NASCIO
Representing Chief Information Officers of the States

NASCIO Staff Contact:
Eric Sweden
Enterprise Architect
esweden@AMRms.com

Electronic Records Management and Digital Preservation: Protecting the Knowledge Assets of the State Government Enterprise

Part III: Management Leads and Technology Follows – but Collaboration is King!

Electronic Record as defined by the Society of American Archivists - Data or information that has been captured and fixed for storage and manipulation in an automated system and that requires the use of the system to render it intelligible by a person.

I. Introduction

This research brief concludes the current NASCIO series on the subject of electronic records management and digital preservation. This subject area is very broad and has multiple dimensions, perspectives and challenges for the state IT community. The objective for the series is to highlight some of the key issues and make relevant recommendations to the state CIO. Ultimately, electronic records management and digital preservation must be a shared responsibility with understanding and support from the state CIO. Everyone within state government must play their part in managing the digital assets of the state.

In keeping with the theme of the previous briefs on this topic, NASCIO maintains that electronic records management and digital preservation initiatives must be led and managed within the broader scope of enterprise architecture. A collaborative team made up of enterprise architects, project managers, electronic records managers, state librarians and archivists will ensure state knowledge assets are managed properly using the most appropriate technology. Further, given the role of electronic records in investigation, discovery and litigation, CIOs would be well advised to invite input from their office of attorney general, state auditor and other legal counsel versed in the specific rules of civil procedure. State CIOs carry a very broad scope of enterprise responsibilities and must rely on the specialized knowledge of these professionals.

Who is responsible?

Managing the knowledge assets of state government is a shared responsibility. The creators of records, records managers, and state archivists must partner with the CIO and the state enterprise architect to ensure that electronic records management and digital preservation issues and

NASCIO represents state chief information officers and information technology executives and managers from state governments across the United States. For more information visit www.nascio.org.

201 East Main Street, Suite 1405
Lexington, KY 40507
Phone: (859) 514-9153
Fax: (859) 514-9166
Email: NASCIO@AMRms.com

discipline are integrated with the states investment process, project and program management, and the overall enterprise architecture strategies for managing *knowledge assets* of the state. State librarians should be consulted relative to public access and use of public information.

Does technology help or hinder?

Digital Preservation systems have a simple goal, that the information they contain remains accessible to users over a long period of time.[1]

Beyond the organizational issues are the technological issues related to digital preservation. The advances in technology create *both* opportunities and challenges. The opportunities reside in the new technologies for managing digital assets. The challenge resides in the fact that as existing technology becomes obsolete, older digital assets will become inaccessible if proactive approaches are not taken to sustain them. Related are the challenges to the authenticity and integrity of electronic documents; demonstrating that an electronic document is the original or "official" document. Multiple initiatives are underway nationally and globally to address this challenge. The best resources for staying up to date on the various preservation efforts are the records managers and archivists within state government. These initiatives are researching methods for long term preservation. Long-term preservation of simple data types – plain text or tabular data sets – is relatively straightforward. However, for more complex knowledge assets the jury is still out regarding specific technology solutions. Technology obsolescence is one of the problems these initiatives hope to solve[2] – but the solutions have not been identified yet. However, technology is only part of the solution. More important is the need to develop a preservation strategy.

Public sector electronic records are being created at an alarming rate and the necessary policies and tools for their management are desperately lacking. Many of these records, possibly most, can be *and should be* disposed of fairly quickly. Some records must be kept in a manner making them available for legal, administrative, or research purposes for many years. A few records – usually estimated at three to five percent of all records created – must be kept for the specified retention period, possibly life. The challenge for long-term management of digital information is preserving the digital record of the item (*bit preservation*: assuring that the bit streams constituting the digital objects remain intact and recoverable over the long-term). There must be the capability of finding and retrieving the digital item. Long term preservation may require digital material to be transformed into new formats, or the use of emulators to produce the original appearance and function so that it can be viewed, played or used as intended by the item's author.[3]

The first step is knowing how long to keep certain records. Preserving the bit stream is irrelevant for ephemeral records. *The second step* is now what to preserve. Do you preserve the entire database, either by taking frequent snapshots or by implementing the technology to roll back transactions to the state of the database at any point in time? More often than not, it is sufficient to preserve views of the data as records.

A vital records database might be used to preserve birth and death records that are presented as individual documents, while the original database can be lost. Even GIS information can be preserved as a select set of maps so that someone in the future can easily see how the data in the GIS was rendered for specific purposes. At the same time, the underlying tables can be preserved for import into and manipulation by future GIS systems. This would be a much more doable strategy versus trying to keep the original GIS software up and running for future data analysis.

There have been multiple approaches proposed. These have been succinctly described by Kenneth Thibodeau, director of the National Archives and Records Administration's Electronic Records Archives (ERA) program, as *emulation*, *migration*, and *persistent object preservation*.[4] We'll examine these later in this research brief. At this point, we simply make the point that there is a lot of activity working in parallel to address the challenges

of managing knowledge assets taking account of authenticity, integrity, privacy, security, and availability.

II. The Technology Issues of Digital Preservation

Technology developments are moving so rapidly that the software on which much of the digital record is stored is obsolete, and even extinct. This issue is so significant that a recent survey conducted by RLG[5] reported that technological obsolescence is the single greatest threat to successfully managing digital assets.[6]

Digital technologies are now the norm for creating, distributing, and storing the state record. Many states have written into law that the electronic records are now recognized as legal documents. Attention to electronic records or ERM in the mid 1990s led to the development of electronic records management software that could integrate with other information systems.[7] Many organizations now recognize the importance of records management and preservation and have implemented policy and technology. However, some of these organizations, in trying to manage risk, *save everything*. They correctly recognize that emails are enterprise records if they carry certain characteristics. Other emails don't merit preservation because they are simply the means for communication events trivial to the enterprise and have only temporary value – e.g., *"Let's go to lunch."* However when that same email says, *"Let's go to lunch and talk about the contract with XYZ Consultants"*, it may merit preservation.[8, 9] One consideration for achieving consistent application of records retention and paring out enterprise records from resources like email is to delegate this process to information systems that incorporate expert or artificial intelligence system technology.[10] However, saving everything also has inherent risks.

