

Digital Preservation Practices across the Alliance

Devin Becker, Chair, Digital Preservation Working Group

Digital Preservation Working Group

Sara Amato, Willamette University

Devin Becker, University of Idaho

Laura Buchholz, Reed College

Moriah Caruso, University of Washington

Brian Davis, Oregon State University

Danielle Mericle, University of Oregon

Jodi Allison-Bunnell, Alliance staff

Survey Summary/Methodology

- Based on work out of the National Digital Stewardship Alliance (NDSA)
 - NDSA Table: “Levels of Digital Preservation”
 - “Staffing for Effective Digital Preservation” NDSA Report, December 2013
 - “USGS Guidelines for the Preservation of Digital Scientific Data.” April 2014
- Delivered to Orbis Cascade Alliance Members in May 2016
- 32 of 39 Members responded
- 2 Parts
 - Current practices, size of archive, staffing, and policies/education
 - NDSA Level Evaluation

General Results - Digital Preservation Practices

- Nearly all members report that they create and/or manage digitized photographs (97%), books and maps (84%), and audio or audiovisual materials (87%). Electronic theses and dissertations are also widely managed (84%).
- 25 of 32 indicate they do engage in digital preservation practices
- Of those 25:
 - 8 have a digital preservation policy in place; 2 are developing one
 - 8 offer digital preservation education
- Staffing:
 - Majority have less than 1 FTE assigned to digital preservation
 - All indicate an increase in staffing would be helpful

Digital Archives Size/Growth

- Libraries in the Orbis Cascade Alliance manage about 149TB of digital content for preservation purposes
 - 9 institutions manage 5 or more TB,
 - 5 manage between 1 and 4.9 TB
 - 10 manage less than 1TB
- Digital archives have grown 25% in last year, 86% in past 5
- Estimated growth for next year: 62%

NDSA

“Levels of Digital Preservation”

Table 1: Version 1 of the Levels of Digital Preservation

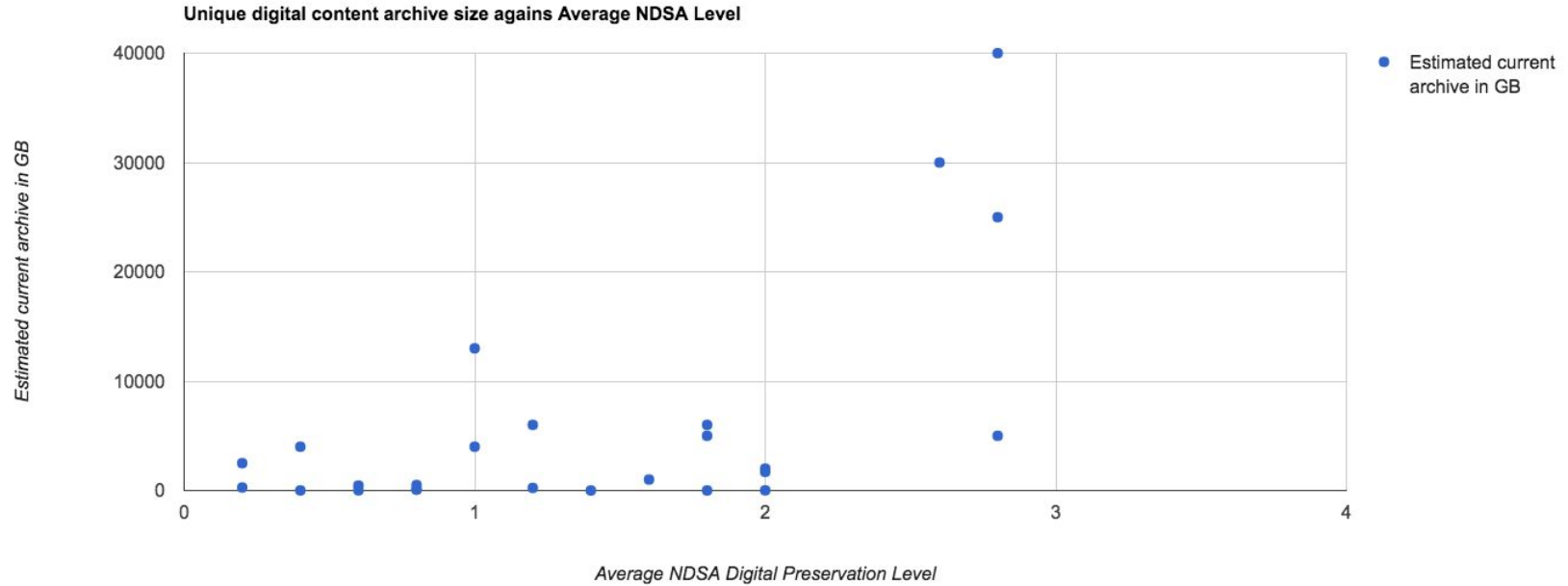
	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	<ul style="list-style-type: none"> - Two complete copies that are not collocated - For data on heterogeneous media (optical discs, hard drives, etc.) get the content 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 individual files - Restrict who has those authorizations to individual files 	<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 and preservation actions 	<ul style="list-style-type: none"> - Perform audit of logs
Metadata	<ul style="list-style-type: none"> - Inventory of content and its storage location - Ensure backup and non-collocation of inventory 	<ul style="list-style-type: none"> - Store administrative metadata - Store transformative metadata and log events 	<ul style="list-style-type: none"> - Store standard technical and descriptive metadata 	<ul style="list-style-type: none"> - Store standard preservation metadata
File Formats	<ul style="list-style-type: none"> - When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs 	<ul style="list-style-type: none"> - Inventory of file formats in use 	<ul style="list-style-type: none"> - Monitor file format obsolescence issues 	<ul style="list-style-type: none"> - Perform format migrations, emulation and similar activities as needed

NDSA Levels

NDSA category	Storage & Geographic Location	File Fixity & Data Integrity	Information Security	Metadata	File Formats
Alliance average level	1.17	0.54	1.29	1.79	2.04
Alliance median level	1	0	1	2	2

- Average NDSA level for Alliance institutions was 1.37 (out of 4), with the mean being 1
 - Strengths: Metadata and File Formats
 - Weaknesses: File Fixity & Data Integrity

Archive Size vs. NDSA Level



Recommendations

1. Investigate a shared digital archive for digital archival content from Alliance members.
 - Why: Making available storage in locations with different disaster threats for Alliance members would greatly improve the levels of digital preservation for many members.

Recommendations cont.

2. Provide in-person or web-delivered education and assistance on file-fixity and data integrity, as well as how to record and store transformation metadata.

- Why: Address a lack of knowledge across alliance regarding file fixity, checksums, etc..

Recommendations cont.

3. Create Alliance portal for digital preservation training kits and checklists/quick-start guides for Alliance institutions looking to improve their digital preservation practices.

- Why: Allows institutions to address/educate themselves on-demand regarding those areas in which they need the most assistance.
- A few specific requests:
 - i. Create an online resource that documents which institutions have readers of obsolete file formats.
 - ii. Provide lists of tools for various tasks and contacts at institutions using th documentation site.
 - iii. Provide a template for a digital preservation plan.
 - iv. Quick / High Level Digital Preservation FAQ - what it is/isn't and why it is important.

Thank you for your time!

Devin Becker

dbecker@uidaho.edu

On behalf of the Digital Preservation Working Group