

CoSA & Preservica Practical Digital Preservation 2017



Protecting & Preserving Long-Term Digital Information for IT Professionals and Practitioners

Tuesday January 24, 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

Council of State Archivists (CoSA)

- 56 state and territorial archives
 - Preservation of and access to records of enduring value
 - Efficient management of government records
- Group Initiatives
 - **State Electronic Records Initiative (SERI)**
 - Continuing Education and Training
 - Digital Preservation Self-Assessments
 - PERTTS Portal (<https://www.statearchivists.org/pertts/>) for sharing electronic records and digital Preservation Resources
 - **Archives Collaborating and Cooperating with External Strategic Stakeholders (ACCESS)**
 - **Intergovernmental Preparedness for Essential Records (IPER)**
 - Online courses developed by CoSA and sponsored by FEMA

Today's Presenters



David Mayer
State Historical Society
of North Dakota



Camille Tyndall Watson
State Archives
North Carolina



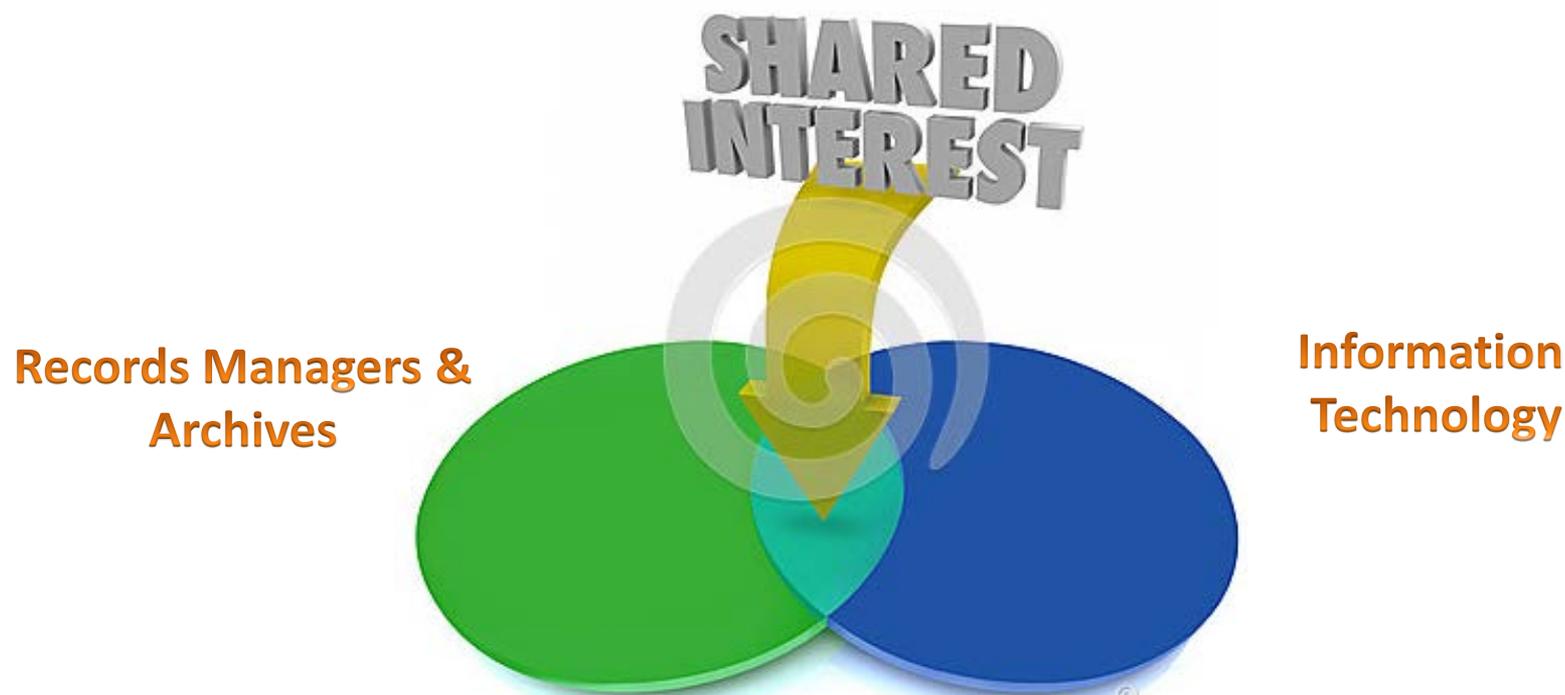
Michael Hope
Preservica

Agenda

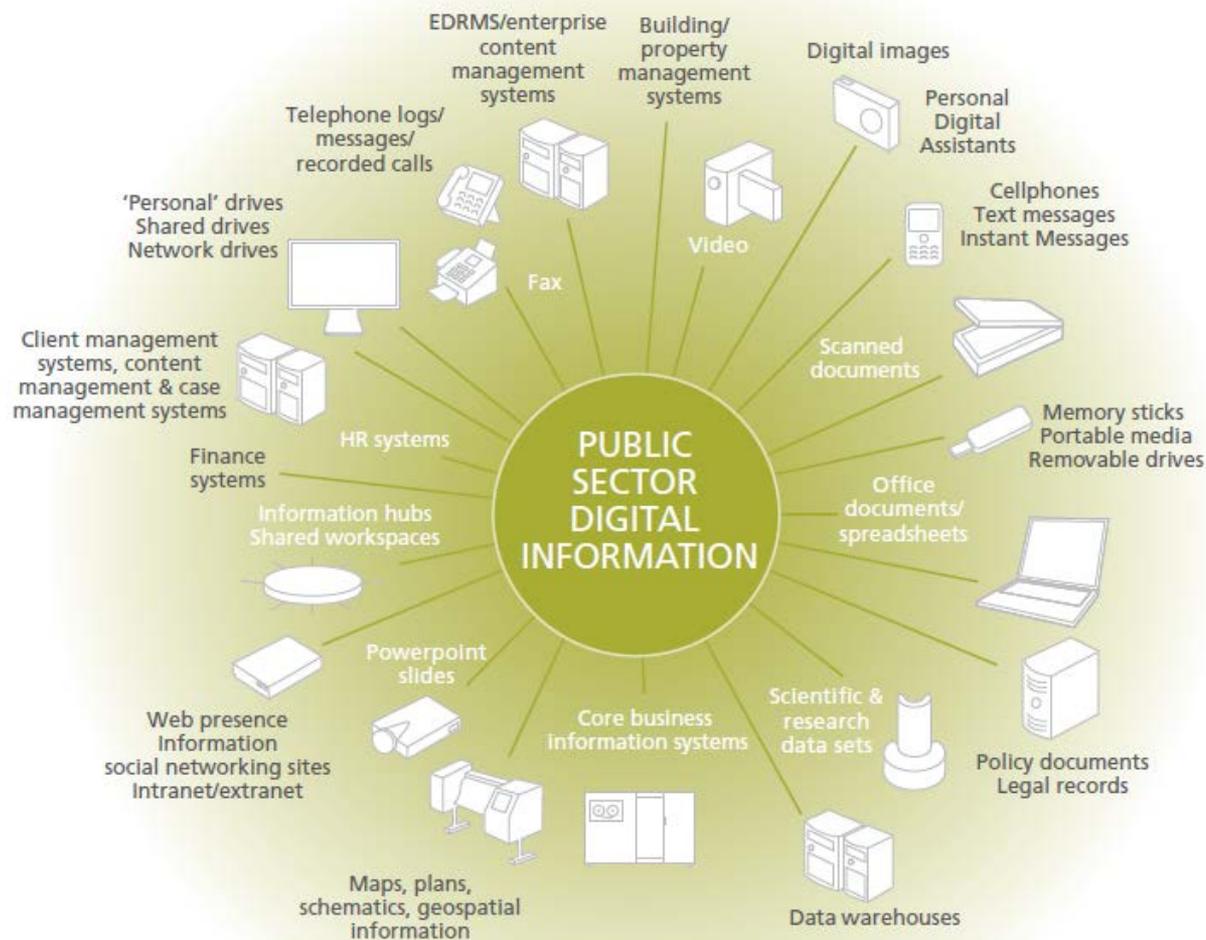
- State of Electronic Government Records
- Protecting Long-Term Digital Information
- Case Studies
- Summary & Next Steps

Records World Meets Data World

Information created, received and maintained as evidence and as an asset by an organization or person, in pursuit of legal obligations or in the transaction of business (ISO 15489-1: 2016)



Public Sector Digital Information



Archives New Zealand
Digital Continuity Action Plan

Challenges of Government E-Records

- Variety of tools being used to conduct the business of state government
- Growing emphasis on transparency and accountability
- General lack of really solid tools
- Increasing volume of records created in or being converted to electronic formats (mass digitization projects)
- Old/proprietary formats and media

Records Management

- Efficient and systematic control of the creation, receipt, maintenance, use and disposition of records
- Processes for capturing and maintaining evidence of and information about business activities and transactions

ISO 15489-1: 2016

Records Schedules

- Identify each category of records an agency creates
- Explain value of each category
 - Business, legal, historical
- Contain disposition instructions for each category
 - Permanent records
 - Kept by agency or transferred to state archives
 - Non-permanent records
 - Kept for a few months, a few years, or a few decades
- Specify special instructions
 - Restricted, PII, HIPPA, FERPA
- May be tied to other data classification schemes

Benefits of Records Management for IT

- Reduced costs
 - Information readily located and retrieved
 - Mission-critical information identified
 - Information with long-term value identified
 - Records transferred / deleted at the appropriate time
 - Obsolete and duplicate information destroyed
- Reduced risk
 - Legal and regulatory compliance
 - Litigation and investigations

