.wp .wks .jpg .pdf .gif .doc .xls .dif .pps .pub
.rtf .bmp .xy3 .png .mm .ppt .wpd .ooxml .mpg
.cgm .txt .html .tif .eps .wmf .xml .csv .dbf .wk1
.prn .wq1 .pic .slk .psd .wp5 .wks .jpg .odf .pdf
.doc .psp .gif .dxf .xls .pps .pub .rtf .bmp .png
.mm .ppt .wpd .mpg .txt .html .tif .eps .wmf
.xml .csv .dbf .wk1 .prn .wq1 .slk .psd .ai .hpgl
.pal .dat .wp .wks .jpg .pdf .gif .doc .xls .dif .pps
.pub .rtf .bmp .xy3 .png .mm .ppt .wpd .ooxml
.mpg .cgm .txt .html .tif .eps .wmf .xml .csv .dbf
.wk1 .prn .wq1 .pic .slk .psd .wp5 .wks .jpg .odf
.pdf .doc .psp .gif .dxf .xls .pps .pub .rtf .bmp
.png .mm .ppt .wpd .mpg .txt .html .tif .eps
.wmf .xml .csv .dbf .wk1 .prn .wq1 .slk .psd .ai
.hpgl .pal .dat .wp .wks .jpg .pdf .gif .doc .xls
.dif .pps .pub .rtf .bmp .xy3 .png .mm .ppt .wpd
.ooxml .mpg .cgm .txt .html .tif .eps .wmf .xml
.csv .dbf .wk1 .prn .wq1 .pic .slk .psd .wp5 .wks
.jpg .odf .pdf .doc .psp .gif .dxf .xls .pps .pub
.rtf .bmp .png .mm .ppt .wpd .mpg .txt .html
.tif .eps .wmf .xml .csv .dbf .wk1 .prn .wq1 .slk
.psd .ai .hpgl .pal .dat .wp .wks .jpg .pdf .gif
.doc .xls .dif .pps .pub .rtf .bmp .xy3 .png .mm
.ppt .wpd .ooxml .mpg .cgm .txt .html .tif .eps
.wmf .xml .csv .dbf .wk1 .prn .wq1 .pic .slk .psd
.wp5 .wks .jpg .odf .pdf .doc .psp .gif .dxf .xls
.pps .pub .rtf .bmp .png .mm .ppt .wpd .mpg
.txt .html .tif .eps .wmf .xml .csv .dbf .wk1 .prn
.wq1 .slk .psd .ai .hpgl .pal .dat .wp .wks .jpg
.pdf .gif .doc .xls .dif .pps .pub .rtf .bmp .xy3
.png .mm .ppt .wpd .ooxml .mpg .cgm .txt .html
.tif .eps .wmf .xml .csv .dbf .wk1 .prn .wq1 .pic
.slk .psd .wp5 .wks .jpg .odf .pdf .doc .psp .gif
.dxf .xls .pps .pub .rtf .bmp .png .mm .ppt .wpd
.mpg .txt .html .tif .eps .wmf .xml .csv .dbf .wk1
.prn .wq1 .slk .psd .ai .hpgl .pal .dat .wp .wks
.jpg .pdf .gif .doc .xls .dif .pps .pub .rtf .bmp
.xy3 .png .mm .ppt .wpd .ooxml .mpg .cgm .txt
.html .tif .eps .wmf .xml .csv .dbf .wk1 .prn
.wq1 .pic .slk .psd .wp5 .wks .jpg .odf .pdf .doc
.psp .gif .dxf .xls .pps .pub .rtf .bmp .png .mm
.ppt .wpd .mpg .txt .html .tif .eps .wmf .xml .csv
.dbf .wk1 .prn .wq1 .slk .psd .ai .hpgl .pal .dat
.wp .wks .jpg .pdf .gif .doc .xls .dif .pps .pub
.rtf .bmp .xy3 .png .mm .ppt .wpd .ooxml .mpg
.cgm .txt .html .tif .eps .wmf .xml .csv .dbf .wk1
.prn .wq1 .pic .slk .psd .wp5 .wks .jpg .odf .pdf
.doc .psp .gif .dxf .xls .pps .pub .rtf .bmp .png
.mm .ppt .wpd .mpg .txt .html .tif .eps .wmf
.xml .csv .dbf .wk1 .prn .wq1 .slk .psd .ai .hpgl
.pal .dat .wp .wks .jpg .pdf .gif .doc .xls .dif .pps
.pub .rtf .bmp .xy3 .png .mm .ppt .wpd .ooxml
.mpg .cgm .txt .html .tif .eps .wmf .xml .csv .dbf
.wk1 .prn .wq1 .pic .slk .psd .wp5 .wks .jpg .odf
.pdf .doc .psp .gif .dxf .xls .pps .pub .rtf .bmp
.png .mm .ppt .wpd .mpg .txt .html .tif .eps
.wmf .xml .csv .dbf .wk1 .prn .wq1 .slk .psd .ai
.hpgl .pal .dat .mm .ppt .wpd .ooxml .mpg .cgm
.txt .html .tif .eps .wmf .xml .csv .dbf .wk1 .prn
.wq1 .pic .slk .psd .wp5 .wks .jpg .odf .pdf .doc
.psp .gif .dxf .xls .pps .pub .rtf .bmp .png .mm
.ppt .wpd .mpg .txt .html .tif .eps .wmf .xml .csv
.dbf .wk1 .prn .wq1 .slk .psd .ai .hpgl .pal .dat
.wp .wks .jpg .pdf .gif .doc .xls .dif .pps .pub
.rtf .bmp .xy3 .png .mm .ppt .wpd .ooxml .mpg
.cgm .txt .html .tif .eps .wmf .xml .csv .dbf .wk1
.prn .wq1 .pic .slk .psd .wp5 .wks .jpg .odf .pdf
.doc .psp .gif .dxf .xls .pps .pub .rtf .bmp .png
.mm .ppt .wpd .mpg .txt .html .tif .eps .wmf
.xml .csv .dbf .wk1 .prn .wq1 .slk .psd .ai .hpgl
.pal .dat .wp .wks .jpg .pdf .gif .doc .xls .dif .pps
.pub .rtf .bmp .xy3 .png .mm .ppt .wpd .ooxml
.mpg .cgm .txt .html .tif .eps .wmf .xml .csv .dbf
.wk1 .prn .wq1 .pic .slk .psd .wp5 .wks .jpg .odf
.pdf .doc .psp .gif .dxf .xls .pps .pub .rtf .bmp
.png .mm .ppt .wpd .mpg .txt .html .tif .eps
.wmf .xml .csv .dbf .wk1 .prn .wq1 .slk .psd .ai

