

CoSA & Preservica Practical Digital Preservation 2017



Digital Preservation 101: State Archives and Agencies Putting Digital Preservation into Action Part 2

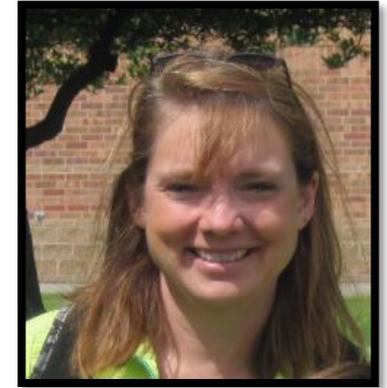
February 28, 2017

Practical Digital Preservation 2016/17

Welcome!

PDP Briefings

- Protecting and Preserving Long-Term Digital Information January 24
- For IT Professionals & practitioners
- The Governance of Long-Term Digital Information May 23
- For Senior Managers & Budget Administrators



Sarah Grimm
Wisconsin Historical Society

PDP “Hot Topic” Webinars Tuesdays 2-3pm Eastern

- Preserving and Protecting Audio-visual Files April 11
- Preserving Digitized State Government Records May 9
- Best Practices in Digital Preservation: International Perspective June 13

PDP Online Workshops - Digital Preservation 101:

‘State Archives and Agencies Putting Digital Preservation into Action’

- Part 1: Practical Training in the Key Concepts February 14
- **Part 2: Practical Training in the Key Concepts February 28**
- Part 3: Case Studies March 14

Sign up today on the CoSA website – PERTTS Portal > Education -Training

Today's Presenters



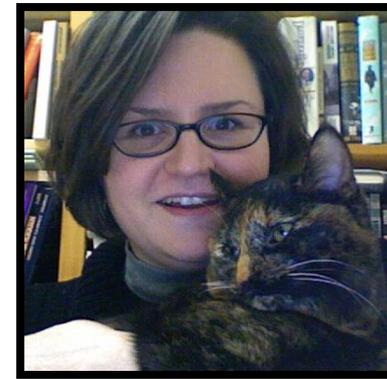
David Portman
Preservica



Lori Ashley
Tournesol Consulting



Tim Hodge
Preservica



Bonnie Weddle
New York State Archives

Council of State Archivists (CoSA)

- 56 state and territorial archives
 - Preservation of and access to records of enduring value
 - Efficient management of government records

<https://www.statearchivists.org/>



Your State Archives

- Records scheduling
- Advice and training
- Transfer of records of enduring value
- State-specific procedures and best practices

<https://www.statearchivists.org/connect/resources-state/>



The screenshot shows the CoSA website's 'Directory of State Archives' page. The header features the CoSA logo (Council of State Archivists) and navigation links: CONNECT, PERTTS PORTAL, PROGRAMS, RESOURCES, and SHRABS. A secondary navigation bar includes ABOUT, MEMBERSHIP, SPONSORS, CAREERS, BLOG, FORUM, SHOP, CART (0), CONTACT, LOGIN, and a search box labeled 'Search this site'. The main content area is titled 'Directory of State Archives' and includes a search bar with a 'Keyword Search' input field, a 'Filter' button, and a 'Show All' button. A left sidebar under the 'Connect' heading lists: Directory of State Archives, CoSA Committees & Task Forces, Volunteer Opportunities, Events, and Listservs.

Agenda



- Introduction
- Retention and Preservation Strategies and Tools
- Taking Action to Protect Long-term Records: Fostering Collaboration with IT
- Summary & Preview of Part 3
- Q&A