Government must put in place deliberate strategies for the preservation of the digital record. That strategy will likely include a variety of tactics specific to different groups of records. Generally, the longer the records must be kept, the more difficult it is to preserve the records. Appropriate tactics will balance the costs of preservation against the risks and value associated with information in the records.

A preservation strategy must address preservation throughout the life of a record. The following questions help frame that strategy and evaluation of alternatives:

> *Is the record created in a manner that lends itself to preservation, or does it rely on technology that will be difficult to sustain?*
>
> *Is the record used in a way that protects its integrity during active use, or is it easy for the record to be intentionally or accidentally changed – with the result that it loses it evidential value?*
>
> *Is the record stored in a manner that facilitates its access?*
>
> *Does the storage system promote the proper disposition of records – either destruction or transfer to an archives – and does the system protect records subject to legal holds?*

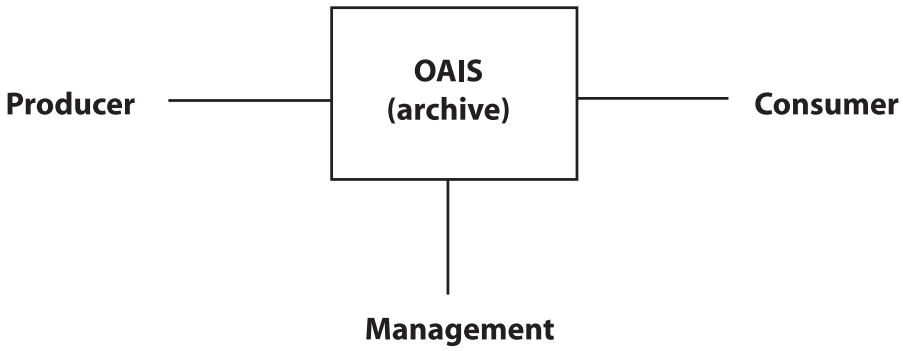
Strategies include:

Migration is the movement of data records from the existing software, and potentially hardware, to new technology. This requires that the data be reformatted with each new iteration of technology. Again, maintaining records over long periods of time presents an escalating problem. Records need to be migrated – and at each migration - there are more records to be handled.[11, 12]

This will be much easier in the future with migration toward a new generation of office applications that use XML as the native file format. New opportunities are then afforded due to digitally stored documents that use XML.

Enterprise Architecture is a management engineering discipline that presents a holistic, comprehensive view of the enterprise including strategic planning, organization, relationships, business process, information, and operations.

Environment Model of an OAIS



Ephemeral Records are records that have a very short term value to the state and have minimal or no ongoing administrative, fiscal, legal, evidential or historical value.[13]

Emulation - this is where the data is archived with the original application, operating system and potentially even hardware. Data can be retained in the original formatting using this approach. This also requires personnel who can either remember or learn to run the original software and hardware on which the records were created. Maintaining records over very long periods of time – centuries – precludes the feasibility of this approach.[11,12]

Persistent object preservation provides a more workable solution for long term digital preservation. In this approach, the record is removed from the computer system that created it. Digital records are preserved, and remain accessible through current technology. Persistent object preservation recreates records in a standard format that are self-describing and readable by newer systems.[11,12,14]

Given the complications of preserving, retrieving and presenting complex digital records there must be deliberate thought regarding what functionality the organization is willing to commit to maintaining in regard to current digital records and archived digital records. An example of a complex knowledge resource is an interactive map showing the change in some aspect such as timber harvesting, incidence of disease, or environmental conditions.

There can be a significant difference in the required investment for the various capability choices available. For example, the difference between retrieving and presenting a preserved knowledge resource with all of its original "look and feel" – *and interactive capabilities* – versus retrieving and presented *essential information* from that resource. There is the issue of maintaining the necessary ongoing investment to ensure the knowledge asset and its characteristics are preserved over the appropriate lifecycle of usefulness. There may be a point in time when certain characteristics are no longer preserved, but essential business information is preserved in a simpler, less expensive form. Notwithstanding this caution, there are some digital knowledge assets that require the preservation of complex interactive capabilities, or animation. Most important is establishing deliberate intent, and following through with a strategy and capabilities for delivering that intent over time. This includes consideration of how records are initially created and stored. These decisions will impact both the economic and the technical feasibility for preservation.

If an archives is to be established, it should be based on a deliberate architecture that includes business intent and outcomes sought, organization, business processes, data architecture, and technology architecture. One of the definitive sources for guidance on archives architecture is the Reference Model for an Open Archival Information System (OAIS). This reference model was developed under the auspices of the Consultative Committee for Space Data Systems (CCSDS). This reference

model addresses a full range of archival information preservation functions including ingest, archival storage, data management, access, and dissemination.[15]

The San Diego Supercomputer Center (SDSC), the University of Maryland (UMD), and the National Archives and Records Administration (NARA) have embarked on a project to develop a technology framework for digital preservation using *data grid technologies*, and to demonstrate these technologies on a pilot archives. This initiative is also funded by the National Science Foundation under the Partnership for Advanced Computing Infrastructure (PACI) Program. Data grid technologies promise to be a significant capability for managing archives across a diverse geography and available for managing a wide variety of knowledge assets including digital records of voice, audio, video, medical records, and scientific research such as genome research.[16, 17] At this juncture, state CIOs should be aware that this research is underway. NASCIO will continue to monitor ongoing progress of this work.

