


Words with Friends: Digital Preservation Peer Assessment

Digital
Commonwealth
Conference
2019



Becky Geller
Preservation Specialist

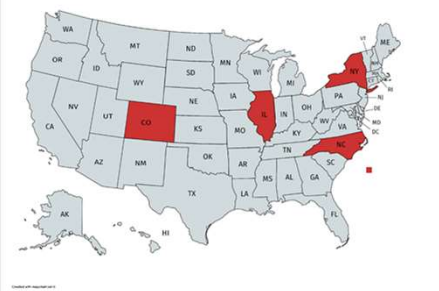
Agenda

- Overview of Digital Assessment Training Program
- Definitions
- Concepts & Challenges in Digital Preservation
- Digital Preservation Assessment
 - Community documents and standards
 - Concepts and challenges
 - Preservation strategies
- Peer Assessment Framework
 - Overview
 - Assessment practice

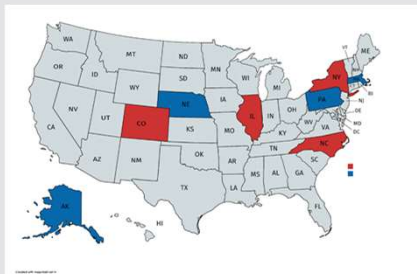
NEH GRANT (JAN – AUG 2017)



- Task Force Meeting – Develop Framework
- Round 1 Assessments – Pilot Framework



NEH GRANT (SEPT 2017 – JUNE 2018)



- RAP Meeting – Revise Framework, Develop Glossary, Create Peer-Assessment Framework & Workshop
- Round 2 Assessments - Include Peer-Assessment Workshops and Observer
- Feedback Interviews with Round 1 & 2 Clients

NEH GRANT (JULY – OCT 2018)



- Training Institute – Cohort of 12 conducting assessments
- Feedback interviews with these clients

FINAL PHASE OF GRANT

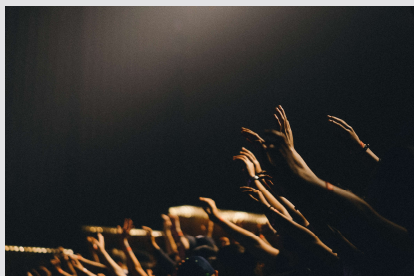


Photo by [Edwin Andrade](#) on [Unsplash](#)

- Symposium (Nov. 2018) – Practitioners and funders in discussion
- Release of Digital Preservation Assessment Handbook
- Release of the Peer-Assessment Workshop Curriculum:

<https://bit.ly/2Up8aY0>

Definitions

Digital Preservation

Digitization for Preservation vs. Digital Preservation

- Digitization for preservation: Activities that result in the creation of digital objects worthy of long-term preservation, including selection, image capture, description, and compilation for delivery
- Digital preservation: Tools, operations, standards, and policies that help ensure that this investment is not squandered

(Paul Conway, "Preservation in the Age of Google")

Digital Collections

- Digital images attached to a record for inventory — NO
- Records that have been created on a word processor and you only receive a digital file — **YES!**
 - [These are considered to be born-digital]
- You actively create digital representatives of your physical objects for users to research — **YES!**
- You are creating exhibition-quality digital reproductions — **YES!**
 - [Can include audio, video, maps, text, etc.]

What Is “Digital”?



- Bits/Bitstream
- Filetype/Format
- Software: source code & machine code
- Middleware
- Platforms/OS
- Firmware & Hardware

Digital Preservation Concepts & Challenges

The Concepts

- Digital vulnerabilities: technological risks
- Digital vulnerabilities: organizational risks
- Preservation strategies



Photo by Benjamin Smith on Unsplash

Digital Vulnerabilities: Technological Risks

Risk: Obsolete Hardware



Risk: Obsolete Hardware?