Preserving the Present:

Creating, accessing and maintaining Minnesota's electronic documents

Report to the Minnesota Legislature

January 15, 2008

Gopal Khanna, Chief Information Officer

State of Minnesota

Officer of Enterprise Technology

Office of Enterprise Technology
State of Minnesota
Innovation. Change. Transformation.

This report was prepared in accordance with Minnesota Laws 2007, Chapter 148, Article 2, Section 77, which directs the chief information officer of the state, in consultation with the state archivist and legislative reference librarian, to study how electronic documents and the mechanisms and processes for accessing and reading electronic data can be created, maintained, exchanged, and preserved by the state in a manner that encourages appropriate government control, access, choice, and interoperability. For more information, contact the Office of Enterprise Technology at 651.296.8888.

Preserving the Present:

Creating, accessing and maintaining Minnesota's electronic documents

Executive Summary 1

Introduction 3

Issue Context 4

Study Process 6

Steering team; Input to the study; Study scope 6

Definition of "Records" 6

Open Formats 7

Records Life Cycle 8

Research Results 9

Attention to stakeholder needs 9

Policies of other states and nations; Management guidelines for state archives as they pertain to electronic documents 10

Privacy and confidentiality of electronic records 11

Standards; Use of the World Wide Web; Better organization of information for public access; Costs of implementation 12

Potential savings; Interoperability and data sharing 13

Long-term preservation; Changes to records laws 14

Conclusions 15

Specific recommendations related to adaptation of standard document format; General recommendations 16

Next steps 17

Appendix A Background Information on the Issues 18

Appendix B Stakeholder Survey 19

Appendix C Stakeholder Survey Responses 19

Executive Summary

In its 2007 session, the Minnesota legislature directed the state's Chief Information Officer, Gopal Khanna, to undertake a study related to preservation of electronic documents. The catalyst for the study and this report was a bill introduced earlier in the 2007 session that called for all government documents to be created, exchanged and preserved in a particular file format. In response to that bill, discussions among interested parties and stakeholders greatly enlarged the scope of study. Through the committee process in the House and Senate, the focus moved from the concept of the mandate of a single technological standard, set in statute, to the much broader perspective of the infrastructure, policies, practices and ancillary legislation that could realize the possibilities for digital preservation expressed in the bill.

The records that government creates must be managed in a way that ensures that they remain trustworthy, complete, accessible and durable. Minnesota's strategy for managing its electronic records must take into account the needs of multiple stakeholders. An effective strategy will be flexible enough to account for these and other uses of records as various points throughout their life cycle.

A survey to solicit stakeholder input to this report sought to answer the question: "What is Minnesota government's plan for preserving and managing records?" Some responders addressed a narrower question related only to the adoption of a document format standard. Some responders, however, did address the fundamental issues related to electronic records raised by the later, broader language. Many advocates for citizen access struck a moderate tone, recognizing the practical impacts of format decisions and placing an emphasis on the proper documentation and organization needed for records access in a variety of formats.

Adopting a standard file format is one way to overcome some of the technical barriers to successful information exchanges between users. Organizations may adopt standard file formats to increase the ability of users with disparate technology environments to exchange data, to use a software program or to retrieve data from a repository.

This report does not recommend the adoption of a particular format standard. The dynamic nature of technology innovation and change make adoption of a single standard problematic. Moving in the direction of a fully documented functional document standard that can do all one wants is desirable. But neither of the competing standards proposed addresses all the government goals and purposes in the law. In any case, the choice or use of a standard must not be to adopt a standard for the sake of adopting a standard. Any choice must be in the context of what value such a decision adds to government.

The report identities several concrete, practical steps that the state can take to address electronic records policy issues so that actions taken support the best interests of the state for responsible stewardship of information resources, including working with other states and stakeholders to seek collaborative approaches to common problems in government electronic records management.

We obviously can't save everything for everybody. *But we can save demonstrably valuable information; we can layer it with other information, in a critical mass; we can use technology to make it all available online; and we can provide ever more sophisticated tools to make that information more and more useful. If we can't do everything for all people, we can certainly focus and prioritize our efforts to deliver useful products and services to key constituencies.*

Robert Horton, State Archivist, Minnesota Historical Society. Notes from the 20th Anniversary of the National Archives and Records Administration Panel Discussion

Introduction

In its 2007 session, the Minnesota legislature directed the state's Chief Information Officer, Gopal Khanna, to undertake a study related to preservation of electronic documents. The pertinent legislative language (Minnesota Laws 2007, Chapter 148, Article 2, Section 77) reads:

> *The chief information officer of the state, in consultation with the state archivist and legislative reference librarian, shall study how electronic documents and the mechanisms and processes for accessing and reading electronic data can be created, maintained, exchanged, and preserved by the state in a manner that encourages appropriate government control, access, choice, and interoperability.*

The law states that CIO is required to report his findings and recommendations to the Legislature by January 15, 2008.

In a democratic society, citizens are entitled to a government that manages information effectively to promote the public good and with the care necessary to protect private rights. The records created must be managed in a way that ensures that they remain trustworthy, complete, accessible and durable.