SDSC has already built a prototype using commodity platforms with significant disk caches coupled with heterogeneous tape libraries. Software tools based on grid technologies to organize and manage the resources at the different sites, as well as novel software tools to load data into the archives for long-term preservation have been developed. Data grids enable the creation of shared "collections" that are distributed across multiple organizations and jurisdictions, physical locations, and storage systems. SDSC is also researching methods and process for characterizing and applying management policy for a given shared collection. These characteristics include assessment criteria, access controls, mechanisms for presenting information, and distribution of data across multiple categories.[18]

Examples of digital libraries that build on data grids to manage distributed collections:[19]

Digital Library Applications:

- ***DSpace digital library – MIT and Hewlett Packard*** [20]
- ***Fedora digital library – Cornell University and University of Virginia*** [21]

Digital Repositories (libraries and archives):

- ***NARA research prototype persistent archives***
- ***California Digital Library – Digital Preservation Repository***
- ***NSF National Science Digital Library persistent archives***

Why so many initiatives that are devoted to solving the same issues? *"There is no 'silver bullet' for solving the multi-faceted challenges of records management and digital preservation."*[22] All of these initiatives will provide a part of the solution, or provide a solution that can be implemented now. Such has happened with the state of Kansas's KSPACe implementation which uses Dspace.[23]

The ongoing issues of records management and preservation eventually point to the challenge of *operationalizing* records management principles. Understanding the importance of records management is just the beginning. How to turn that recognition into an ongoing business operation that is efficient and economical will require leveraging technology that is currently under development. The National Archives and Records Administration (NARA) is currently working on research and development to eventually present the capability for managing records over time. This important initiative, if successful, will revolutionize how records are preserved.[24]

Vital records are the essential agency records that are needed to meet operational responsibilities under national security emergencies or other emergency or disaster conditions (emergency operating records), or to protect the legal and financial rights of the Government and those affected by Government activities (legal and financial rights records) - 36 CFR 1236.14

III. The Promise of Extensible Markup Language (XML)

The primary goal in digital preservation is maintaining knowledge assets over time, ensuring it is accessible going forward decades and even centuries. The advent of

digital conversions and "born digital" has brought great capabilities, and also inherent challenges. The issue results from the coupling of content with specific software. Software is coupled with obsolescence, and versioning. Over time, content can become inaccessible because the software version used to create the content is obsolete, and new versions of the software can't access the old content. This occurs with prevalent office productivity software applications such as word processing.

In addition to maintaining synchronism between content and application software, there is the relationship between the application software, the operating system, and the hardware on which it runs. In preserving the accessibility to content, some organizations archive the software *and* the operating system *(and* if necessary the hardware) that will run that software.

Born Digital: Materials originally created in an electronic form and may only exist in electronic form.

XML provides a viable alternative for digital preservation

XML has been identified as a potential solution to this dilemma.[25] Some advantages of XML are:

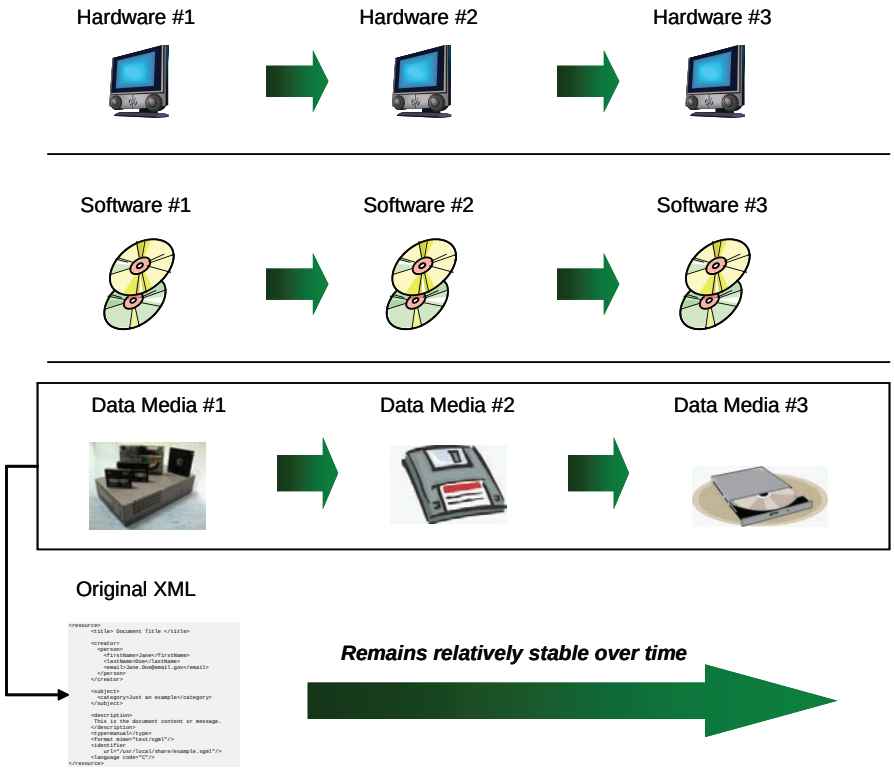
- XML is an open standard established by the World Wide Web Consortium (W3C)
- XML is not tied to any specific platform
- Because XML is a markup language, it provides more than simple text preservation approaches such as ASCII. That is, the formatting of text can be preserved.
- XML can be read by humans as well as machines using a simple text editor
- XML documents can be self-describing. Document structure, field names, and content can be stored in the document. *However, information about the structure of the document may be stored in a separate DTD or schema.*
- XML has been designed to accommodate computer processing
- Many digital files can be converted directly into XML
- Because XML is not tied to any specific platform, it is well suited for digital preservation of text. Conversions going forward will be drastically reduced, if not eliminated, which will help ensure *authenticity* of digital records
- Existing software development environments are XML ready now
- Next generation office applications use XML natively

Some disadvantages of XML are:

- XML can provide a wrapper for complex filetypes or binary objects such as spreadsheets, graphics, video, audio, and even specially or proprietary formatted text. Preservation strategies will still need to address how to preserve, retrieve and present these wrapped objects. In other words, not everything can be converted to XML.
- XML namespaces and tags become an issue when there are multiple "vocabularies" for naming or labeling something. If there is not a common vocabulary, then a mapping of vocabularies is required. This becomes especially difficult when there are two or more ways used to label the same concept. *Solving this challenge was one of the original intentions of the Global Justice XML Data Model (GJXDM) and the National Information Exchange Model (NIEM). See www.niem.gov*
- XML does not preclude the need to develop a preservation strategy – which starts with how records are originally created and preserved. This includes the anticipated demands of those who will use the archives. The enterprise will still need to practice good administrative controls as part of managing records.
- Presentation of a previously archived XML document may not *perfectly* match the original rendering in terms of appearance. Presentation will depend on the XML interpreter used.

