactively monitored and updated to reflect internal changes, changes in other policies or other external factors.
		Scope of collection is defined and documented (e.g. collections, development policy, retention schedules).	All relevant staff are aware of the digital preservation policies, strategies, and procedures.	
IT Capability: Information Technology capabilities for supporting digital preservation activities.	The organization is aware of the need for IT capability to support the digital archive, and has an understanding of basic options.	Basic IT support is available to the digital archive.	Advanced IT support is available to the digital archive.	An enhanced level of IT support is available to the digital archive.
		IT systems are documented at a basic level.	IT systems are regularly patched and updated.	IT demonstrates good understanding of, and engagement with, digital preservation issues.
			New tools and systems are deployed when required.	Digital preservation requirements are taken into account when sourcing new IT systems.
			IT systems are comprehensively documented.	A detailed roadmap exists for future development of IT systems.
			Contracts and services with third party service providers are well managed and documented.	Potential new tools and systems are proactively identified and tested.
Continuous Improvement: Processes for the assessment of current digital preservation	The organization is aware of the need to understand current position and define goals.	An initial benchmarking exercise has been carried out.	Goals have been established and agreed with senior managers.	Continuous/improvement reviews have been achieved and maintained as appropriate.
		Steps in digital preservation capabilities have been identified.	Readiness is in place to reach goals.	Recommendations for improvement have been acted upon.
		There is an understanding of where the organization is relative to peers.	Benchmarking exercise is repeated periodically.	Goals and roadmap are reviewed periodically.

Each rubric provided functional areas with metrics defined under each level. By assigning each metric a color based on its completion we easily and visually understood our starting point. In addition to the rubric, the final report included a written summary assessment of each functional area and any work that is being done to fulfill the metric.

The DPC RAM organizational capabilities rubric assesses six functional areas: Organizational Viability, Policy and Strategy, Legal Basis, IT Capability, Continuous Improvement, and Community, using four maturity levels. This is just a portion of our complete rubric which shows our work in some of the functional areas

NDSA's Levels of Preservation



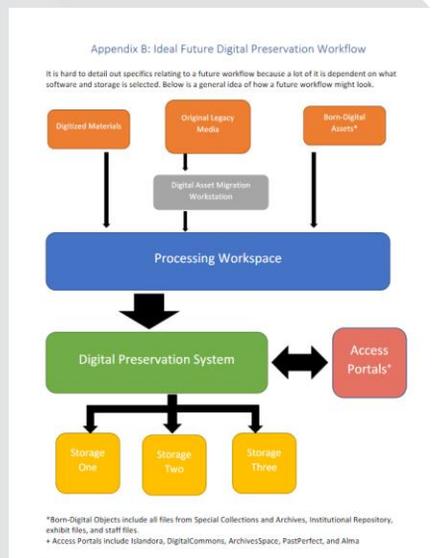
Green – Complete
 Yellow – In Progress
 Red – Not Started

Functional Area	Level 1 - Know	Level 2 - Protect	Level 3 - Monitor	Level 4 - Sustain
Storage	Have two complete copies in separate locations.	Have three complete copies with at least one copy in a separate geographic location.	Have at least one copy in a geographic location with a different disaster threat than the other copies.	Have at least three copies in geographic locations, each with a different disaster threat.
	Document all storage media where content is stored.	Document storage and storage media indicating the resources and dependencies they require to function.	Have at least one copy on a different storage media type.	Maximize storage diversification to avoid single points of failure.
	Put content into a stable storage.		Track the obsolescence of storage and media.	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.	Verify integrity information when moving or copying content. Use write-blockers when working with original media.	Verify integrity information of content at fixed intervals. Document integrity information verification processes and outcomes.	Verify integrity information in response to specific events or activities. Replace or repair corrupted content, as necessary.
	Virus-check all content, isolate content for quarantine as needed.	Back up integrity information and store copy in a separate location from the content.	Perform audit of integrity information on demand.	
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.	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.	Record preservation actions associated with content and when those actions occur.
	Backup inventory and store at least one copy separately from content.		Find and fill gaps in your metadata to meet those standards.	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.	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.
		Build relationships with donors to encourage sustainable file choices.		

The NDSA's Levels of Preservation rubric assesses technical infrastructure in five functional areas: Storage, Integrity, Control, Metadata, and Content, also using four maturity levels. This complete look at the technical rubric shows that while we have started work towards digital preservation, we are still working to achieve Level 1 in any of the functional areas.

Action Plan

- Areas of Need
 - Digital Preservation Policy
 - Intellectual Control
 - Digital Preservation System & Workflow
 - Staff Training



Once the assessment was complete, the library identified and evaluated the areas needing the most support and developed an action plan for the repository to work towards best practices. We identified four areas needing the most attention. First was the creation of policy which would help establish organization support and infrastructure. Second is gaining intellectual control of our collections, what do we have, how much do we have and what media is it currently stored on. This will be accomplished through inventories and processing of our collections. The third area is to investigate software and workflows to provide the technical infrastructure required. Finally, staff training is needed to build more in-house knowledge of digital preservation best practices.

Next Steps

- Action plan approved in March 2021
- Preservica selected as preservation system
- Digital preservation is incorporated into the library's strategic plan.

Leon S. McGoogan Health Sciences Library Digital Preservation Plan

Table of Contents	
Section One: Executive Summary	3
Section Two: Standards and Assessment Tools	4
Objective	4
Standards	4
Digital Preservation Philosophy	5
Assessments/Benchmarks	5
Section Three: Current Environment	6
Organizational Capabilities	6
DPC RAM Organizational Infrastructure Chart	8
Technical Infrastructure	11
NDSA's Levels of Preservation Chart	13
Section Four: Best Practices and Moving Forward	15
Organizational Infrastructure	15
Technical Infrastructure	17
Section Five: Technical Preservation Need by Content	20
Section Six: Current Applications	22
Section Seven: Discussion, Timeline, and Budget for Moving Forward	24
Timeline	24
Budget	25
Appendix A: Current SCA Digital Preservation Workflow	26
Appendix B: Ideal Future Digital Preservation Workflow	29
Appendix C: Processing Tools	30
Digital Preservation Systems	30
Preservica	30
Archivermatica	31
Estimated costs	32
Digital Object Storage	32
Storage Media Migration Workstation (Digitization Workstation)	33
XML Editor	33
Appendix D: Glossary	34
Appendix E: AAHSL Digital Preservation Survey Results	36
Bibliography and Works Consulted	40

A detailed plan, including the assessment, preliminary budget, and timeline, was developed to provide a look at the projects, resources, and time needed to complete all the metrics for a sustainable digital preservation program. This was presented to library administration and approved in March. We have completed an evaluation of software programs and begun the acquisition of the Preservica system. Other metrics have been incorporated into the library's strategic plan and with priority given to the project we anticipate meeting national standards within two years. We then hope to begin discussions on digital preservation best practices and workflows throughout the university campus.



To learn more about digital preservation and the assessment tools used, please see the resource section of my poster. Thank you for your time.