Minnesota's strategy for managing its electronic records must take into account the needs of multiple stakeholders. Some records are valuable for only the short time that they are needed for a given transaction; others must be preserved and accessible in original form for use in litigation; others must be preserved for their historical value. An effective strategy will be flexible enough to account for these and other uses of records as various points throughout their life cycle.

Issue Context

The catalyst for this report was a bill introduced earlier in the 2007 session. The original proposal, titled the Preservation of State Documents Act, read:

> *Effective July 1, 2008, all documents including text, spreadsheets, and presentations of the state of Minnesota shall be created, exchanged, maintained, and preserved in an open, XML-based file format, as specified by the chief information office of the state, that is:*
>
> *(1) interoperable among diverse internal and external platforms and applications;*
> *(2) fully published and available royalty-free;*
> *(3) implemented by multiple vendors; and*
> *(4) controlled by an open industry organization with a well-defined inclusive process for evolution of the standard.*
>
> *By that date, the state of Minnesota shall be able to accept all documents received in open document format for office applications and sha*ll *not migrate to a file format currently used by only one organization.*

In response to that bill, discussions among interested parties and stakeholders greatly enlarged the subject of study. Through the committee process in the House and Senate, the focus shifted from mandating a single technological standard to a broad consideration of the infrastructure, policies, and practices needed for digital preservation. Instead of, "Which standard for office documents is better for preservation?" the question became, "What is Minnesota government's plan for preserving and managing government records?" And, "What are the implications for Minnesotans?"

Records serve as the organization's memory and are evidence of past events and the basis for future actions. When created, maintained and disposed in an orderly and systematic manner, records are a tremendous asset. When treated in a haphazard and disorderly manner, they can reduce an organization's effectiveness, hinder understanding of policy decisions, and increase costs and liabilities. Information technology innovation has resulted in an exponential increase in the production and accessibility of government records and makes proper planning for their management increasingly critical. Not having a plan to solve a problem does not make the problem go away.

A survey to solicit stakeholder input sought to answer the question: "What is Minnesota government's plan for preserving and managing records?" Some responders addressed a narrower question related only to the adoption of a document format standard. Some responders, however, did address the fundamental issues related to electronic records that were raised by the later, broader language.

Responders to the survey differed dramatically in their priorities and concerns, but all agreed on the importance of preserving the state's electronic documents and records.

Given the impossibility of keeping everything, the state should build a system to identify and

preserve documents and records of most value to future generations. Roles among agencies should be defined and build on existing programs. For example, as a depository for state government publications, the Legislative Reference Library archives paper and pdf copies of publicly available documents from all state agencies, with an emphasis on documents mandated by the Legislature. (Read more at: www.leg.state.mn.us/lrl/mndocs/mndocs.asp) The State Archives selects records for their historical value. Many agencies retain deep archives of public and private electronic records, but lack clear guidelines as to what should be kept long term.

Many advocates for citizen access struck a moderate tone, recognizing the practical impacts of format decisions and placing an emphasis on the proper documentation and organization needed for records access in a variety of formats.

The state needs to ensure access to and preservation of electronic documents. The state's information architecture policies recommend the use of open format without standardizing on a single format. The dynamic nature of technology innovation and change make adoption of a single standard problematic. Furthermore, a decision about standards is not simply a matter of choosing between Standard A or Standard B. Other options could include the decision not to adopt a standard at all or to adopt multiple standards. Any decision about standardizing must be made based on the best interests of citizens and on the business requirements of state and local government. Those requirements include operational requirements, access controls, costs of implementation and records retention. Such a decision must also be made within the context of the state's overall plan for managing records throughout their life cycle.

Creation of an effective system *for archiving electronic records needs participation across all branches and levels of government. Strong technical leadership by the state CIO is important, but IT expertise should be augmented with professionals who are expert and experienced content managers — archivists, records managers, librarians. This study is a great start to identifying needs, and hopefully an impetus for obtaining information technology funds at the state level to facilitate collaborative planning and benefit all agencies.*

Robbie LaFleur
Director, Minnesota Legislative
Reference Library

Study Process

Steering team

The study was coordinated by the Office of Enterprise Technology (OET), working through a project steering team comprising the state archivist, legislative reference librarian and OET's director of strategic planning. The plan for the study was discussed with major stakeholders prior to initiation.

Input to the study

To allow all potential stakeholders a fair and equal opportunity to comment on the topic before the data gathering and analysis were completed, OET conducted an electronic survey in September and October 2007, with structured questions tied to the legislative requirements. Alternative forms of input were accepted from stakeholders who did not have ability to access the web form. Nearly 50 representatives of government agencies, libraries, companies, and citizen interests provided responses to the survey. For the survey questions, see Appendix B. For the survey responses, see Appendix C.

Study scope

The study considered, among other issues,

- the policies of other states and nations,
- management guidelines for state archives as they pertain to electronic documents,
- public access to information,
- expected storage life of electronic documents,
- costs of implementation,
- potential savings,
- industry history and trends, and
- state information architecture.

The team reviewed current literature related to the study, looking particularly for material relevant in a government records environment. Several recent reports summarize the current state of knowledge, articulating and identifying practical options. As such, they are excellent reference points for Minnesota as it moves forward and they provided a common framework for the discussion that informed this report. For the reports, see Appendix A.

Definition of "Records"

The terms "records," "data," and "documents" are often used interchangeably; indeed, the language that calls for this study itself uses "documents" and "data" without distinguishing the two. But all of these terms have separate legal meanings and significance in Minnesota and it is important to keep in mind that stakeholders may have one particular legal definition in mind, or may not distinguish among the meanings.

For example, the word "document" —which most people think they understand—could, depending on the circumstances, be considered a document, a record, or data, subject to a variety of Minnesota laws. The resulting confusion can lead to unintentional misuse or noncompliance.