Much of XML potential is just starting to be realized. There is a developer learning curve with XML metalanguage - as with

Digital Preservation Enabled Using XML[26]



any new technology or capability - but the potential for reuse, data sharing and work flow advances make it worthwhile.

At this juncture, the state CIO should maintain a healthy skepticism regarding XML for addressing the digital preservation challenge. XML has the potential to be a long term preservation solution, however there will always be something to replace current technology – including XML. Business professionals as well technology professionals are well aware of the many past situations where organizations have given undo reliance to specific technology solutions. XML does not provide a solution for imbedded complex objects such as graphics, and interactive information such as interactive geospatial objects. These complex objects will require preservation of the software that allows them to function or a strategy for simplifying migration to future software packages. Again, there is no "silver bullet" for digital preservation. Notwithstanding this sage advice, XML does *appear* to be an excellent available means for managing some types of electronic records over a long term time horizon with an emphasis on text.

IV. The Necessary Link to Enterprise Architecture

The most assured way for records management and preservation concerns to gain the proper attention is to link these enterprise requirements and enabling capabilities with the state's enterprise architecture program. Enterprise architecture presents and maintains a comprehensive view of the enterprise. This view must recognize the knowledge assets of the enterprise and manage them as *enterprise assets*.[27]

Integrate records management principles into EA program, Project Management, and Change Management

Enterprise architecture was highlighted in the baseline report published by the Center for Technology in Government.[28] One of the observations from that report was that enterprise architecture and records management are not well connected. That report went further to recommend the leveraging of state enterprise architecture efforts to establish the centrality of digital preservation to enterprise-wide information management. Further, results of a series of workshops with state librarians, archivists, and CIO representatives sponsored by the Library of Congress presented that state enterprise architecture initiatives remain an untapped resource for library, records management, and state preservation functions to partner with state information technology and the state CIO's office in support of digital preservation.

Overall respondents appear to be aware of their state's EA efforts (66.7%) while only 37.1% reported any involvement in those efforts.[28]

So why should they be connected?

First, literally everything in the enterprise *is* connected to – in fact, a component of – *THE ENTERPRISE ARCHITECTURE*. This is true whether that enterprise architecture is documented or not. As pointed out in the Center for Technology in Government Baseline Report:

State Library, Archives, Records Management, and CIO representatives identified EA initiatives as having the potential to positively influence digital preservation initiatives by integrating the issues of digital preservation into the ongoing creation of enterprise governance bodies, reference models, business processes, and accountability strategies under development and in use in many states.

This recommendation specifically referenced digital preservation. EA should embrace the full scope of records management – that is the full information lifecycle. This would fit with the vision of EA as an enterprise approach to managing operations, business processes, and enterprise assets. From the records management perspective, the only way to truly accomplish the vision of managing records as corporate assets is to imbed the discipline of records management in every business process.

Records Management Must Be "baked into" the Business Process

The Federal Enterprise Architecture Program Management Office recognized and supported this relationship and worked with the NARA to create a document that reflects the necessary touch points between enterprise architecture and records management within the context of the federal reference models.[29]

As presented earlier in this series, the roles and responsibilities for setting technology standards may reside outside of state records management and archival organizations that have responsibility for providing that service. Yet, this is where statutory authority for records management resides. Enterprise architecture entails the planning and design of functional and organizational structure. Enterprise architects will objectively review who should set standards and who should provide electronic records management and digital preservation services.

Further, even as it is recommended that enterprise architecture partner with records management functions – records management functions should bring enterprise architecture to the table when establishing and implementing standards. The most effective path for implementing records management standards is via the enterprise architecture standards promulgation process. This ensures standards of any kind are maintained, communicated,

periodically reviewed and implemented with an enterprise perspective. Additionally, standards setting will include necessary collaboration with the various allied disciplines of enterprise architecture, project / program management, the capital investment process, procurement, and security.

The Center for Technology in Government research report makes five recommendations for states to help form collaborative arrangements to pursue the combined goals regarding digital preservation and records management:

1. Identify and build on existing knowledge and expertise.

2. Build digital preservation partnerships within and among states.

3. Clarify roles and responsibilities between and among state archives, records management, IT, and other interested and responsible parties.

4. ***Use state Enterprise Architecture efforts to establish the centrality of digital preservation to enterprisewide information management.***

5. Continue to invest in knowledge sharing initiatives across the digital preservation community.

V. Managing Digital Assets

Digital preservation must not be treated as an event, or a reaction to avoid impending disaster.[30] Management and preservation of digital assets must be seen as a *routine operation* that is part of the ongoing management of information and enterprise knowledge assets. Records management must be integrated with the operations of the enterprise. For example, records retention rules must be adopted by all state government agencies as an ongoing *cost of doing business*. In many states, adherence to records retention rules is required by law. As presented by Brian Lavoie and Lorcan Dempsey, digital preservation initiatives must be pre-emptive. The systems on which knowledge assets reside is temporal. The enterprise must anticipate this temporal nature of software programs, operating systems, hardware, and build into their knowledge management strategy a combination of archiving and disposal that recognizes this temporal and fragile nature of electronic information systems.