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>

Protecting Long-term Digital Information

Michael Hope
Preservica

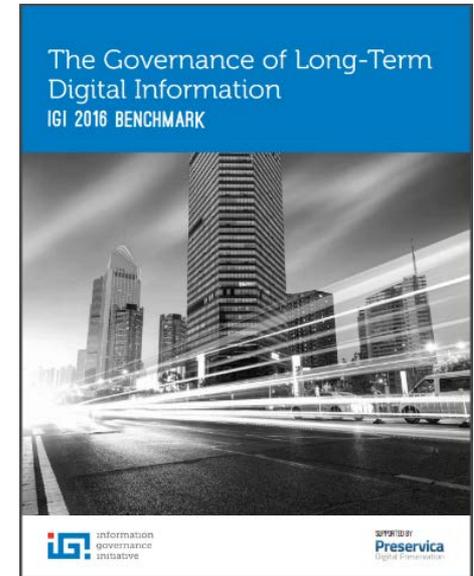
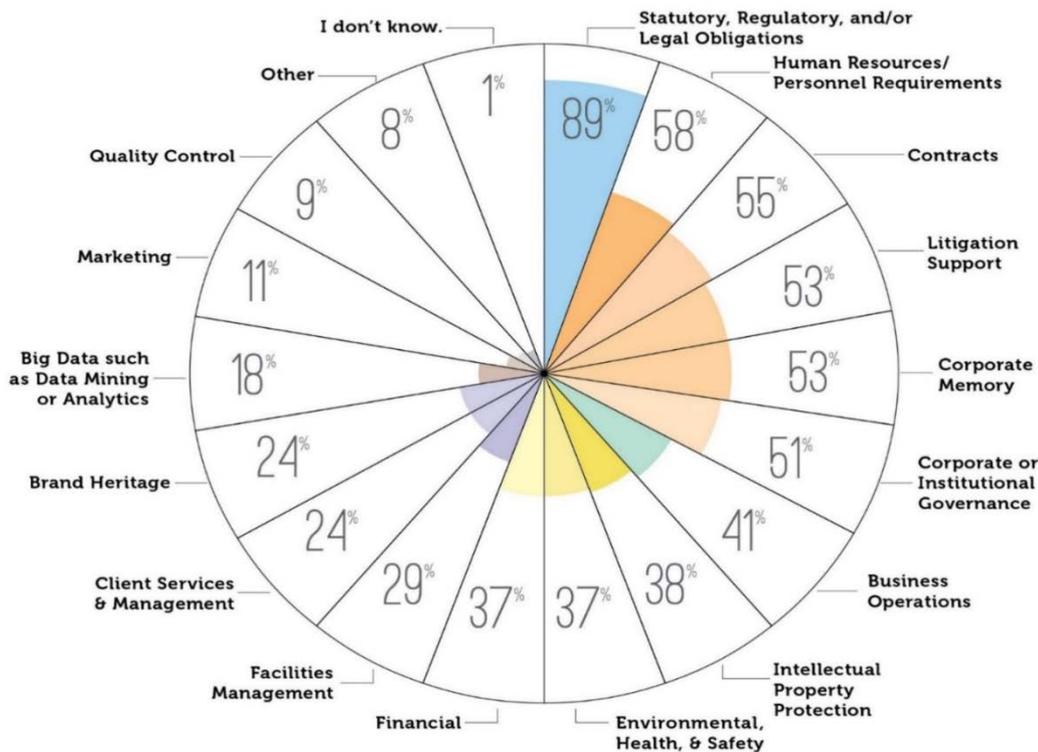
IGI 2016 Benchmark

“The Governance of Long-Term Digital Information”



MOST ORGANIZATIONS HAVE DIGITAL RECORDS AND INFORMATION
THEY KEEP LONG TERM BECAUSE OF THEIR IMPORTANCE

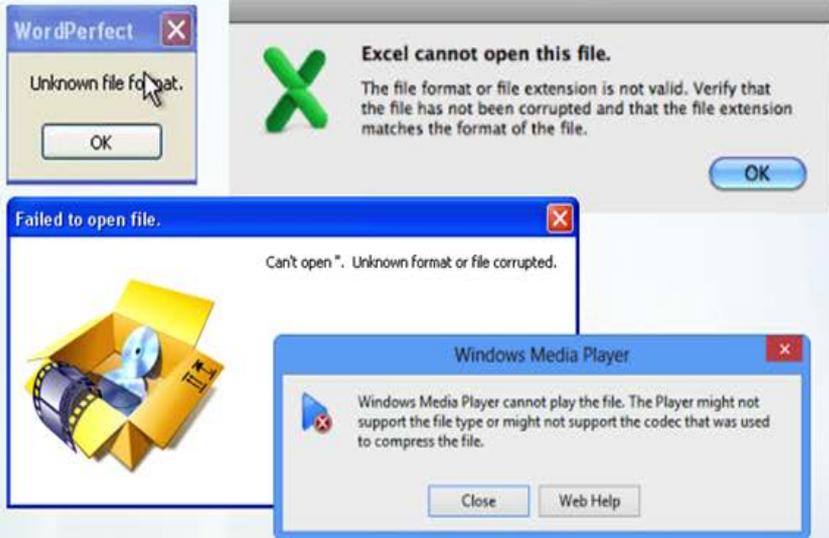
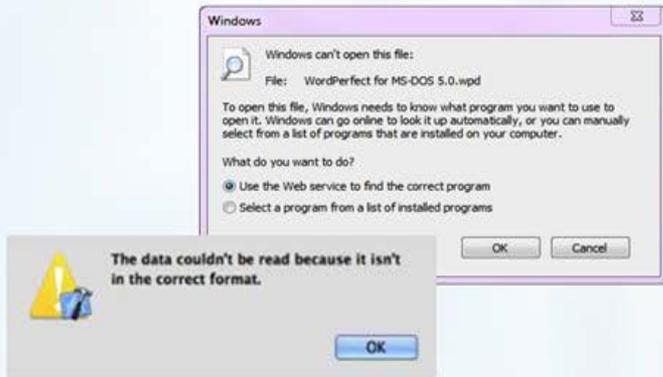
Organizations Report a Variety of Reasons Why They Keep Digital Information



<http://preservica.com/wp-content/uploads/sites/3/2016/05/The-Governance-of-Long-Term-Digital-Information-IGI-Benchmark-2016.pdf>



Digital Content older than 10 years is **at risk**



98%
Yes

Organizations that need to keep digital records and information for the long-term because of its importance

97%
Yes

Information Managers are aware that technology obsolescence puts long-term digital records at risk

Source: Information Governance Initiative, April 2016

CURRENTLY USED STORAGE SOLUTIONS ARE PUTTING LONG-TERM DIGITAL RECORDS AND INFORMATION AT RISK

WHERE ARE DIGITAL RECORDS AND INFORMATION BEING STORED?	
Shared Network Drive	68%
Line of Business Applications (e.g. CRM, ERP, Manufacturing, HR Systems, etc.)	52%
Enterprise Content Management System (ECM)	47%
Disk or Tape Backup Systems	44%
Records Management System (e.g. EDRMS)	43%
Application-specific Archiving (e.g. email)	33%
Removable Media (e.g. CD or USB)	22%
Enterprise Information Archiving System (EIA)	14%
Purpose-built, Long-term Digital Preservation System	11%
Other	9%
Commodity Cloud Storage (e.g. Amazon)	8%
I don't know.	1%