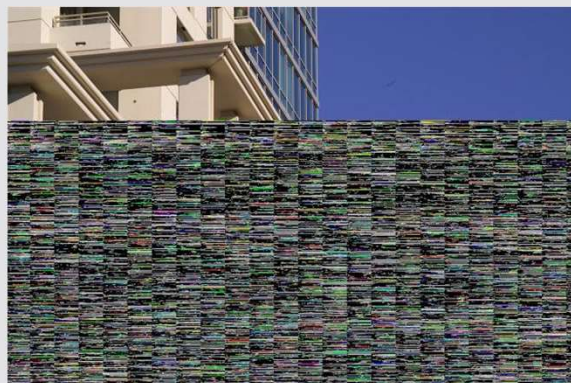
Risk: Obsolete Software



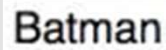
Risk: Hardware/Software Failure



Risk: Data Corruption



Risk: (In)Authenticity



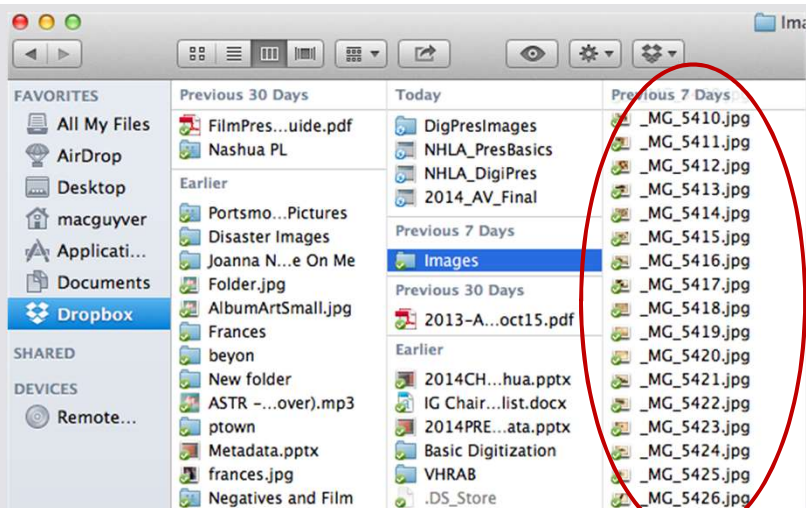
From Wikipedia, the free encyclopedia

This is an old revision of this page, as edited by Guitar Godd23 (talk | contribs) at 20:55, 14 June 2007. It may differ significantly from the current revision.

(diff) ← Previous revision | Current revision (diff) | Newer revision → (diff)

DUN NUH NUH NUH DUH NUH NUH NUH NUH
BATMAN! DUN NUH NUH NUH DUH NUH NUH
NUH NUH BATMAN! DUN NUH NUH DUH NUH
NUH NUH BATMAN! DUN NUH NUH NUH
DUH NUH NUH NUH BATMAN! BATMAN!
BATMAN! DUN NUH NUH NUH DUH NUH NUH
NUH NUH **BATMAN**

Risk: Loss of Context



IP/Copyright Restrictions



Digital Vulnerabilities:
Organizational risks

Mandate

- Do you have a right to collect and care for these objects over time?
- Do you have an organizational commitment to these activities?



Roles and Responsibilities



- Are all stewardship tasks explicitly assigned?
- Is your staff trained to do these roles?

Selection/Creation Policy



- Media types
- File formats
- File sizes
- Content types

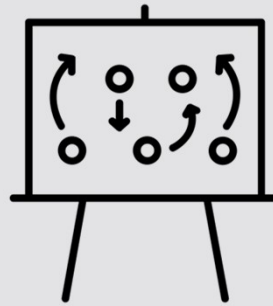
Loss (simply doing nothing)





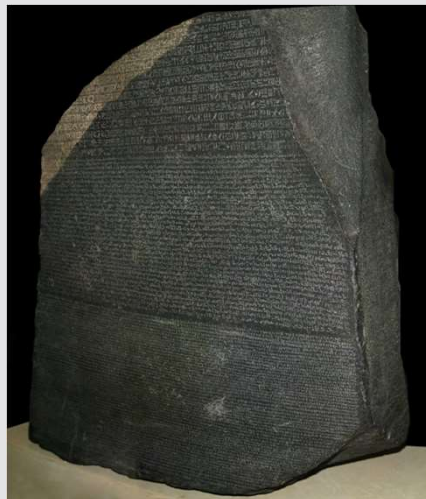
Strategy: Planning

- Strategic planning
- Succession planning
- Tied to organization's mission and strategic initiatives