> *. . . But digital materials generally do not afford the luxury of procrastination. The fragility of digital storage media, combined with a high degree of technology dependence, considerably shortens the "grace period" during which preservation decisions can be deferred. Issues of long-term persistence can arise as soon as the time digital materials are created: for example, in choosing between a widely-used, stable digital format, and one that is obscure or on the verge of obsolescence. This sense of urgency is driven largely by the fact that it is problematic to apply digital preservation techniques ex post—i.e., after deterioration has set in. While a print book with a broken spine can be easily re-bound, a digital object that has become corrupted or obsolete is often impossible (or prohibitively expensive) to restore. Digital preservation techniques are most effective when they are pre-emptive.*[30]

Electronic documents assume multiple iterations over time. They can be authored, transmitted, edited, resent, edited yet again, and distributed innumerably, only to be reedited and resent. It must be determined by the organization what constitutes the *authoritative source* for *specific record types*, what will constitute *convenience copies*, what will be maintained, shared, kept, period of usefulness, regulatory requirements, destruction methods, etc. Knowledge, information and data must be treated holistically using a "cradle to grave" lifecycle management.

Managing digital assets requires consideration and application of the following:

1. Records constitute government knowledge assets. As such they must be managed over a lifecycle with the valuation of that record changing over time.

2. Records management cannot be a reactive activity – rather it must be proactive as part of every capital IT investment decision.

3. Records management discipline will be difficult to invoke using only administrative controls. Business rules that apply to a record must be maintained so that records can be properly managed and disposed of using automated means.

4. When evaluating software options, enterprise architects and procurement must consider what capabilities are being offered by the vendor for managing records going forward.

5. When *developing* or *procuring* a system, a critical requirement is to establish requirements for how the related records / knowledge assets will be managed during the systems lifecycle and beyond. The project team must deliberately anticipate and plan for possible extraction/migration to a new system at some point in the future if and when the system is decommissioned.

   This also entails appropriate risk management considerations. The two sides of the balance equation are: the risk associated with long term preservation – not only the value of the information to the enterprise, but also the liability associated with eDiscovery liabilities - *versus* – the potential loss of knowledge and possibly less risk of eDiscovery liabilities. It may often be cheaper to lose records (even records of some value) than to export them.

6. Much of records management policy has associated regulatory compliance. Know what regulations are applicable within each government line of business.

7. Understand the relationship between technical solutions, record content, and records retention rules. The content and records retention rules may far outlast specific media on which the content currently resides.

VI. Summary – The Top Ten List for CIOs

This series of NASCIO research briefs has covered a lot of ground in electronic records management and digital preservation. The state CIO can't possibly be intimately aware of all of the nuances, interpretations, classifications, and decision making rationale in this very broad subject area. The key recommendation for the state CIO is to collaborate and team with the state experts: records managers archivists and state librarians. Enlist their help in establishing policy, reviewing options, assessing technology solutions along with procurement, project management and enterprise architecture.

In summary the following guidance is provided:

1. Most, if not all, states and the federal government have laws that provide a definition of what constitutes a "record." These laws may establish regulations for records retention and disposition. Records should not be destroyed arbitrarily or capriciously by a government employee. Some process must exist to authorize the destruction of records to ensure that it is done systematically to protect the state from charges of spoliation.[31]

2. All records have value. That value may be short term (until received and read) or long term (forever). Records managers and archivists evaluate administrative, fiscal, legal and historical value of records to determine how long records need to be kept and when destruction should be authorized.

Records managers and archivists must maintain an enterprise view, and compare one record to all other records created by an enterprise when appraising relative value.[31]

3. Record keeping has become distributed. This has a huge impact on the volume, variety and maintenance of electronic records. IT Operations and Data Management need to collaborate with records managers, and archivists regarding how to address these challenges.[31]

4. Desktop recordkeeping issues are different than database/system issues. There is no one-size-fits-all solution.[31]

5. If an electronic record needs to be retained for more than 10 years, the original technology that was used to create/store it will most likely become obsolete during its lifecycle. IT Operations will be responsible for migrating/maintaining the electronic records during this time. This will require an investment in resources and skills. IT Operations needs to work with records managers and archivists to understand, assess and address this situation. State archivists will eventually take custody of some of these electronic records for permanent preservation. They need help from IT Operations and new funding for this task. Otherwise, essential records that protect the rights and interests of the state and individuals, as well as the history and culture of the people will disappear.[31]

6. Electronic Records Management must be an integral part of any project or investment proposal. Project managers must work with records managers and state archivists to determine appropriate plans and associated investments required to maintain the records generated and/or referenced by any business process, or system. These plans must be part of presenting the Total Cost of Ownership associated with any project plans to deliver business processes, or systems.

7. Electronic Records Management discipline must be viewed as an integral part of the state enterprise architecture. Records Management (and Knowledge Management) touch every aspect of Enterprise Architecture – Business, Information, Process, Organization, Technical, Program & Project Management, Security. NASCIO has made a similar case regarding the integral nature of enterprise architecture with project/program management, security, and procurement.

8. Records Management has inherent risk management issues. What to keep and what to destroy will always constitute a balancing act. No one can fully anticipate what knowledge and information will be sought now or at some point in the future. State government cannot keep everything. Records Management policy will have to be established to drive the decision making process for managing records.

9. State CIOs and Enterprise Architects must partner with Records Managers and State Archivists, and State Librarians in order to establish necessary elements for managing digital assets. These elements include policy, responsibility, capacity, and, understanding and awareness among state employees. The desired outcome is to ensure the state has the information it needs today – and tomorrow.

10. State CIOS are focusing more on business strategy as their roles, demands and expectations are expanding. Governors rely on them to interpret, organize and effectively harness information technology to serve the state. In the area of electronic records and digital preservation, NASCIO stresses the need for cooperative, collaborative relationships – particularly with state offices that have statutory responsibility for these lines of business. State

CIOs and Enterprise Architects must partner with the appropriate officials in order to ensure Records Management Policy, Principles and Best Practices are implemented effectively.