Data is derived from The Governance of Long-Term Digital Information: An IGI 2016 Benchmark. More information at www.iginitiative.com © 2016 Information Governance Initiative.
 © 2016 IGI. Licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. <https://creativecommons.org/licenses/by-nd/4.0/>

Additional Threats to Digital Content

- Limited budgets
- Lack of specialized expertise
- Insufficient or outdated technical infrastructure
- Format obsolescence
- Creator mismanagement or loss of materials
- Insufficient metadata
- Bit rot or other forms of media corruption

Future-proofing Digital Content is a Priority

“many government bodies and industries must retain data for an individual's life span, perhaps 100-plus years, or indeed forever as might be the case for historical records. These time frames require a different and more systematic approach to preserve both the accessibility and the value of these records. This is the emerging market for digital preservation”

***Market Guide for Cloud-Based
Preservation in Government
Neville Cannon***

“As formats change, software is retired and hardware becomes obsolete, **the data that organizations might want to keep can be lost forever.**”

Gartner[®]

“Long-term retention of digital content means preservation issues must be addressed. **Hardware, software, and file format obsolescence risks will haunt us if not taken seriously**”

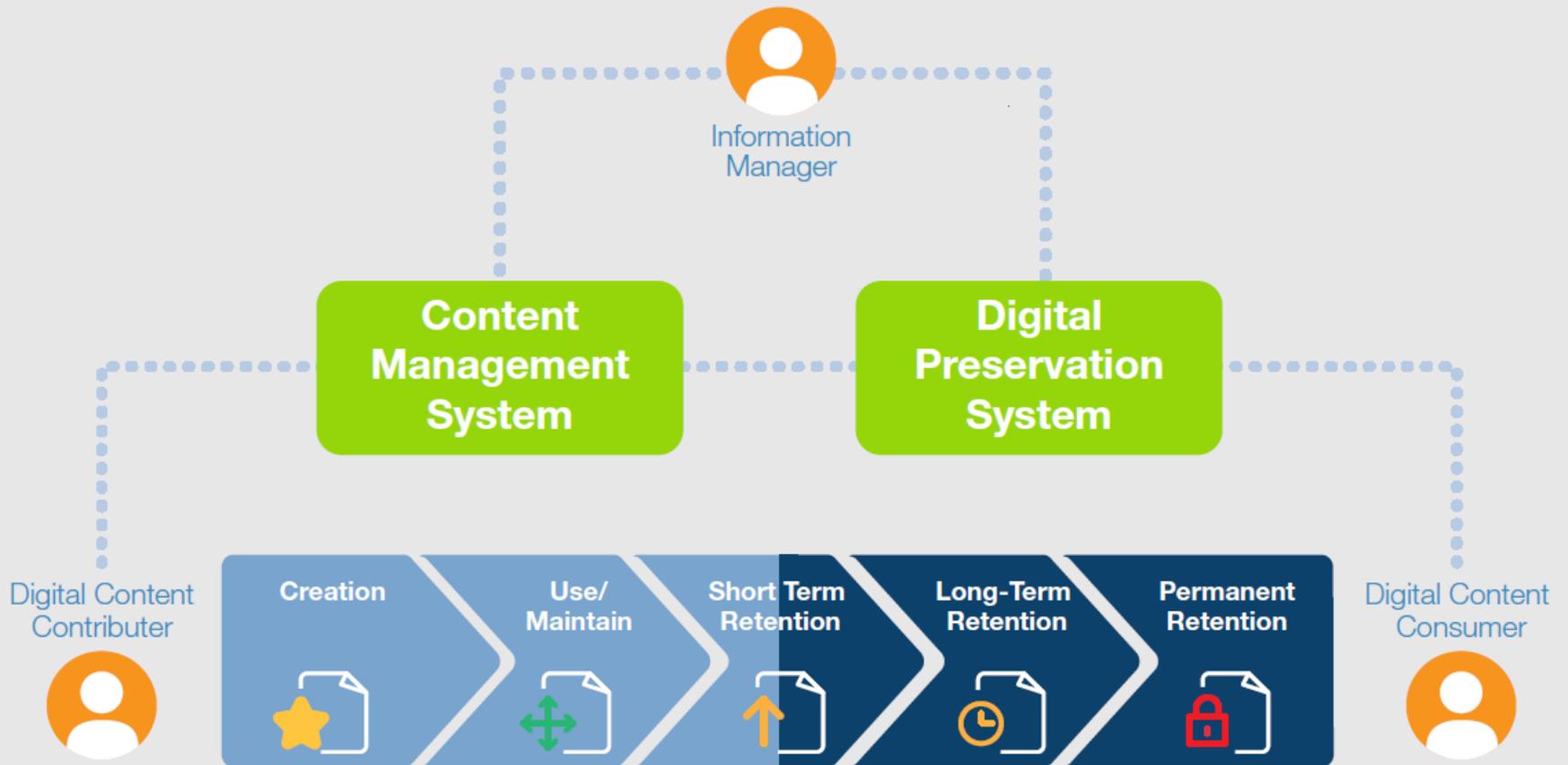
FORRESTER[®]

Digital Preservation System : Requirements



Enterprise Content and Records Management	System Storage and Backup	Digital Preservation System
Accepted Industry Practices or DoD 5015.2 (DoD Records Management) Compliant	Accepted Industry Practice	Conforms to ISO 14721, Open Archival Information Systems (OAIS) reference model
Content collaboration (retrieval), governance, policy, version control, self-healing, retention and disposition	Storage using accepted industry tools, including tiered storage management, multiple copies, and self-healing	Long-term availability, usability, and trustworthiness of digital content
Bit level storage	Bit level storage	Bit level storage, protection against file format obsolescence, integrity protection, preservation metadata
Short-term retention (< 10 years)	Long-term retention (> 10 years)	Long-term and permanent retention (> 10 years)