Created by Made
from Noun Project

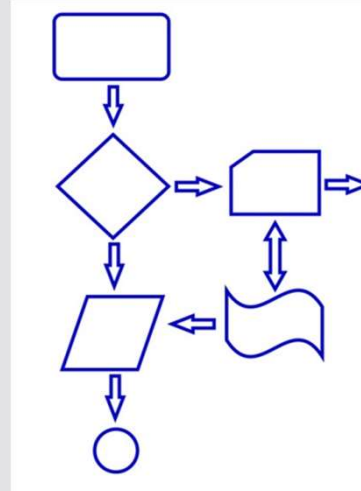
Strategy: Policies



- A few good ones:
 - Designated Community
 - Collection Development (with Rights)
 - Preservation Plan
 - Preservation Action Plans

Strategy: Procedures

- Support long term goals
- Support **unknown** long term goals
- Enable consistent implementation
- Should be written down and fully documented



Strategy: Normalization & Migration

File Format Sustainability factors:

- Documentation is available
- Widely adopted
- Easily rendered by different programs and operating systems



Strategy: Authentication & Fixity

- User authentication and permissions protect against vandalism and accidental deletion
- Periodic tests for fixity using automated systems will verify the integrity of objects after transfer or save

```
d131dd02c5e6eec4 693d9a0698aff95c 2fcab58712467eab
55ad340609f4b302 83e488832571415a 085125e8f7cdc99f
d8823e3156348f5b ae6dacd436c919c6 dd53e2b487da03fd
e99f33420f577ee8 ce54b67080a80d1e c69821bcb6a88393
```

MD5 Checksum algorithm

Strategy: Metadata

VRA Core

MODS
DCMI
METS

PBCORE
PREMIS

Standards

Schema

Organization

Interoperability

Peer Assessment Framework



Community Documents and Standards

- Curation Lifecycle Models
- OAIS Open Archival Information System reference model

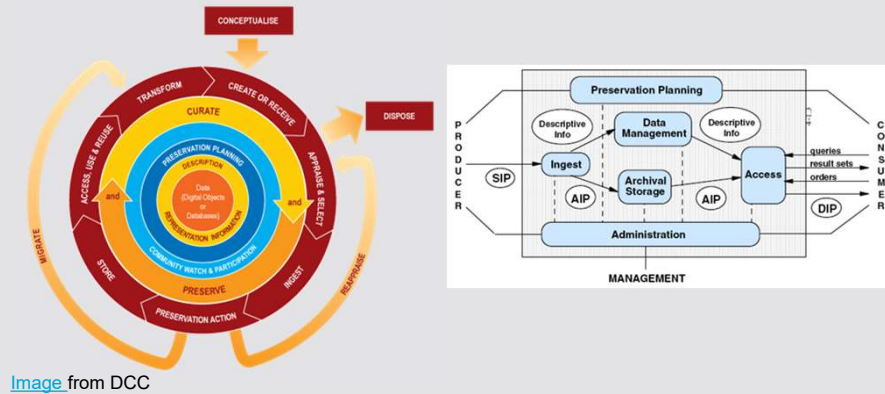
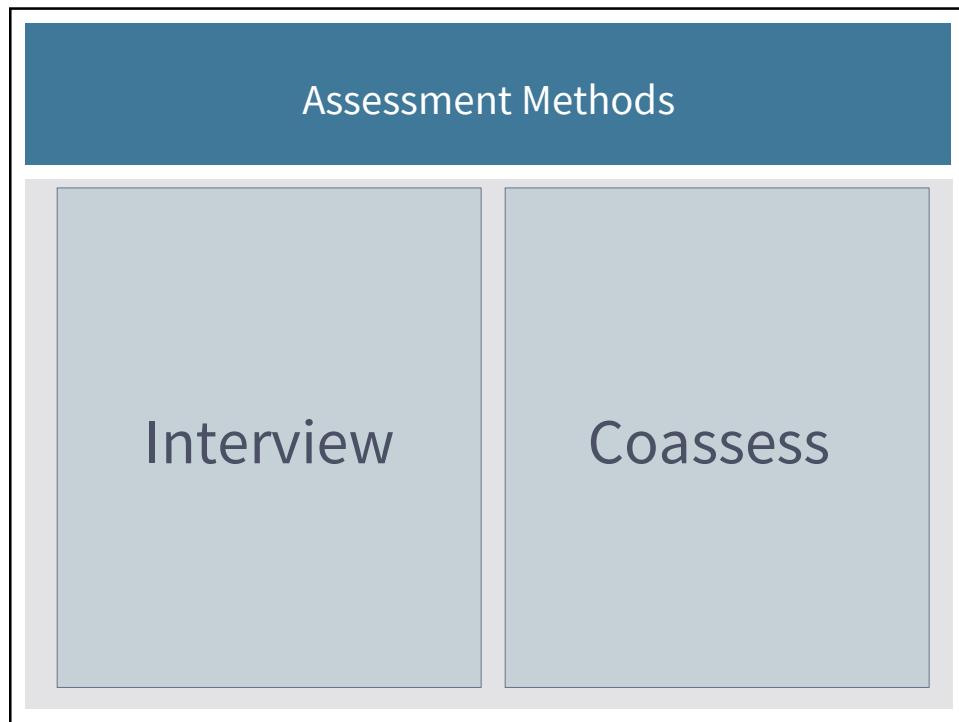