Appendix A: Contributors and Reviewers

Mary Alice Baish
Associate Washington Affairs Representative
American Association of Law Libraries
E.B. Williams Law Library
Georgetown University Law Center
111 G Street, NW
Washington, DC 20001-1417
P: 202-662-9200
E: baish@law.georgetown.edu

Jim Bryant
Chief Information Officer
State of South Carolina
Budget and Control Board
4430 Broad River Road
Columbia, SC 29210
P: 803-896-0222
E: jabryant@cio.sc.gov

G. Brian Burke
Senior Program Associate
Center for Technology in Government
University at Albany – State University of New York
187 Wolf Road, Suite 301
Albany, New York 12205
P: 518-442-3892
E: bburke@ctg.albany.edu
www.ctg.albany.edu

L. Reynolds Cahoon
Director-Advanced Programs
Civil Missions Solutions
Lockheed Martin Transportation & Security Solutions
9231 Corporate Boulevard
Rockville, MD 20850
(301) 640-2844
E-mail: reynolds.cahoon@lmco.com

David Carmicheal,
Past President
COSA
The Georgia Archives
5800 Jonesboro Road
Morrow, GA 30260-1101
P: 678-364-3714
F: 678-364-3859
E: dcarmicheal@sos.state.ga.us

Robert Chadduck
Director of Research
Electronic Records Program
National Archives and Records Administration
P: 301-837-1585
E: Robert.Chadduck@nara.gov

Mark Conrad
Archives Specialist – Research Division
Electronic Records Archives (ERA)
NHER
The National Archives and Records Administration
Building 494 Second Floor
310 State Route 956
Rocket Center, WV 26726
Phone: 304-726-7820
Email: mark.conrad@nara.gov

Dr. Bruce Dearstyne
Lecturer
University of Maryland, College of Information Studies
Principal
University of Maryland Center for Information Policy
23 Venezio Avenue
Albany, NY 12203
Phone: (518) 456-0872
E-mail: dearstyne@verizon.net

Stephen Fletcher – Co-chair
Chief Information Officer
State of Utah
Office of the Governor
East Capitol Complex Bldg
Ste E-220
Salt Lake City, UT 84114
P: (801) 538-3298
E: sfletcher@utah.gov

Debbie Gearhart, CRM
Director
Records Management Services
Michigan Historical Center
Department of History, Arts, and Library
3405 N. Martin Luther King Jr. Blvd.
PO Box 30026
State of Michigan,
Lansing, MI 48909
P: 517-335-9145
E: Gearhartd@michigan.gov
http://www.michigan.gov/recordsmanagement

Margaret Hedstrom, Ph.D.
Associate Professor
University of Michigan
School of Information
3210 SI North
Ann Arbor, MI 48109
P: (734) 647-3582
E: hedstrom@umich.edu

William (Bill) LeFurgy
Project Manager Digital Initiatives
Office of Strategic Initiatives
National Digital Information Infrastructure and Preservation Program
(NDIIPP)
The Library of Congress
P: 202-707-8618
E: wlef@loc.gov

Scott A. Leonard
Electronic Records Specialist
Kansas State Historical Society
6425 SW Sixth Avenue, Topeka, KS 66615
P: 785-272-8681, ext. 280
E: SLeonard@kshs.org
www.kshs.org
ERM at KSHS: www.kshs.org/government/records/electronic/index.htm

Reagan Moore
Associate Director
Data-Intensive Computing
San Diego Supercomputer Center (SDSC)
UC San Diego, MC 0505
9500 Gilman Drive
La Jolla CA 92093-0505
P: 858-534-5073
F: 858-534-5152
E: moore@sdsc.edu

Richard Pierce-Moses
Deputy Director for Technology and Information Resources
State of Arizona
Arizona State Library, Archives, and Public Records
1700 West Washington
Suite 200
Phoenix, AZ 85007
Phone: 602-926-4035
Email: rpm@lib.az.us

Stephen Newell
IT Architect
IBM
2401 W Jefferson St
Springfield, IL 62702-3495
P: (217) 793-5825
E: srnewel@us.ibm.com

Doug Robinson
Executive Director
NASCIO
201 East Main St., Suite 1405
Lexington, KY 40507
P: (859) 514-9153
E: drobinson@AMRms.com

Bill Roth
Chief Architect
State of Kansas
Landon State Office Building
900 Southwest Jackson
Room 751 South
Topeka, KS 66612
P: (785) 296-2108
E: Billr@ksdot.org

Michael P. Ryan
Enterprise Architect (retired)
State of Minnesota
mryan dot net LLC
St. Paul, MN 55105
michael.ryan@mryan.net
Phone 651.699.7159

Russ Saito
Comptroller
State of Hawaii
Department of Accounting & General Services
1151 Punchbowl St Rm 412
Honolulu, HI 96813-3007
P: (808) 586-0400
E: russ.k.saito@hawaii.gov

Dr. Timothy Sprehe
Sprehe Information Management Associates, Inc.
5504 Surrey Street
Chevy Chase, MD 20815-5524
P: 301-657-2481
E: jtsprehe@jtsprehe.com
http://www.jtsprehe.com

Chris Walls
Senior Website & Publications Coordinator
AMR Management Services
201 E. Main St., Suite 1405
Lexington, KY 40507
P: 859-514-9174
E: cwalls@AMRms.com

Caryn Wojcik
Government Records Archivist
Records Management Services Michigan
Historical Center Department of History,
Arts and Libraries
3405 N. Martin Luther King Jr. Blvd.
PO Box 30026
Lansing, Michigan 48909
P: (517) 335-8222
F: (517) 335-9418
E: wojcikc@michigan.gov

Appendix B: Resources

NASCIO
www.nascio.org

Seek and Ye Shall Find? State CIOs Must Prepare Now for E-Discovery
September 2007
This issue brief explains the impact for State CIOs of e-discovery requests and encourages State CIOs to pursue a holistic approach to enterprise records management as part of a team of state government stakeholders, including state legal counsel, archivists, records managers, and agency business leaders.
http://www.nascio.org/publications/index.cfm#88