Enterprise Content Lifecycle



Benefits of Active Digital Preservation for IT

- Builds lifecycle management considerations into system planning and development
- Orderly and efficient records and data disposition
- Frees up IT resources
- Promotes IT as a strategic enterprise capability
- Supports FOI search and retrieval requirements
- Technology watch capabilities help mitigate technology obsolescence



Forging Relationships & Managing Expectations

David Mayer
Information Technology Administrator
State Historical Society of North Dakota



HISTORY FOR *everyone.*

North Dakota's 63rd Legislature

(January 2013)

- Passed SB 2021 authorizing a one-time \$100,000 appropriation for a Digital Archiving Study.
 - Through a competitive procurement process Tournesol Consulting was awarded contract to conduct the study. The consultants interviewed multiple agencies, researched available technologies, and examined what other states are doing and reported their findings.
- Authorized 1 additional FTE to SHSND for an Information Technology administrator (me- December 2013)
- Prior to my hiring the agency had no “IT Department”, 1 individual served as desk top support among other duties. IT support was contracted or provided for a fee through ITD.



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

North Dakota's 64th Legislature

(January 2015)

- The results of the study were used to petition additional funding and FTE's
- Authorized 1 additional FTE for Digital Archivist at SHSND
- Did Not Authorize the Digital Records Manager at ITD
- SHSND was able to secure some one-time funding for consultants and network storage (which by law must be purchased and managed by ITD)
- SHSND also secured ongoing funding for support of its network storage.



The Digital Archives Study Report

Consultant's Recommendations

- Develop Digital Preservation Strategy
- Add Digital Preservation Archivist to SHS Staff
- Add Dedicated Electronic Records Analyst to ITD Staff
- Integrate Digital Preservation Domain into Enterprise Architecture Governance
- Adopt and Promote an Information and Records Life Cycle Management Model
- Inventory Digital Information Systems and Map to Record Series
- Update Appraisal and Scheduling Techniques for Electronic Records of Permanent Value
- Migrate Electronic Records Accepted by State Archives to Network Storage
- Proactively Address Archives Ambiguity through Education and Engagement
- Configure and Implement Email Archive
- Initiate Development of Digital Preservation Policy Framework
- Establish Centralized Standards-Based Digital Preservation Repository for State Government Records of Historical Value
- Technical Staffing for the Preservation Environment
- Implement ISO 14721 Conforming Preservation Repository
 - Develop Preservation Environment Specifications
 - Conduct Periodic System Operator and Preservation Environment Audits
 - Ensure Long-Term Storage and Procedural Suitability
 - Implement Security in the Preservation Environment
 - Design and Test Submission Information Packages (Ingest)
 - Institute Technology Watch Program to Mitigate Technology Obsolescence



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

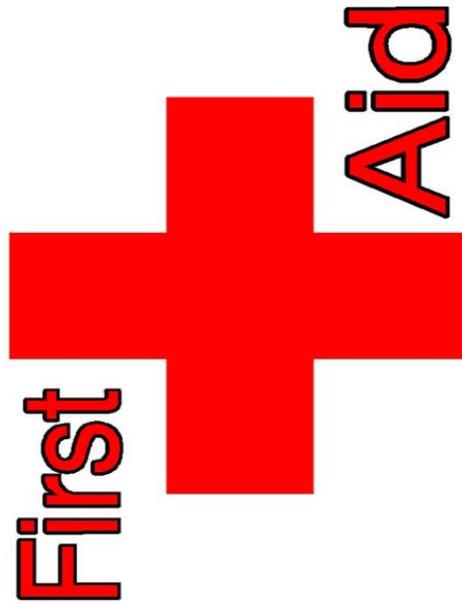
Clock's Ticking – Gotta' Do Something



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

Digital Archivist





STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

Started Building a Rapport With Agencies



So what was I doing while Lindsay was dealing with that?



- *“From Theory to Action: Good Enough Digital Preservation for Under-Resourced Cultural Heritage Institutions”*
 - authored by the Institute of Museum and Library Services and the POWRR group.



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

Digital POWRR

Tool Evaluation Grid

Digital POWRR Tool Evaluation Grid	Ingest				Processing					Access		Storage				Maintenance			Other					
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost

“Trusted” Is an Important Word



Trusted and Secured

- “Trusted” digital repository is not the same as “Secured” network storage.
- “Secured” network storage is a requirement of a “Trusted” digital repository.

IT vs. Archivist

A look at duplicate copy



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

- The “IT” world can make backups, do mirroring, etc.
 - Its processes and scripts ensure data/file integrity.

Trusted Digital Repositories use workflows and metadata to “send” copies to multiple locations, do a fixity/check sum, and document its activity/results for provenance.

Future workflows insure that copy placed in location A is same as copy placed in location B.

- **Building a secure IT environment is the role of ITD**

- enterprise management
- server security
- network security
- redundancies
- data backup, etc.

- **Building a Trusted Digital Repository is the responsibility of the Archivist**

- ISO compliant policies
- procedures
- audit trails
- fixity
- provenance
- governance
- mitigating technology obsolescence.

My role was to serve the needs of the Archivist with a solution set that conforms to the requirements of the IT enterprise.



What are we preserving in our repository?

Documents of permanent or historical value



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

Basically, if it comes to us we keep it forever unless our future selves determine it is no longer relevant / worth keeping.



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

Functional Requirement

(an oversimplified definition)

- If I was subpoenaed and required under oath to attest to the provenance and authenticity of a digital file in my possession could I do it and what documentation / proof do I have to support my testimony.
 - And as an extension of that question can my successors, 5, 10, 50 years from now do the same and will it be available and useable.