In Minnesota Statutes 3.03, the Legislative Reference Library is designated as repository of government documents. In this case, "documents" means official publications. Minnesota Statutes 15.17 defines "official records" as records that are necessary to a full and accurate knowledge of government's official duties. Definition of these records includes documents. Minnesota Statutes 138.17 defines "government records"

and describes a process to be followed to obtain authority to dispose of them. Some documents are records and some are not. Chapter 13 defines "government data" broadly as all data collected, created, received, maintained or disseminated by government entities regardless of physical form, storage media or conditions of use. Chapter 13 describes the circumstances under which the data may be accessed or protected from access.

The upshot of all of this is that all these (and other) statutory references have to be taken into account when determining the proper use and disposition of a "document."

In the stakeholder survey, in an effort to avoid confusion, the term "electronic records" was used inclusively. The definition was borrowed from the Uniform Electronic Transactions Act (UETA) (Minnesota Statutes, Chapter 325L):

> *Electronic records "mean" records created, generated, sent, communicated, received, or stored by electronic means.*
>
> *Records "mean" information that is inscribed on a tangible medium or that is stored in an electronic or other medium and is retrievable in perceivable form."*

Responders to the survey were asked to indicate if the use of any term other than "electronic records" carries any special meaning or has any special implications. Most responders, however, used "records," "data," and "documents" interchangeably.

Open Formats

Adopting a standard file format is a way to overcome some of the technical barriers to successful information exchanges between users. Organizations may adopt standard file formats to increase the ability of users with disparate technology environments to exchange data, to use a software program or to retrieve data from a repository.

An open format is a published specification for storing digital data, usually maintained by a non-proprietary standards organization, and free of legal restrictions on use. For example, an open format must be implementable by both proprietary and free and open source software, using the typical licenses used by each. In contrast to open formats, proprietary formats are controlled and defined by private interests. Open formats are a subset of open standards. (The relationship between open formats and free and open source software is frequently misunderstood. Many proprietary software products readily use open formats, and free and open source software can often use proprietary formats.)

The Open Document Format (ODF) is an XML-based open source file format for saving and exchanging text, spreadsheets, charts, and presentations. ODF was developed by a committee formed under the OASIS (Organization for the Advancement of Structured Information Standards) consortium. The standard is published as ISO/IEC 26300:2006.

ISO is a worldwide federation of national standards bodies from some 100 countries, with one standards body representing each member country. The American National Standards Institute (ANSI) represents the United States. Member organizations collaborate in the development and promotion of international standards.

OOXML (Office Open Extended Markup Language), also called Open XML, is a file format

for documents, spreadsheets and presentations that is intended for use with the 2007 and later versions of the Microsoft Office suite. The specification is currently undergoing fast-track standardization within ISO/IEC as DIS 29500 (Draft International Standard 29500), but the draft text has failed (as of the September, 2007 ballot) to reach sufficient approval from ISO/IEC national body members to be accepted as a standard. A ballot resolution process will allow for the text to be amended and a final decision to be reached on its acceptability for standardization.

There are significant functional and philosophical differences between ODF and OOXML. A "standards war" has arisen concerning the relative openness and flexibility of OOXML compared with ODF. It is relevant to note that neither provides a complete solution.

(This section on open format standards was based on standard reference sources.)

***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.*

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

Records Life Cycle

The life cycle of a record starts when it is created for a business purpose, maintained for a period of time that is determined by its administrative, legal, historical or fiscal value; when it is disposed of because it no longer has value, the cycle ends. During its life, the record may need to be exchanged as part of its use. Some records may be available for public access or they made need to be protected from access. A small percentage of all government records are preserved long-term because of their historical value. Records management as a professional discipline is related to managing the record throughout its life cycle, identifying the record, determining value, assigning a retention period based on the value, and disposing of the record according to an approved retention schedule.

***Digital preservation must not be treated as an event,** or a reaction to avoid impending disaster. 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.*

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

Research results

Attention to stakeholder needs

In addressing public access, responders urged the state to pay attention to multiple stakeholder needs and to balance the needs of many different audiences with different technology capabilities and different abilities to maintain the latest version of software. The state was urged to embrace a range of goals, in government control, access, privacy, and security. Local entities mentioned the need for the state to provide centralized services for storage, conversion, scanning and other resources beyond the reach but not the need of local entities.

User needs were cited, such as public internet access at public libraries throughout the state, and a reminder that some users will have limited English language proficiency. The state was urged to pay attention to the needs of blind and visually impaired users, and to ensure that PDF or other formats include searchable text rather than image files which cannot be read by translators.

Responders expressed concern about deterioration and inaccessibility of records over time. They said it is critical to take into account all aspects that contribute to a trustworthy system.

Citizens, small business, real estate, regional economic development organizations, nonprofits, and other jurisdictions were identified as key stakeholders that may be overlooked in developing a state-level approach. The impact on local government was identified as a concern when establishing state level systems. Others stressed the unique needs for handling health records.

There are multiple public interests in the management of government information. To date, strategic information systems planning, data classification and access management, records identification and disposition scheduling are not integrated. Separate laws and policies govern data access, sharing and privacy, records retention and disposition, data and systems security, disaster recovery and oversight of system development and operations. The public interest would be better served by consolidation of information management planning, reporting and monitoring, and the state would benefit from the streamlined government operations that would result. Finally, consolidation would encourage a better understanding of the multiple interests among government officials and the public.

***The key stakeholders who should play critical role**s in current digital archiving efforts include:*

- *elected officials such as legislators, governors, and secretaries of state;*
- *IT professionals such as chief information officers (CIOs) and agency IT managers and staff;*
- *a diverse group of information creators at both the local and state level agencies; and*
- *IT and digital archiving solutions vendors from the private sector.*

The business case that must be made to gain the support of these "potential investors" includes a compelling analysis of costs, benefits, and risks in the language that each of the stakeholders understands and that speak to the technology, policy, political, and management realities that they face when having to make decisions.