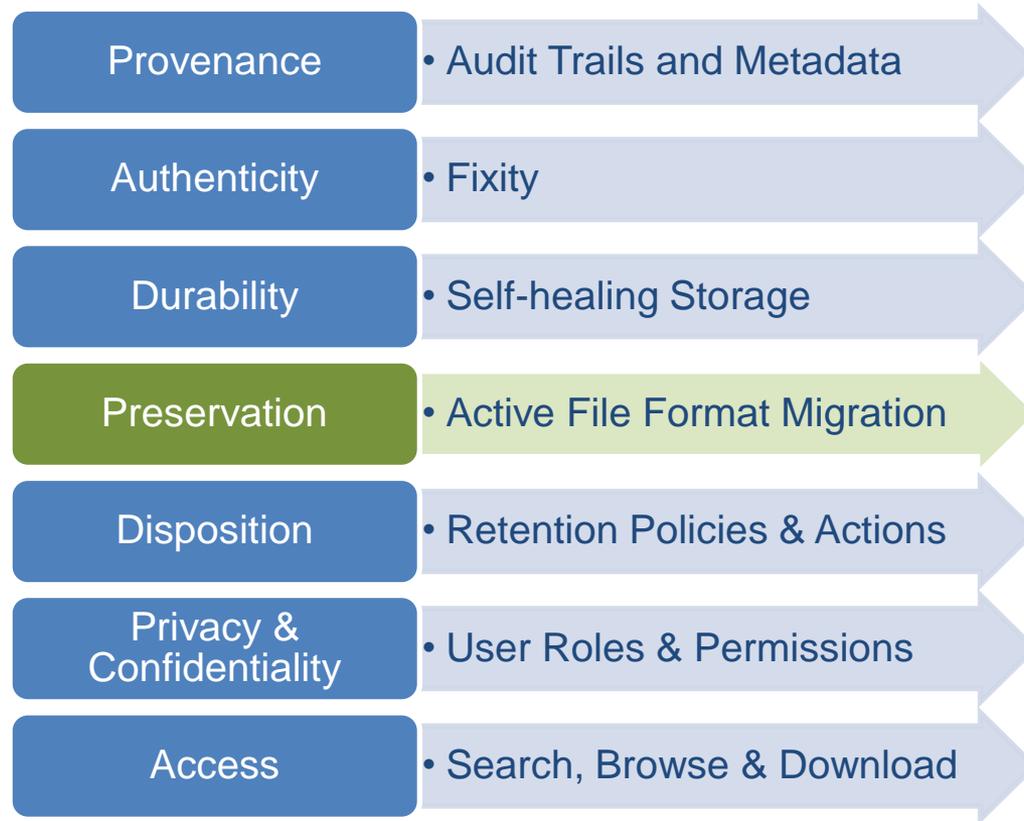
Retention and Preservation Strategies and Tools

Lori Ashley & Tim Hodge

Recap Part 1: Long-Term Record Requirements

- Long-term electronic records (10+ year retention) must be actively managed
- Start as early in the lifecycle as possible, ideally when planning to create records and/or configuring systems
- Decisions include formats and essential documentation about the systems and information objects (metadata)
- Integrate into IT infrastructure and architecture wherever possible
- Specialized expertise required for addressing the scope, scale and diversity of long-term records and metadata

Long-Term Digital Archiving and Preservation



Retention Schedule Examples: Permanent Records

STATE AGENCY RECORDS RETENTION SCHEDULE		
Series	Records Title and Description	Function and Use
U0243	Capital Construction Disbursements and Receipts File	This series documents the disbursement of undertaking which is to be financed or funded by federal funds; (b) any undertaking which is a construction item, or combination of items resulting in the acquisition of equipment estimated to cost \$200,000 or more; or (c) any item of movable equipment if the total cost of purchase, upgrade, or replacement of an information system exceeds \$200,000. The Capital Planning Advisory Board of the State is consulted as to state spending for capital projects.
	Access Restrictions	None
	Contents	Series may contain: Date of progress payments; it also may contain supporting documentation.
	Retention and Disposition	Retain permanently in creating unit