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

Do Something Today That
Your Future Self Will
Thank You For



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

Making a Move



HISTORY FOR *everyone.*

Digital POWRR Tool Evaluation Grid	Ingest				Processing				Access		Storage				Maintenance			Other						
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

There Is No Right Answer

$$f(x) = 3x + 4$$

$$x = 3y + 4$$

$$y = \frac{1}{3}x + \dots$$

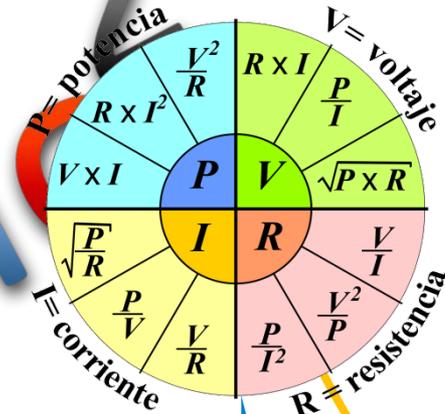
Handwritten mathematical notes on a blue grid background, including:

 $\frac{d}{dx} \int_a^b f(x) dx = f(x)$

 $\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=0}^n \Delta x f(x_i)$

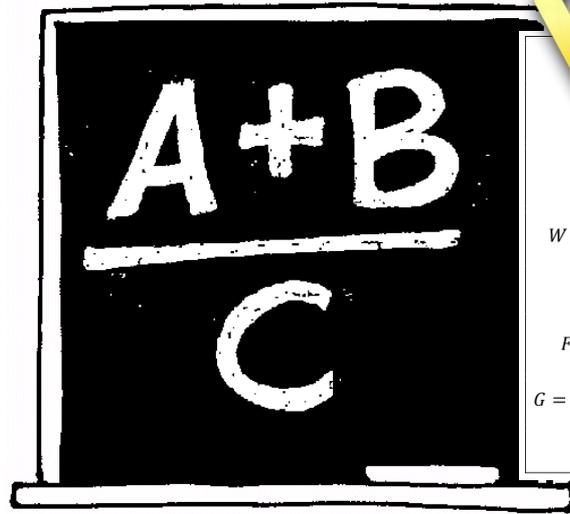
 $\frac{d}{dx} \int_a^b f(x) dx = f(x)$

 $\int_a^b f(x) dx = \frac{1}{n} \sum_{i=0}^n f(x_i) \Delta x$

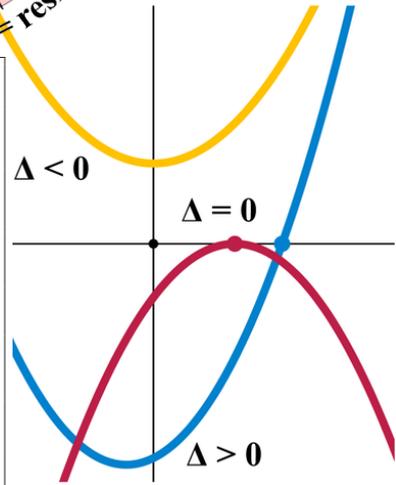


7

Large stylized letters 'E', 'm', and 'h' in yellow, blue, and green respectively, overlaid on the diagram.



$x = vt$	$\Delta x = \frac{1}{2}at^2 + v_0t$	$\Sigma F = ma$	$F_f \leq \mu N$	$\tau = r_1 F = r F_L$
$v^2 - v_0^2 = 2a\Delta x$	$F_c = \frac{mv^2}{r}$	$F_G = \frac{GMm}{r^2}$	$T = \frac{2\pi r}{v} = \frac{1}{f}$	
$E_0 + W = E_f$	$K = \frac{1}{2}mv^2$	$U_g = mgh$	$U_s = \frac{1}{2}kx^2$	$F_s = kx$
$W = F_{\parallel} \Delta x = F \Delta x_{\parallel}$	$P = \frac{W}{t}$	$AMA = \frac{F_{out}}{F_{in}}$	$IMA = \frac{x_{in}}{x_{out}}$	$\epsilon = \frac{AMA}{IMA} = \frac{W_{out}}{W_{in}}$
$\vec{p}_0 + \vec{J} = \vec{p}_f$	$\vec{p} = m\vec{v}$	$\vec{J} = \Delta \vec{p} = \vec{F} \Delta t$	$v_1 - v_2 = v'_2 - v'_1$	
$F_{E,pt\ chg} = \frac{kQq}{r^2}$	$\vec{E} = \frac{\vec{F}_E}{q}$	$\Delta V = \vec{E} \Delta x$	$V_{pt\ chg} = \frac{kQ}{r}$	$\Delta V = \frac{-W}{q} = \frac{\Delta U_E}{q}$
$G = 6.67 \times 10^{-11} \frac{N \cdot m^2}{kg^2}$	$g = 10 \frac{m}{s^2}$	$k = 9 \times 10^9 \frac{N \cdot m^2}{C^2}$	$q_e = 1.6 \times 10^{-19} C$	$1\ hp = 745.7\ W$
$\sin(\theta) = \frac{opp}{hyp}$	$\cos(\theta) = \frac{adj}{hyp}$	$\tan(\theta) = \frac{opp}{adj}$		



HISTORY FOR everyone.