Digital Archiving: From Fragmentation to Collaboration
National Association of Secretaries of State (NASS) and
National Electronic Commerce CoordinatingCouncil

Policies of other states and nations

Many governments and organizations have studied these issues. To date, no one entity has developed a comprehensive, proven and sustainable model, although many have identified practical steps to take that would better position an organization to manage its digital content, to respect its legal mandates and to realize the advantages technology offers. Minnesota is collaborating with other states in seeking solutions for shared problems.

Industry lobbyists have touted standard (but competing and not entirely interoperable) document formats as the panacea for managing and preserving government records. Proponents of a given standard promote their favorite as the universal remedy problems of document access and interoperability. However, responders noted that many new software implementations lack backward compatibility, meaning users who are not able to purchase new software would be left behind.

Governments of other nations have passed laws or instituted policies requiring or recommending particular document formats. Smaller countries such as Norway and Croatia and larger countries such as Poland and Japan are among nations using procurement preferences to influence the use of standard formats. Some local government units have made commitments to open document standards.

Generally, other states have resisted pressure to require agencies to use a particular document format standard. Proposals for state laws specifying formats were defeated in California, Florida, Texas, Oregon and Connecticut. Many have looked to the Massachusetts experience where a broad commitment to the open document format in 2005 was later modified after further study of assistive technology support, cost and other implications. An opportunity exists for states to work together and benefit from collaboration on these issues; a newly formed working group of NASCIO (National Association of Chief Information Officers) will address Electronic Records Management/Digital Preservation.

Management guidelines for state archives as they pertain to electronic documents

The Council of State Archivists Archives Resource Center lists over 20 states with electronic records management guidelines or committees. Typically, the guidelines cover requirements for preservation of official records as detailed in each state's laws, which universally define records based on content and use, not on format.

Minnesota's own state archives program, which is managed by the Minnesota Historical Society, has been a leader in developing guidance on electronic records management. The program provides extensive electronic records management guidelines with detailed information and resources on its website:mnhs.org/preserve/records.

Government programs for electronic records management polices and practices vary in scope. No single government has a complete answer. An organization undertaking a global attempt to schedule all recorded information for, and

---

Sufficient information was not available *to explain the value to be attained or how it would be measured, or to identify the impact and cost of pursing the initiative, opportunity costs, IT infrastructure impacted, technical and non-technical resource requirements, impact on human resource management, assigned responsibilities, points of accountability, and the impact of not pursuing the proposed initiative. In addition, there was no documented strategy to support a managed implementation of ODF-compliant office suite products with specifics that could be measured to set milestones, and demonstrate progress and value achieved.*

Commonwealth of Massachusetts , Auditor Of The Commonwealth, Office Of The State Auditor's Report On The Examination Of The Information Technology Division's Policy For Implementing The Open Document Standard, September 20, 2007

to control the disposition of, every bit of data soon discovers that is an impossible task. Global electronic records management programs are never adequately resourced and remain difficult to sell. One reason for failure is that such broad reaching programs do not prioritize their efforts based on the value of specific records sets. The plain fact is that that some government information is more valuable than other information.

Programs of smaller and more defined scope stand a better chance of success because they recognize the need to set priorities based on the value of the records to be managed. With targeted programs, it is more likely that financial records subject to audit are protected from premature destruction, and agencies with significant private or confidential data can devote more resources to training employees in data protection requirements.

Privacy and confidentiality of electronic records

In Minnesota, public access to information and well as protection from unauthorized access to private information is governed by MS Chapter 13, the Minnesota Government Data Practices Act. The act applies to government data in any form, electronic or otherwise.

Many responders identified the need for good business practices necessary to carry out the government's responsibility to protect privacy of records in any form. Several noted that these practices are not universally applied now and urged attention to rigorous processes, clear procedures, and input from all stakeholder groups, and planning based on multiple goals. Some stated that it is not easy for most employees to understand what is private and urged simplification and clarification of existing standards. A central document management repository was suggested as a way to control access. A state contract for data privacy consulting was suggested to ensure consistency in advice. Oversight of records and information management systems by qualified records managers and experts in HIPAA, data practices and other privacy laws was identified by several responders, who noted that an effective records management system can use records retention schedules to indicate classification of data.

Some responders focused on taking advantage of technological capabilities to aid in protecting privacy by tagging data elements with classifications, tracking viewership of records, maintaining audit logs, and ensuring that systems have redaction capability and the ability to place temporary hold on records. Others suggested using the capabilities of digital rights management systems and content management systems to manage access and modification authorities. Graphical verification systems were identified as good for security but as problematic for the visually impaired. Audio or other alternatives must be provided for those who cannot use a graphical interface.

Among other good business practices cited were due diligence regarding security best practices, passwords, firewalls and encryption; the need to ensure that uniform procedures are followed by

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.*

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

all government agencies; and the importance of integrating state and county computers. Some put a priority on managing human behavior, including providing annual training for all employees, conducting comprehensive security monitoring and assigning responsibility for breaches.

Few responders cited document format standards as a way to protect privacy and confidentiality of government records in electronic form, although one suggested the use of XML to capture instructions on whether data can be shared, and with whom. One stated that ODF does not prevent the protection of privacy.

Standards

- Survey responders who believed that standards could encourage public access suggested that the state adopt the nonproprietary open XML standard because it is developed by international standards bodies, openly documented and available, and developed through consortia.
- Some responders believe that locking into a single standard stifles competition and innovation. Others noted that that formats such as ASCII, PDF, rtf, txt and MS Word doc are pervasive already. Some view PDF as an appropriate document standard because users may download for free the software that allows viewing of the document. (However, it is a proprietary format and a license is required for the software needed to create a document in PDF.)
- Most responders did not distinguish between the format used to store and the application used to display and manipulate documents. Some responders focused on the need for sharing content to be edited and transformed by others, and promoted open document formats to achieve that goal. Others noted that many records are historical and should not be editable.
- Others stated that there already are standards that are not currently followed and urged the state to enforce Data Administration Rules (1205.1500) already on the books, and to follow the record-keeping metadata standards that are already part of the state's enterprise technical architecture.