Retain permanently
in creating unit

Retain in Agency
until disposition
trigger; transfer to
State Archives

04560	Governor's Executive Order File - Miscellaneous - (Includes employment contracts, court orders, non-merit appointments, paroles, re-organizations, special notaries) (V)	This series documents the written action or order of the Governor, which has the force of law, and which is based upon his authority as the chief executive of the Commonwealth. The state's Constitution requires that the Secretary of State receive the original order signed by the Governor. After signature by the Secretary of State, the state seal is then affixed. This series incorporates orders categorized as miscellaneous. These include Writs of Habeas Ad Testificandum/Ad Prosequendum, court orders, employment contracts, non-merit position appointments, conditional releases, paroles, special commission notaries, re-organization of executive branch offices, activation of the National Guard, resignations, and removals from office. The orders are initially created by the General Counsel, Office of the Governor. Much of the reference activity to Executive Orders is by courts requiring certification of the official acts of the state. This series reflects original signatures not included in the Executive Journal (04564).
	Access Restrictions	None
	Contents	Series contains: Original order; supporting documentation
	Retention and Disposition	Retain in Agency four (4) years after the end of the gubernatorial term; transfer to State Archives Center for permanent retention

Permanent Records: From Governor's Office to State Archives

The screenshot shows the Texas Digital Archive website. At the top left is the logo for the Texas State Library and Archives Commission, featuring a stylized building and a star. To the right of the logo, the text reads "TEXAS STATE LIBRARY AND ARCHIVES COMMISSION" and "TEXAS DIGITAL ARCHIVE". A blue button with a magnifying glass icon and the text "explore archive" is located in the top right corner. Below the header, there is a navigation breadcrumb "Home > Root" and a search bar with the placeholder text "Search Archive" and a magnifying glass icon. A instruction line reads: "Click on an image to enter the next level of the archive, click on 'description' for more information about that level". Below this are four main content cards, each with a red border and a document icon in the top right corner. The first card shows the Seal of the State of Texas and is titled "Texas State Agencies" with a "description" link. The second card shows the Seal of the Texas Judicial Branch and is titled "Texas Judicial Branch" with a "description" link. The third card shows a collection of framed prints and photographs and is titled "Prints and Photographs" with a "description" link. The fourth card shows a portrait of Texas Governor Rick Perry and is titled "Texas Governor Rick Perry" with a "description" link.



Jelain Chubb
Texas State Archivist

“We as the Texas State Library & Archives Commission have a statutory responsibility to collect, preserve and make records accessible for the future”

<https://www.tsl.texas.gov/texasdigitalarchive>

Four Preservation Strategies You Can Use

1. Preferred File Formats
2. Metadata Management
3. Content Integrity Checks
4. File Format Migration

Strategy 1: Preferred File Formats

- Characteristics of good preservation formats include:
 - Public and open documentation
 - Non-proprietary
 - Widespread adoption
 - Can be opened, read and accessed using readily- available tools
- Examples of Open Standard/Neutral Formats:
 - HTML, Plain Text, XML, ODF and PDF/A for text
 - PDF/A, PNG and TIFF for scanned images
 - CSV for spreadsheets
 - SVG for vector graphics

Strategy 2: Managed Metadata



Published online before print
 March 18, 2013, doi:
 10.1177/0165551513480107

Journal of Information Science March
 18, 2013 0165551513480107

(10) Patent No.: US 6,539,396 B1

Major Land Resource Area coverages



Continental United States	Metadata (PDF; 0.1 MB)
Alaska	Metadata (PDF; 0.1 MB)
Hawaii	Metadata (PDF; 0.1 MB)
Puerto Rico	Metadata (PDF; 0.1 MB)



Unique Identifiers can also store descriptive data
 EX: MFM121587

Stable URL: <http://www.jstor.org/stable/3051911>

Identification_Information

Citation Information
 Originator: Dean Oman
 Publication Date: 2002
 Title: Cartographer
 Publication Information
 Publication Place: Washington DC
 Publisher: USDA/NRCS/NHQ/RAD
 Other Citation Details: NRI Map Room Coverage
 Online Linkage: https://prod.nrcs.usda.gov/Internet/FSP_DOCUMENTS/nrcs143_013066.zip

Description
Abstract
 Major Land Resource Area (MLRA) boundaries for Puerto Rico and the U.S. Virgin Islands.

Purpose
 This coverage is intended as a basemap for US national level natural resource analysis.