Department of Justice, Office of Justice Programs
Global Privacy and Information Quality Working Group
http://www.it.ojp.gov/topic.jsp?topic_id=55

Center for Technology in Government
Preserving State Government Digital Information: A Baseline Report
http://www.ctg.albany.edu/publications/reports/digital_preservation_baseline

Records Management / Digital Preservation Related Web sites
Library of Congress / Digital Preservation
http://www.digitalpreservation.gov

The National Archives
http://www.archives.gov/

Knowledge Management World
http://www.kmworld.com/

Council of State Archivists
http://www.statearchivists.org/

National Association of Government Archives and Records Administrators
http://www.nagara.org/

Association of Records Managers and Administrators
http://www.arma.org/

The Society of American Archivists
http://www.archivists.org/

Digital Preservation Coalition
http://www.dpconline.org/graphics/whatsnew/

Tutorial on Digital Preservation Management, Cornell University
http://www.library.cornell.edu/iris/tutorial/dpm/eng_index.html

Online Course – email Retention Guidelines – State of Michigan
http://www.michigan.gov/documents/hal/mhc_rm_email_class_175020_7.pdf

Articles by Dr. Timothy Sprehe
Sprehe Information Management Associates, Inc.
http://www.jtsprehe.com/newpage6.htm

NECCC Best Practices for eDiscovery
www.ec3.org/Downloads/2004/Effectively_Man_Discovery_of_El_Records.pdf

InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems
www.interpares.org
The InterPARES Project was launched in 1999 to address a widespread concern. In the last quarter of the century, organizations and individuals have come to rely in a fundamental manner on the creation, exchange and processing of digital information.

Department of Defense
Electronic Records Management Software Applications Design Criteria Standard
www.dtic.mil/whs/directives/corres/pdf/501502std.pdf

eDiscovery 2004 eC3
http://www.ec3.org/Downloads/2004/Effectively_Man_Discovery_of_El_Records.pdf

The San Diego Supercomputer Center
The San Diego Supercomputer Center (SDSC) enables international science and engineering discoveries through advances in computational science and high performance computing. Continuing this legacy into the era of cyberinfrastructure, SDSC is a strategic resource to science, industry and academia, offering leadership in the areas of data management, grid computing, bioinformatics, geoinformatics, high-end computing as well as other science and engineering disciplines.

The mission of SDSC is to extend the reach of scientific accomplishments by providing tools such as high-performance hardware technologies, integrative software technologies and deep inter-disciplinary expertise, to the community.
http://www.sdsc.edu/

The Fedora Project
Fedora is a general purpose repository system developed jointly by Cornell University Information Science and the University of Virginia Library . The Fedora Project is devoted to the goal of providing open-source repository software and related services to serve as the foundation for many types of information management systems. The Fedora software is available under the terms of the Educational Community License 1.0 (ECL).

The Fedora Project is based on previous research at Cornell University Computer Science that was funded by DARPA and the National Science Foundation.

The Fedora Project is currently supported by generous grants from the Andrew W. Mellon Foundation.
http://www.fedora.info/about/

Reports
Federal Enterprise Architecture (FEA) Records Management Profile, version 1.0
http://www.archives.gov/records-mgmt/policy/rm-profile.html

American Association of Law Librarians
53 West Jackson Boulevard, Suite 940,
Chicago, Ill. 60604
http://www.aallnet.org/

- State-By-State Report On Permanent Public Access To Electronic Government Information
- State-By-State Report On Authentication of Online Legal Resources

The Consultative Committee for Space Data Systems (CCSDS)
1801 Alexander Bell Drive, Suite 500
Reston, VA 20191-4344
http://public.ccsds.org/about/default.aspx

- Reference Model for an Open Archival Information System (OAIS)

Online Computer Library Center (OCLC)

- Trusted digital repositories : attributes and responsibilities : an RLG-OCLC report. by RLG/OCLC Working Group on Digital Archive Attributes.; Research Libraries Group.; OCLC.
  Type: Book
  Language: English
  Publisher: Mountain View, CA : RLG, 2002.
  Editions: 2 Editions
  OCLC: 49802387

The Society of American Archivists
527 South Wells St., 5th Floor
Chicago, IL 60607
312/922-0140
http://www.archivists.org/

- *A Glossary of Archival and Records Terminology*
  http://www.archivists.org/glossary/index.asp

Appendix C: Endnotes

[1] D. Rosenthal, T. Robertson, T. Lipks, V. Reich, S. Morabito, Requirements for Digital Preservation Systems, D-Lib Magazine, November 2005, Vol 11., No. 11, ISSN 1082-9873, retrieved on October 12, 2007 from http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html

[2] R. Lombardi, "Lost in Space", CIO Government Review, June 14, 2007, InterGovWorld.com. Retrieved on September 25, 2007, from http://www.intergovworld.com/article/2bd1d5580a010408005313802405e293/pg1.htm

[3] See Lavoie, B., & Dempsey, L. (July/August 2004), "Thirteen Ways of Looking at...Digital Preservation", D-Lib Magazine, Volume 10, Number 7/8, ISSN 1082-9873, Retrieved September 1, 2006 from http://www.dlib.org/dlib/july04/lavoie/07lavoie.html

[4] Walker, Richard W., "For the record, NARA techie aims to preserve", Government Computing News, July 30, 2001. Retrieved on March 21, 2007, from http://www.gcn.com/print/vol20_no21/4752-1.html?topic=news

[5] See http://www.rlg.org/en/page.php?Page_ID=2

[6] See Digital Preservation Coalition, "Technology Issues", Retrieved September 7, 2006, from http://www.dpconline.org/graphics/digpres/presissues.html

[7] See Sprehe, J.T., "Enterprise Records Management: Strategies and Solutions", September 2002, Retrieved on September 21, 2006, from http://www.jtsprehe.com/newpage6.htm, secondary source: http://www.hummingbird.com/alt_content/binary/pdf/collateral/wp/rmstrategies.pdf

[8] Contributory writing from Caryn Wojcik, Government Records Archivist, state of Michigan.