Use of the World Wide Web

Increasingly, governments are using the web as a primary vehicle for access to and dissemination of public information. This can include "push" and "pull" strategies: actively "pushing" or disseminating information, as well as using it as a passive repository from which users may "pull" desired information. Responders to the survey supported this development and suggested the state use the web as the primary communications vehicle. They encouraged creation of websites that comply with W3C standards, and would support requiring agencies to post documents on websites. They would like to see a consistent look and feel for state websites and the use of "wikis," a popular technology for collaboration, for increased participation.

Better organization of information for public access

When asked what mechanisms and processes Minnesota could establish for accessing and reading electronic records to encourage public access, stakeholders believe public access may be enhanced by better organization and documentation of electronic data. As examples, they cited standard taxonomy, better indexing and including electronic documents in library catalogs. Also mentioned were using metadata (literally, "data about data") tagging to identify whether data is public or whether there are access restrictions. Inhibitors to access may include formats, technologies, and changes in use.

Costs of implementation

Although the survey questions asked about costs of implementing a comprehensive plan for managing electronic records, many of the cost-related comments referred to costs of open document formats, conversion and software.

All responders recognized that there are costs to any system or approach to managing electronic

records. Costs include training, consulting, system integration and maintenance. Human resources, storage, backup, conversion, processing are other costs cited as being needed. Specific costs for indexing, a library function, and data retrieval mechanism for easier retrieval, were also identified.

Those writing in support of an open document format standard stated that open standard costs are an investment in the future, and that conversion costs could be managed by phased implementation. Others writing against a requirement for all documents to conform to one open standard noted that costs escalate when choices are restricted; they advised a free market approach. Others saw the free market approach as a windfall for a single vendor and urged a prohibition on the use of licensed software by government agencies. Recognizing that even free software has implementation and maintenance costs, supporters of free or open software stated that costs would be less in the long run than constant upgrades to newer versions of proprietary software.

Potential savings

Although the survey asked about savings related to implementing a comprehensive plan for managing electronic records, many of the comments received related to benefits of Open Document formats. No responder to the survey specifically identified hard savings, focusing instead on a shift of resources away from a proprietary vendor in favor of a perceived increase in control over upgrades with "free" software.

Related to benefits of open document formats, responders cited the benefits of free upgrade path with free software and stressed that open access is the primary concern with cost a secondary consideration. Supporters of implementing the ODF standard said that costs should be downplayed relative to benefit of not being held hostage to a particular vendor, and viewed migration costs to open source as exit costs from the proprietary vendor. Others urged caution about implementing an open document format, given the pace of innovation in technology.

Related to the benefits of a comprehensive plan for managing government records, responders believed that it would make it easier for end-user accessibility and easier for state employees to locate records. Other cited a lowered risk of lawsuits that may occur when agencies do not follow established records retention policies. One noted that it has been some time since there was any active records management at the state level and it would take some time reestablish the principles.

One responder suggested that costs could be recovered by charging users a fee to submit and retrieve government data unless required to do so.

Interoperability and data sharing

One survey question related to mechanisms and processes the state should use to allow interoperability and data sharing. One responder suggested the state bolster support for and enhance capabilities of data sharing currently going on in Minnesota using a centralized document management system.

Some responders observed that interoperability is not the only goal to consider and should not be the litmus test for state IT investments.

The creation, management and digital archival activities should be addressed by the Enterprise Architecture program *and be defined as a domain within the enterprise IT architecture. Under this umbrella, the business, economic and technical considerations could be presented and debated at the enterprise level.*

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

Other goals include government control, access, choice, privacy, security. The state was urged not to choose something that limits procurement choices, but to let the market choose the winners and losers; interoperability alone may stifle innovation.

Still other responders suggested that interoperability and data sharing may be accomplished by creating systems that are accessible to the widest audience and by designing systems from the beginning to accommodate sharing, web publishing, formats that are readable by any standards-compliant internet browser, wiki or blog creation, PDFs and hyperlinks to free open software.

The question again drew opposite responses from those in favor of establishing a document format standard and those opposed to the establishment of any standard. Supporters of an XML standard stated that it is good for data sharing between disparate applications, and ODF should be used because it is governed by a vendor-neutral standards committee. Others stated that the standard was still in infancy, limited and incomplete, and urged the state to remain flexible, and not stifle innovation.

Long term preservation

Survey responders identified the need for a state plan for archiving with input by information technology professionals, archivists, records managers, and librarians.

Although some responders believe that "storage is cheap" and thus every record should be preserved because you never know what researchers will want next, others say that storage is expensive and good records management and archival practices are to archive only what is really necessary.

Responders urged the state to follow widely used and long established professional best practices (National Archives and Records Administration guidelines; State Archives and Legislative Reference Library standards) with policy-based rules rather than technology standards because the need to the preserve some records (birth death, land) will outlive any technology. Supporters of this approach considered it unrealistic to think that each technology must have a guarantee of "forever" readability and that it is better to plan for migration to each new technology as it emerges.

Changes to records laws

Some survey responses suggested that records laws are long and cumbersome and they should be reduced and compressed, only covering what is necessary. One suggested adding a citizen to the records disposition panel. Another suggestion was to institute a panel to review formats.

Some responders believe records management laws are adequate, just not enforced. Others familiar with the laws, particularly the overlapping definitions, say they are confusing and need to be revised for clarity and consistency.