Supplemental Information
Procedures Used
 This coverage contains the Major Land Resource Area (MLRA) boundaries for the U.S. Virgin Islands and Puerto Rico. MLRA boundaries were revised in the mid-1990s from earlier versions established in the 1970s and 1980s by the U.S. Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS), formerly the Soil Conservation Service (SCS). The revised MLRA boundary coverage was digitized at the Soil Survey Center in Lincoln, Nebraska, the

Strategy 3: Integrity Checks

- A checksum is a mathematical algorithm run on/against a file and its resulting value
- The resulting “fixity” value uniquely identifies the file based on its content down to the bit level
- The resulting value is called:
 - Fingerprint
 - Hash Value
 - Cryptographic Hash Function
- A checksum can be recalculated and compared at any time to determine if there have been changes



Strategy 4: File Format Migration

- File format is characterized and captured in digital preservation systems
- Original record in native format is retained as the preservation master
- Records are transformed from one file format to another as needed and the metadata captured
- Guards against file format obsolescence

DEMO: Format Migration Workflow



Long-Term Records Management Partners

Archives

- Appraisal support
- Expertise in digital preservation techniques
- File format monitoring
- Access rights
- Transfer processes and procedures
- Search and retrieval
- Best practices & guidance
- Education & training

Information Technology

- Apply conversion and migration techniques
- Document records destruction and transfer
- Metadata management
- Interoperability protocols
- Integrity checks and chain of custody
- Redundancy in storage architecture
- Data security

Taking Action to Protect Long-Term Records: Collaboration with IT

Bonnie Weddle



New York State Archives

Records Management Is Collaborative

- We build relationships with
 - Records creators
 - Senior management
 - Legal counsel
 - State archives or records management agency
 - Information technology

“But I’m not tech-savvy!”

- You don’t have to become an IT expert
- Include records creators in the conversation
- You can always ask questions
- The Internet is your friend

IT Speaks a (Slightly) Different Language

Term	Information Technology	Records Officers and Archivists
Record	A complete item in a database	Any item, “regardless of physical form or characteristics,” that meets the definition outlined in relevant statute or regulation
Archive	Move to less frequently used storage medium (e.g., tape)	Transfer to state archives -or- Active preservation care
Permanent	About 10 years	Forever

Need to Emphasize Benefits to IT

- Increased efficiency
 - Information readily located and retrieved
- Reduced costs
 - Records transferred / deleted at the appropriate time
 - Obsolete and duplicate information destroyed
 - Appropriate storage choices made
- Reduced risk
 - Protection of mission-critical information
 - Protection of information of long-term value
 - Legal and regulatory compliance
 - Litigation, audit, and investigation

2017 State CIO Priorities

1. **Security and Risk Management**
2. **Consolidation/Optimization**
3. Cloud Services
4. **Budget, Cost Control, Fiscal Management**
5. Legacy Modernization
6. **Enterprise IT Governance**
7. **Data Management and Analytics**
8. Enterprise Vision and Roadmap for IT
9. Agile and Incremental Software Delivery
10. Broadband/Wireless Connectivity

<http://www.nascio.org/topten>

“We’ve Got Storage and Backup”

- Storage and backup characteristics
 - Wide range of systems and practices
 - Multiple copies
 - Bit-level storage
- Challenges
 - Locating needed information
 - Ensuring appropriate disposition
 - Records that must be kept for >10 years