[9] Interview of Julie Gable, Principal, Gable Associates LLC, see contact information under *Contributors.* Also see article "What CIOs Should Know About Records", by Julie Gable, March 23, 2005, Searchcio.com, retrieved on October 30, 2006, from http://searchcio.techtarget.com/originalContent/0,289142,sid19_gci1070614,00.html

[10] Federal News Radio, "Trends in Technology", Records Management – Interview with Christopher Dorobek and Timothy Sprehe. August 28, 2006, Retrieved on September 21, 2006, from http://www.federalnewsradio.com/index.php?nid=277&sid=895215

[11] Walker, Richard W., "For the record, NARA techie aims to preserve", Government Computing News, July 30, 2001. Retrieved on March 21, 2007, from http://www.gcn.com/print/vol20_no21/4752-1.html?topic=news

[12] See Digital Preservation Coalition, "Technology Issues", Retrieved September 7, 2006, from http://www.dpconline.org/graphics/digpres/presissues.html

[13] Definition from the Curtin University of Technology, Records and Archive Office Recordkeeping Manual – Glossary of Recordkeeping Terms. Retrieved on October 12, 2007 from http://www.records.curtin.edu.au/recordkeeping_manual/glossary.html

[14] Society of American Archivists, "A Glossary of Archival and Records Terminology." Retrieved July 7, 2007 from http://www.archivists.org/glossary/term_details.asp?DefinitionKey=1784

[15] "Reference Model for an Open Archival Information System (OAIS)", CCSDS, CCSDS 650.0-B-1, January 2002, available at http://public.ccsds.org/default.aspx

[16] See Chadduck, R., Moore, R., Joseph, J., "Digital Archiving and Long Term Preservation Early Experience with Grid and Digital Library Technologies", Retrieved September 7, 2006 from http://www.archives.gov/era/papers/thic-04.html

[17] The Transcontinental Persistant Archives Prototype (TPAP) extends to 3 NARA locations (as of 5/25/2007). See http://www.sdsc.edu/News%20Items/PR051707_tpap.html
In 2006 the TPAP won an inaugural Internet2 IDEA Award.
See http://www.internet2.edu/idea/2006/transcontinental_persistent_archives_prototypes.html

[18] See San Diego Supercomputer Center, http://www.sdsc.edu/

[19] See Presentation by Reagon Moore, "Managing Simulation Output", retrieved on September 25, 2006, from http://www.sdsc.edu/user_services/training/workshops/2006cihass/docs/2006cihass_Reagan-Moore-Summer.ppt#776,16,Generic Infrastructure_

[20] See DSpace website at http://dspace.org/index.html

[21] See http://www.fedora.info/

[22] Interview with Scott Leonard, Kansas Historical Society

[23] See legislative update from NAGARA at http://www.nagara.org/displaycommon.cfm?an=1&subarticlenbr=16

[24] Federal News Radio, "Trends in Technology", Records Management – Interview with Christopher Dorobek and Timothy Sprehe. August 28, 2006, Retrieved on September 21, 2006, from http://www.federalnewsradio.com/index.php?nid=277&sid=895215

[25] See Slats, J., Verdgem, R., "Practical experiences of the Dutch Digital Preservation Testbed", February 2004. Vol. 34, Iss. 2; pg. 56-65, The journal of information and knowledge management systems, retrieved September 25, 2006 from http://www.digitaleduurzaamheid.nl/bibliotheek/docs/Article_in_VINE_2004.pdf#search=%22netherlands%20digital%20preservation%22

[26] Adapted from *Figure 4. Conversion to XML requires fewer conversions than migration* from "Practical experiences of the Dutch Digital Preservation Testbed" by J. Slats & R. Verdegem, VINE, Vol 34, No. 2, 2005 issue 135, page 56-65. Retrieved on 10/15/2007 from http://www.digitaleduurzaamheid.nl/bibliotheek/docs/Article_in_VINE_2004.pdf#search=%22netherlands%20digital%20preservation%22

[27] See Pardo, T., Burke, B., April 7, 2006, *Partnering for Preservation*, Public CIO Magazine. See response from Doug Robinson, NASCIO Executive Director. Retrieved on May 25, 2007 from http://www.govtech.com/pcio/99093?id=99093&story_pg=1

[28] See Pardo, T., Burke, B., July 2006, "Preserving State Governmental Digital Information: A Baseline Report", Page 5 and Section Six, retrieved September 19, 2006, from http://www.ctg.albany.edu/publications/reports/digital_preservation_baseline

[29] See http://www.archives.gov/records-mgmt/policy/rm-profile.html, retrieved October 30, 2006.

[30] See Lavoie, B., & Dempsey, L. (July/August 2004), "Thirteen Ways of Looking at...Digital Preservation", D-Lib Magazine, Volume 10, Number 7/8, ISSN 1082-9873, Retrieved September 1, 2006 from http://www.dlib.org/dlib/july04/lavoie/07lavoie.html

[31] Contributory writing from Caryn Wojcik, Government Records Archivist, state of Michigan

This report and the NASCIO Enterprise Architecture Program are funded by a grant from the Bureau of Justice Assistance, Office of Justice Programs, U.S. Department of Justice.

The opinions, findings, conclusions, and recommendations contained in this publication are those of the contributors, and do not necessarily reflect the official positions or policies of the Department of Justice.

Disclaimer

NASCIO makes no endorsement, express or implied, of any products, services, or websites contained herein, nor is NASCIO responsible for the content or the activities of any linked websites. Any questions should be directed to the administrators of the specific sites to which this publication provides links. All critical information should be independently verified.