“institutions that have some financial resources available for redirection, but do not expect to add personnel anytime soon”





HISTORY FOR *everyone.*



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*



STATE
HISTORICAL
SOCIETY
OF NORTH DAKOTA

HISTORY FOR *everyone.*

Thank You

David Mayer

Email: damayer@nd.gov

Phone: (701) 328-1490



HISTORY FOR *everyone.*

Building Relationships with IT Staff in NC

Camille Tyndall Watson

Digital Archivist

State Archives of NC



State Archives of North Carolina
NATURAL AND CULTURAL RESOURCES

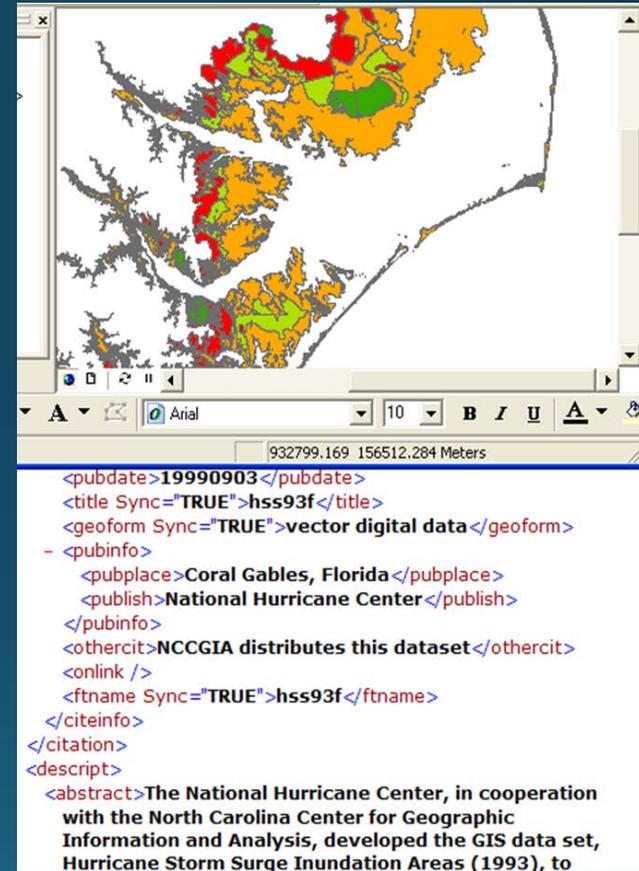
Making Friends and Influencing People

- Cultivating relationships with our business group opened up ways to communicate needs and solve problems between our division and IT
- When other departments understand what you are doing, they can help communicate state-wide needs to IT
- Lesson Learned: Networking is important! Building relationships across the enterprise opens up the door to additional collaborations



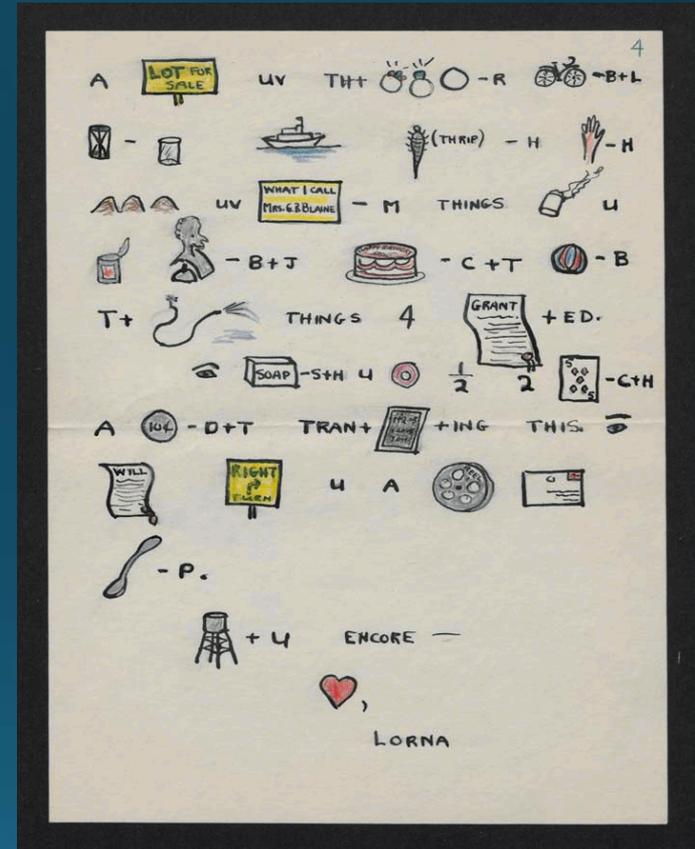
A Seat at the Table

- SANC staff sit on boards and committees for GIS data and archiving, such as the NC GICC and the SMAC
- Result of being open to partnership when reached out to by CGIA and NCSU
- Lesson Learned: Collaborations across non-traditional groups provides new opportunities to solve real-world problems



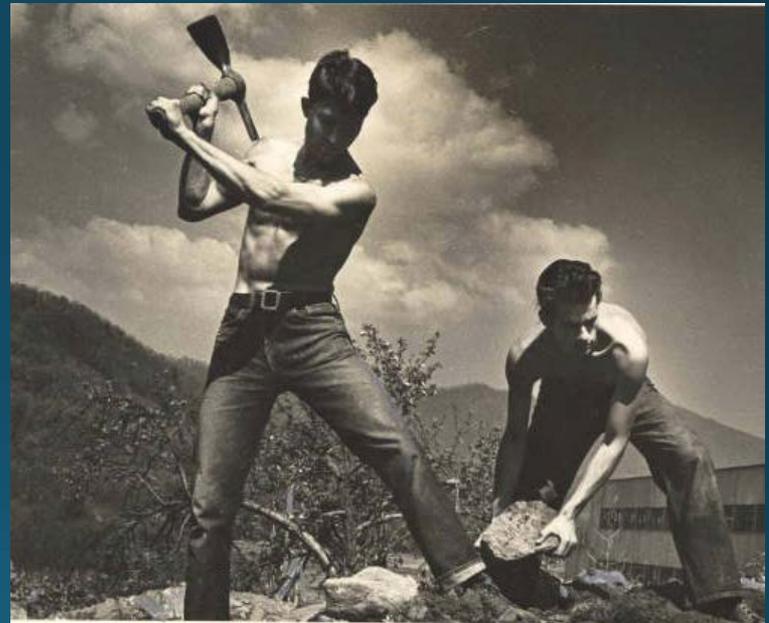
It Can't Hurt to Ask

- SANC and SLNC are collaborating on a Functional Requirements document for state IT regarding our digital repository, which is a collaboration between the institutions
- Developed a testing matrix for digital preservation systems to justify purchase and demonstrate suitability in a quantitative manner
- Lesson Learned: Don't assume that everyone is on the same page. You may need to do education in addition to outreach.
- Lesson Learned (part 2): Create strong partnerships and build on existing expertise