Conclusions

1. Although the study was prompted by a bill that proposed a particular document format standard for all government documents, the final language called for a study that addressed broader government business goals and purposes including access, preservation, choice and cost.
2. The state's websites, using web standards and available to standard browsers, are the best vehicle for public access. Web standards for viewing documents through Internet browsers are widespread.
3. A common format for the storage of documents that may be opened and used in the application of a user's choice sounds like a good thing for users who need to manipulate the documents. But the marketplace is still in flux, and it is not certain that a single standard will emerge
4. Neither of the proposed standards addresses all the government goals and purposes in the law. In any case, the choice or use of a standard must not be to adopt a standard for the sake of adopting a standard. Any choice must be based solely on the value such a decision would add to government.
5. The Legislature has given the state CIO authority to promulgate information technology standards. If a standard were to be adopted, that is the place to do it, not in statute. Minnesota already has laws and standards on the books that are not complied with. Compliance with existing Minnesota laws related to authorized disposition needs to be addressed.
6. Specialized formats (CAD, geospatial, medical imaging) would need to be exempt as there is no standard that either provides interoperability with those types of formats or is supported by the highly specialized software used for those purposes.
7. Wholesale implementation of any document standard would be costly for the state. How costly is only a guess at this point without doing the sort of detailed costs estimates that are called for in the Commonwealth of Massachusetts Auditor's report. Certainly an effort to convert all existing documents and shift to a common format for all documents moving forward would cost several hundred million dollars. The Legislature, always forced to choose between competing needs, has not shown an interest in such an investment when information systems projects critical to conducting government business and proposed for a fraction of that cost remain unfunded.
8. Minnesota should be involved as appropriate in national standards committees; national standards committees should be involved in international standards bodies. It is the state's obligation to acquire proper tools to conduct government business, but becoming overly involved in marketplace conflicts is not the function of the state's procurement process and distracts from state purposes.
9. Statutory definitions and concomitant confusion are not helpful in achieving integrated goals for information management. Most of the public and most employees don't understand the distinctions between authorized access to *data* and preservation and disposition of *records* and the requirement to deposit *documents* that are part of Minnesota's statutory and organizational legacy. The statutes have to be read together before anyone can figure out whether a given instance of recorded information is a record, data or a document, and thus how it must be treated.
10. Records management, which includes regular disposition of records according to an approved retention schedule, is not widespread or consistent across government entities in Minnesota.
11. Employees need practical advice on what to do with this "stuff" from the time it's created,

through use, and disposition: who has access, how long it needs to be kept, when it can be disposed of. Some agencies carry out this work effectively. Most do not. There is no state-level comprehensive enterprise-wide program that coveres the policy, training and administrative aspects of all of these matters.

12. Preservation is important. Most government records do not need to be preserved long term. Most government records do not need (and in fact should not) be preserved in a format that allows editing. Long-term preservation will require regular refreshing and migration.

Specific recommendations related to adoption of standard document format

1. Technologies are dynamic and change quickly; statutes are more difficult to update. Do not put a technology standard in statute.
2. The business functions of state offices vary widely. "One size does not fit all" with respect to the ways that agencies transact business with their customers. The state must allow for multiple, overlapping standards to cover reasonable breadth of needs in the agencies, including diversity in methods of interaction with customers. Whenever practical, the state should rely on independent standards-setting organizations to develop or adopt open standards, eschewing vendor or other proprietary standards in favor of truly open and independently adopted and maintained mature standards.
3. Particular applications enhance the value of digital content. Reducing all digital content to the lowest common denominator of formats or standards may help to preserve it in some raw state, but it will not preserve all its value. The effort and objectives of document creators must be respected in storing them for future reference. In particular, documents that integrate graphical elements, tables and other components should not be forced into overly simplified formats unable to accurately represent the document purpose and the meaning and intent of the content.
4. Given current unsettled and rapidly changing technologies or document areas, adoption of a standard may not be a wise choice. Premature adoption of a standard in an evolving market may discourage both product evolution and successful implementation. The state must be responsive to market changes, customer needs and evolution of technology
5. Focus on the value added by a standard, and not simply mandate standards for the sake of having the standards or to satisfy other interests beside those of government, its citizens and its business partners.
6. If the implementation of a document format standard is contemplated, a comprehensive business case must be created before any adoption, including TCO (total cost of ownership) cost-benefit analysis, risk analysis and stakeholder analysis. The cost-benefit analysis needs to take into consideration the costs of conversion, training, maintenance, and support, as well as the financial and other impacts on citizens and business partners.
7. If collaboration among the states leads to the adoption of a national standard, that standard should be phased in over a reasonable time period, and existing records that have proven to be of limited or only theoretical value to end users should be "grandfathered" in.
8. Stay active in standards setting bodies so that the state will benefit from opportunities to standardize when it makes good business sense.

General recommendations

1. Update and revise confusing statutes related to government and official records, documents, and data. Clear and integrated definitions can serve as the basis for efficient management of current government information and for rational planning for long-term preservation.
2. Determine what is worth preserving and what is not. The State Archives acquires for permanent preservation only a very small percentage of the records government creates. A good records management program will help government entities determine what

records they need for operational, legal, financial and historical purposes, as well as determine how long those records need to be maintained. It is critical that government entities distinguish operational and in-progress documents from permanent archival record documents in both policy and practice.

3. Restore important records management functions within agencies and at a coordinated statewide level in order to manage government information throughout the entire life cycle of creation, access and final disposition. Standards, guidelines and education can help all agencies and prevent inefficient duplication of effort.
4. Provide adequate resources for thorough training and support of employees who manage government information (data, records, and documents) at every stage in their life cycle.
5. The links between state and local governments are deep, and solutions to issues of format or preservation require analysis of intergovernmental needs and cooperation. Local governments should benefit from the economies of scale, whether it is statewide implementation of technology or research done at the state level.
6. Implement enterprise-wide document management capabilities to facilitate effective document storage, access, and retrieval to support retention and timely disposition practices and ease backup, future conversion and recovery of documents over time. Invest in the creation and implementation of a comprehensive information architecture upon which both policy and standards decisions could be based. These capabilities will require significant planning and resources at the state level. In order to preserve government information, we have to invest in the architecture, hardware, and software to organize it, describe it, back it up, maintain security copies, and migrate material to new formats and applications, as well as to create a front-end interface to provide access.
7. Revise and streamline laws related to authorized access to government data and authorized disposition and preservation of government records in order to improve understanding and compliance. Ensure that system developers are aware of those laws and of the laws and standards pertaining to access to information by persons with disabilities.