Image from DCC

NDSA Levels of Preservation - 2013

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	<ul style="list-style-type: none"> - Two complete copies that are not collocated - For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system 	<ul style="list-style-type: none"> - At least three complete copies - At least one copy in a different geographic location - Document your storage system(s) and storage media and what you need to use them 	<ul style="list-style-type: none"> - At least one copy in a geographic location with a different disaster threat - Obsolescence monitoring process for your storage system(s) and media 	<ul style="list-style-type: none"> - At least three copies in geographic locations with different disaster threats - Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems
File Fixity and Data Integrity	<ul style="list-style-type: none"> - Check file fixity on ingest if it has been provided with the content - Create fixity info if it wasn't provided with the content 	<ul style="list-style-type: none"> - Check fixity on all ingests - Use write-blockers when working with original media - Virus-check high risk content 	<ul style="list-style-type: none"> - Check fixity of content at fixed intervals - Maintain logs of fixity info; supply audit on demand - Ability to detect corrupt data - Virus-check all content 	<ul style="list-style-type: none"> - Check fixity of all content in response to specific events or activities - Ability to replace/repair corrupted data - Ensure no one person has write access to all copies
Information Security	<ul style="list-style-type: none"> - Identify who has read, write, move and delete authorization to 	<ul style="list-style-type: none"> - Document access restrictions for content 	<ul style="list-style-type: none"> - Maintain logs of who performed what actions on files, including deletions 	<ul style="list-style-type: none"> - Perform audit of logs

<http://www.digitalpreservation.gov/ndsa/ndsa-glossary.html>



Resources for Digital Preservation Planning

Selected Resources

Staff Training & Professional Development

- Digital POWRR Institutes: <http://digitalpowrr.niu.edu/institutes/>
- Digital Preservation Management Workshop: <https://dpworkshop.org/>
- NEDCC Digital Directions: <https://www.nedcc.org/preservation-training/dd19-home>
- LYRASIS online classes: <https://www.lyrasis.org/services/Pages/Classes.aspx>
- Digital Archives Specialist Program, Society of American Archivists: <http://www2.archivists.org/prof-education/das>

Selected Resources

Tools, Systems, Practices

- AVPreserve. 2015. Fixity storage tool. <https://www.avpreserve.com/tools/fixity/>
- AVPreserve. 2015. MDQC metadata tool. <https://www.avpreserve.com/tools/mdqc/>
- Community-Owned Digital Preservation Tool Registry (COPTR). <http://coptr.digipres.org/MainPage>
- Duracloud storage. <http://www.duracloud.org/>
- Van Malssen, Kara. "Implementation of systems for Media/Digital Asset Management in 10 Steps." <http://www.slideshare.net/kvanmalssen/implementation-mam-10steps/>
- POWRR Tool Grid. <http://digitalpowrr.niu.edu/digital-preservation-101/tool-grid/>

Selected Resources

Workflows

- "You've Got to Walk Before You Can Run" OCLC. <https://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf>
- Archivematica: <https://wiki.archivematica.org/MainPage>
- Community Owned Workflows (COW): <http://coptr.digipres.org/Workflow:CommunityOwnedWorkflows>
- OSSArcFlow: <https://educopia.org/research/ossarcflow>

Questions?

Becky Geller
Preservation Specialist
978-470-1010
bgeller@nedcc.org