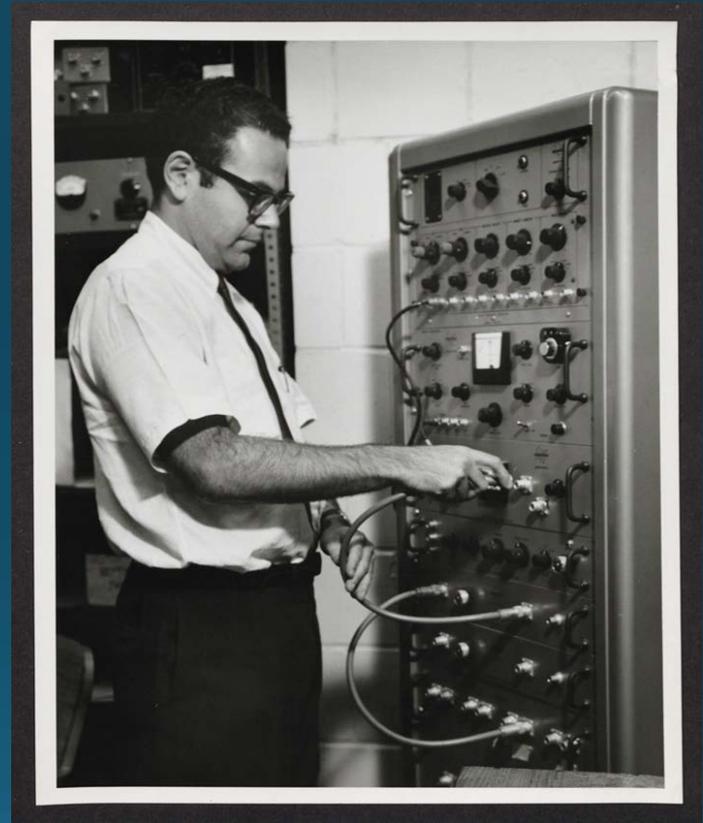
It Can't Hurt to Ask (Part 2)

- Currently, SANC is working on TOMES, an NHPRC grant regarding email archiving
- Reached out to IT and OSHR regarding automated tagging of email accounts by position number...and it worked!
- Lesson Learned: Don't be afraid to directly address your problems with other groups; they may be dealing with the same issues.



Challenges

- Centralized IT
- High Staff Turnover Rate
- IT and Archives can be separated by a common language (but we're a lot closer than you might think!)



Responsibility for Records Management and Digital Preservation

- Records Creators
- Agency Management
- Records Management
- CIO
- State Enterprise Architect
- Domain Managers and System Administrators
- State Archivists
- State Librarians
- Local Government

Potential Collaborations

- Form a community of practice
- Develop framework of standards and guidance
- Identify high-value information that must be kept long-term
- Update retention schedules and disposition workflows
- Support preservation systems design and procurement
- Raise awareness of digital information risks and benefits
- Support information governance initiatives

Summary

- Digital information will pass through a number of technology environments over its lifespan.
- As public records and information become digital, there is a corresponding expectation from the public that those will be easily accessible.
- Collaboration between non-traditional partners and implementation of best practices will help ensure the long-term preservation of digital content.
- Enterprise level digital preservation reduces costs and duplication of effort.

Resources - Reports

- National Association of Secretaries of State (NASS) White Paper:
Harnessing the Power of Digital in State Records Management (Feb16)
<http://www.nass.org/reports/surveys-a-reports-2/>
- National Association of State Chief Information Officers (NASCIO) –
State CIO Top Ten Policy and Technology Priorities for 2017
<http://www.nascio.org/topten>
- Council of State Archivists (CoSA) –
**Archives Collaborating and Cooperating with External Strategic Stakeholders (ACCESS)
Action Plan**
https://www.statearchivists.org/files/3314/5383/8754/CoSA_ACCESS_Action_Plan.FINAL.pdf
- National Association of Government Archives and Records Administrators
(NAGARA)
Archives for Local Governments (2012)
<https://www.nagara.org/Public/Resources/Publications/Public/Resources/Publications.aspx?hkey=b2245c8d-7283-4bd5-a6d3-3a8615202cbe>

Resources - Assessment Tools

- ARMA's Information Governance Maturity Model –
<https://www.lva.virginia.gov/agencies/records/psrc/documents/Principles.pdf>
- National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation
http://www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf
- Digital Preservation Capability Maturity Model (DPCMM)
<https://www.statearchivists.org/resource-center/resource-library/digital-preservation-capability-maturity-model-dpcmm/>

* Register for the Digital Preservation Capability self-assessment survey used by CoSA at www.DigitalOK.org

Next Steps

- Practical Digital Preservation Program:
<https://www.statearchivists.org/pertts/education-training/cosa-preservica-practical-digital-preservation/>
- Upcoming webinars : Register at www.preservica.com
 - LIVE Demo of Preservica v5.7 – **February 16**
 - How Not to Build a Digital Archive: Lessons from the Dark Side of the Force – **February 2**
- Discover the Preservica Resources Center
<http://preservica.com/resources/>
- CoSA PERTTS Portal – Electronic Records Training, Tools and Standards
<https://www.statearchivists.org/pertts/>



Questions?



Thank You!

www.preservica.com

info@preservica.com

@preservica

@dPreservation

www.statearchivists.org/