Next steps

1. The Office of Enterprise Technology should engage actively with other states, stakeholder groups, and standards groups in developing collaborative approaches to solutions for common problems in government electronic records management.
2. The Office of Enterprise Technology should convene a work group with membership from the Revisor of Statutes, the Legislative Reference Library, and the State Archives to streamline confusing definitions in statutes of records, data and documents. The work group should examine the UETA definition of records as a clear and comprehensive model for refining other definitions.
3. The Office of Enterprise Technology should ensure that a records management component is included in the enterprise document management initiative now moving forward. Given the exigencies of legal discovery and business continuation, it is less costly and disruptive for the state to identify and address records retention requirements sooner rather than later.
4. The Office of Enterprise Technology should identify and address other electronic records policy issues so that actions taken support the best interests of the state for responsible stewardship of information resources.

Appendix A: Background information on the issues

Commonwealth Of Massachusetts, Auditor Of The Commonwealth, *Office Of The State Auditor's Report On The Examination Of The Information Technology Division's Policy For Implementing The Open Document Standard, September 20, 2007* www.mass.gov/sao/auditingpage2007.htm

Ditch, Walter. JISC TechWatch: XML-based Office Document Standards. (August 2007). www.jisc.ac.uk/whatwedo/services/services_techwatch/techwatch/techwatch_ic_reports2005_published.aspx

Minnesota State Archives Electronic Records Management Resources
www.mnhs.org/preserve/records/electronicrecords.htm

Minnesota Legislative Reference Library. Minnesota State Documents, Resources
www.leg.state.mn.us/lrl/mndocs/mndocs.asp

National Association of State Chief Information Officers (NASCIO) All three publications are available at:
www.nascio.org/committees/ea/pubArchive.cfm

Electronic Records Management and Digital Preservation: Protecting the Knowledge Assets of the State Government Enterprise PART I: Background, Principles and Action for State CIOs, May 2007

Electronic Records Management and Digital Preservation: Protecting the Knowledge Assets of the State Government Enterprise PART II: Economic, Legal, and Organizational Issues, July 2007

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, October 2007

National Electronic Commerce Coordinating Council (NECCC). *Digital Archiving: From Fragmentation to Collaboration.* (December 2006) www.ec3.org/Pubs/2006NASS_WhitePaper.pdf

Notes from the 20th Anniversary of the National Archives and Records Administration Panel Discussion, comments by Robert Horton, State Archivist, Minnesota Historical Society May 20, 2005 http://www.archives.gov/about/history/anniversary/panel/bob-horton.html

Pardo, Theresa A., G. Brian Burke and Hycukbin Kwon. P*reserving State Government Digital Information: A Baseline Report.* Center for Technology in Government, SUNY-Albany. (July 2006)
www.ctg.albany.edu/publications/reports/digital_preservation_baseline

Minnesota State Archives Electronic Records Management Resources
www.mnhs.org/preserve/records/electronicrecords.htm

Appendix B Stakeholder Survey

Electronic Documents Study Stakeholder Survey

The terms "records," "data," and "documents" are often used interchangeably; the language that calls for this study itself uses "documents" and "data" without distinguishing the two. But all of these terms have a separate legal meaning and significance in Minnesota, so it will be important to understand the exact distinctions any respondent makes in this survey.

In this study and in the survey, to avoid confusion, the term "electronic records" will be used inclusively. The definition is borrowed from the Uniform Electronic Transactions Act (UETA) (Minnesota Statutes, Chapter 325L). It reads:

> *"Electronic records" mean "records created, generated, sent, communicated, received, or stored by electronic means."*
>
> *"Records" mean "information that is inscribed on a tangible medium or that is stored in an electronic or other medium and is retrievable in perceivable form."*

1. Name Affiliation E-Mail address

2. What mechanisms and processes can the State of Minnesota establish for accessing and reading electronic records to encourage public access?

3. What mechanisms and processes can the State of Minnesota establish for accessing and reading electronic records to encourage interoperability and data sharing with citizens, business partners and other jurisdictions?

4. What mechanisms and processes can the State implement to ensure the privacy and confidentiality of electronic records?

5. Are there mechanisms and processes the State of Minnesota can establish that are specific to the management of electronic records in its various life cycle stages (creation, maintenance, exchange, preservation and disposal)?

6. How should the State address the long term preservation of electronic records? What should the State consider regarding public access to such archived content?

7. What changes, if any, should be made to the government records management provision in Minnesota Statutes? (MN Stat. 138.17-138.226)

8. What constraints and benefits should the State consider regarding the costs of implementing a comprehensive plan for managing electronic records?

9. What should the State consider regarding highly specialized data formats such as CAD, medical imaging, digital art and multimedia?

10. What constraints and benefits should the State consider regarding potential savings or additional costs associated with defined formats?

11. What existing programs, in the private or public sector, for the management of electronic records are appropriate for the State to examine? Please cite specific examples.

12. What standards for the management of electronic records should the State consider adopting or evaluating?

13. What else should the State consider about this subject?

Appendix C Stakeholder Survey Responses

The Office of Enterprise Technology conducted an electronic survey in September and October 2007, with structured questions tied to the legislative requirements. Alternative forms of input were accepted from stakeholders who could not access the web form. Nearly 50 representatives of government agencies, libraries, companies, and citizen interests provided responses to the survey. Responses can be found at:

http://www.state.mn.us/portal/mn/jsp/content.do?subchannel=-536894135&programid=536915535&sc3=null&sc2=null&id=-536894133&agency=OETweb