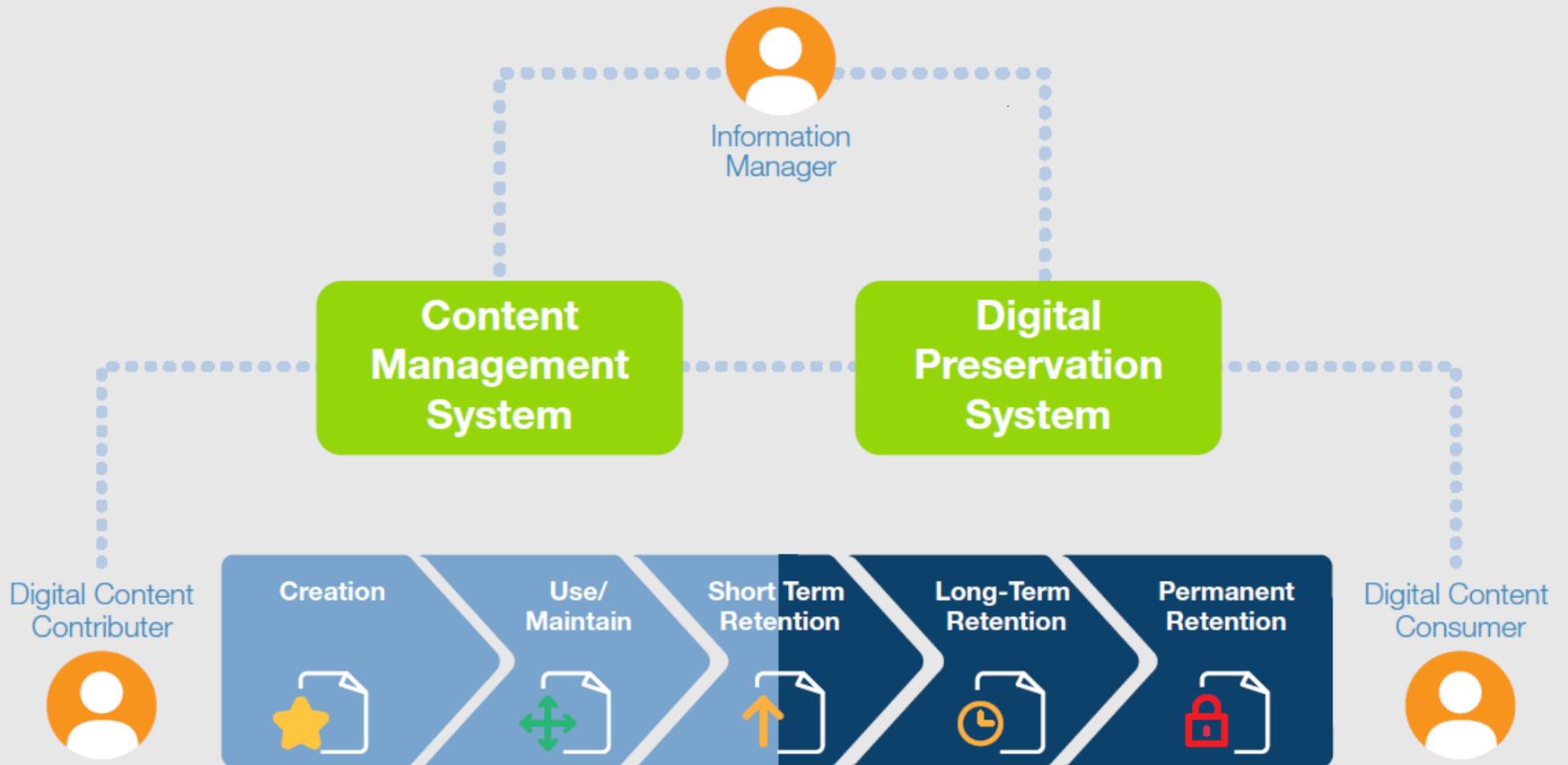
“We Have an ECMS”

- Electronic Content Management System characteristics
 - Widely accepted strategies, tools, and practices
 - Content capture and management
 - Access mechanisms and controls
 - Version control
 - Bit-level storage
 - May offer records management functionality
 - DoD 5015.2 compliance
- Challenges
 - Records management modules add to cost
 - Records that must be kept for >10 years

“What’s a Digital Preservation System?”

- Key characteristics
 - Long-term availability, usability, and trustworthiness
 - Protection against file format obsolescence
 - Integrity safeguards
 - Preservation metadata
 - ISO 14721 (Open Archival Information System Reference Model) compliance
 - Optimal environment for records that must be kept for >10 years
 - Moving toward integration with ECMS
- Challenges
 - Making the case to IT and senior management

Enterprise Content Lifecycle



Getting Your Foot in the Door

- Keep emphasizing
 - Increased efficiency
 - Risk and cost reduction
 - Storage/backup and ECMS don't support long-term preservation
- Keep looking for potential allies
 - Information security officer
 - Chief data officer
 - Past partnership with records management
 - Chaotic freedom of information or e-discovery experience
 - Senior management
 - Legal counsel
- Contact your state archives

Potential Collaborations

- Develop standards and guidance
- Identify information with high, long-term value
- Update retention schedules
- Develop disposition workflows
- Support preservation systems design and procurement
- Raise awareness of electronic records risks and benefits

Summary

- Effective records management through time, systems and custodians
- Preservation Strategies to use for long-term electronic records:
 - File integrity checks
 - Preferred file formats
 - Robust metadata management
 - Periodic transformation of electronic records
- Minimizing barriers to cross-agency records transfer will improve collaboration and reduce duplication of effort
- Capabilities for long-term digital preservation that are integrated into public sector IT infrastructure and architecture should result in significant economic savings.

PERTTS Portal

Education-Training

SERP Framework

Related/Additional Framework Topics

Electronic Records Resources

Glossary

Bibliography

Case Studies and Examples

PERTTS Portal

Welcome to the Program for Electronic Records Training, Tools, and Standards - the PERTTS Portal.

Education - Training

Webinar and training opportunities available through CoSA. Information on and links to opportunities from other organizations are also included.

SERP Framework

The State Electronic Records Preservation (SERP) Framework provides information and guidance on 15 different areas on how to move forward in areas associated with digital preservation. The SERP Framework is based on the Digital Preservation Capability Maturity Model which states are using in the CoSA requested Self-Assessments.

Electronic Records Resources

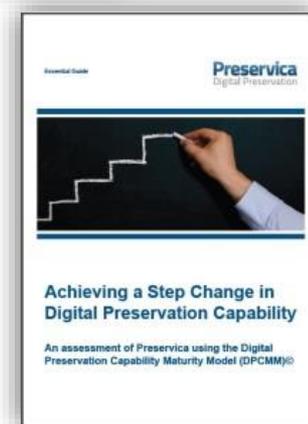
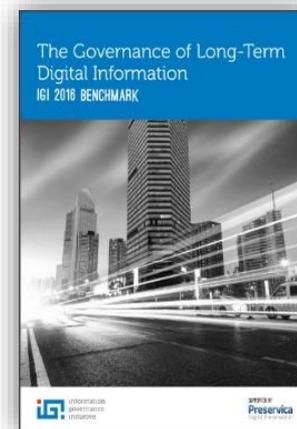
Searchable database of resources including tools, standards, policies/guidelines, and other informational guides related to digital preservation and records management topics can be found in the CoSA Resource Center.

* Add resources! **Login to submit resources** that you find most helpful. When ready to add a resource, click the Add Resource button on the Resource Center Home page.

CoSA thanks the National Historical Publications and Records Commission and Preservica for their substantial support of the PERTTS Portal.

Next Steps & Resources

- CoSA/Preservica Practical Digital Preservation Program:
<https://www.statearchivists.org/pertts/education-training/cosa-preservica-practical-digital-preservation/>
- Preservica Resources Center
<http://www.preservica.com/resources/>
- Preservica Webinars
 - Live Demo: March 2 @ 10am Eastern
Register at www.preservica.com
- Upcoming Events
 - CoSA: Boise, ID (10-13 July)
 - NAGARA: Boise, ID (12-15 July)
 - SAA: Portland, OR (23-29 July)



Preview of Part 3 – Case Studies

Tuesday March 14 @ 2pm Eastern

- Introductions & Recap
- State Archive Use Cases
 - *Examples from Arkansas & Alabama State Archives*
- *Resources*
 - *CoSA Digitization Guidance*
 - *Tools & Resources for your agency*
- Next Steps
- Summary
- Q&A

Questions?



Thank You!

www.preservica.com

info@preservica.com

@preservica

@dPreservation

www.statearchivists.org/