Electronic Records Management Guidelines
Version 5, March 2012

State Archives
Minnesota Historical Society
345 Kellogg Boulevard West
Saint Paul, Minnesota, 55102-1906
651-259-3260

Shawn Rounds
State Archivist
shawn.rounds@mnhs.org
651-259-3265

Carol Kussmann
Collections Assistant
carol.kussmann@mnhs.org
651-259-3262

The State Archives would like to acknowledge Signorelli & Associates, Inc. for writing and editing the first version of the guidelines.

Table of Contents

Introduction

Summary

You routinely create, use, and manage information electronically in your daily work as you use computers to send e-mail, create spreadsheets, publish web pages, manage databases, and create digitized materials. Because you work for a government agency, Minnesota and federal laws mandate that you treat that information as official government records.

You probably already have a strategy to manage your paper records; because of the pervasiveness of digital files, you must also have a plan to manage electronic records.

Common Questions

As you begin the process of developing an electronic records management strategy, you will find yourself asking many questions, including:

- Which Minnesota laws apply to electronic records?
- How can we use electronic records to help ensure public accountability while protecting non-public records?
- Who is responsible for developing our electronic records management strategy?
- How do we dispose of electronic records?
- Should we manage our electronic records differently from our paper records?
- How do we know what information is an electronic record?
- Is an electronic copy of a record an acceptable substitute for the original?
- Does an electronic record have the same legal significance as a paper record?

Electronic Records Management Guidelines

Because records management laws do not always translate easily into specific technological terms, the State Archives of the Minnesota Historical Society has developed a series of guidelines on basic electronic records management topics.

Purpose of the Guidelines

These guidelines should serve as a starting point as you review your electronic records management practices and develop an appropriate strategy. Each chapter provides an overview of key concepts within the applicable legal framework, questions to spark discussion, and an annotated list of resources as a guide for more detailed research. We recommend that you begin by reading the *Electronic Records Management Strategy* guidelines for a general introduction to key concepts.

Guidelines in the Series

Guidelines in the series include:

- *Legal Framework*. A description of the laws that apply to electronic records in Minnesota.
- *Electronic Records Management Strategy*. Read this set of guidelines first for basic, key concepts in electronic records management.
- *Long-Term Preservation*. Learn how to develop a long-term electronic records preservation plan.
- *Business Case for Digital Preservation*. Learn about the elements that go into making a successful business case (for digital preservation).
- *Metadata*. Become familiar with metadata, its functions, and its importance in managing electronic records, as well as specific metadata standards.
- *File Naming*. Learn about the importance of including a file naming policy in your electronic records management strategy.
- *File Formats*. Review descriptions of common file formats and a summary of the issues surrounding converting or migrating files.
- *Digital Media*. Review digital media storage options for your electronic records.
- *Digital Media Storage*. Learn about physical storage space options and access procedures.
- *Digital Imaging*. Learn about digital imaging, its uses, and legal considerations. Review recommendations for under taking digital imaging projects.
- *Electronic Document Management Systems*. Familiarize yourself with electronic records issues that may arise if you implement an electronic document management system.

- *E-mail Management*. Consider the issues involved in extending your electronic records management strategy to your e-mail messages.

- *Web Content Management*. Learn how to develop a policy for managing your web content that meshes with your electronic records management strategy.

- *Electronic and Digital Signatures*. Learn about the distinction between electronic and digital signatures, and the legal considerations surrounding their use.

Legal Framework

Summary

State and federal laws govern records management practices, whether records are in paper or electronic form. It is important to know about and understand the laws that may affect your records management policies and procedures.

Legal Framework in Minnesota

Electronic records, just like paper records, are subject to specific Minnesota statutes that you must understand and comply with, including general records laws and electronic records laws. Therefore, your understanding of existing Minnesota statutes is crucial as you begin to develop your electronic records management strategy.

General Record Laws

Official Records Act; Minnesota Statues, Chapter 15.17[1]

The Official Records Act is a general records law that mandates that "all officers and agencies" at all levels of government "shall make and preserve all records necessary to a full and accurate knowledge of their activities." This mandate reflects a concern for accountability: since government spends public money on public services, government agencies must be accountable to citizens, government administrators, courts, the legislature, financial auditors, and to history—that is, to future generations. Under the Official Records Act, your agency's chief administrative officer is responsible for creating and preserving government records, including electronic records. This statute also allows you to copy records to another format or storage medium and still preserve the authenticity, reliability, and legal admissibility of the record, as long as the copies are made in a trustworthy process.

Records Management Act; Minnesota Statues, Chapter 138.17[2]

The Records Management Act recognizes that creating comprehensive records and preserving them forever would be an impossibly expensive burden. Instead, the Act creates a mechanism for the orderly and accountable disposition of records in the form of the Records Disposition Panel.

Records Disposition Panel Members
The Records Disposition Panel includes the:

- Attorney General, for expertise on the legal value of records

[1] Minnesota Office of the Revisor of Statutes. *Minnesota Statutes, Chapter 15.17.* http://www.revisor.leg.state.mn.us/stats/15/17.html

[2] Minnesota Office of the Revisor of Statutes. *Minnesota Statutes, Chapter 138.17.* http://www.revisor.leg.state.mn.us/stats/138/17.html

- Director of the Minnesota Historical Society, for expertise on the historical value of records

- Legislative Auditor (for state agencies) or State Auditor (for local agencies), for expertise on the accounting value of records

Records Disposition Panel Functions

The panel reviews and approves or disapproves record retention schedules as well as non-routine requests to dispose of or transfer records. Fundamentally, the panel provides oversight, but does not initiate any actions. If your agency wants to keep records forever, then you never have to work with the panel. However, if your agency wants to do anything else legally with your records, you must submit your proposal to the panel for approval. See *Preserving and Disposing of Government Records*[3] for more information.

Minnesota Government Data Practices Act (MGDPA); Minnesota Statutes, Ch. 13[4]

The MGDPA assumes that government records (including electronic records) should be accessible to the public. However, government agencies create some records that are confidential or sensitive, such as child protection records and adoption records. So, while in theory all records are presumed to be publicly accessible, many exceptions exist. Only the Minnesota state legislature defines these exceptions. Any organization, public or private, that improperly releases data covered by the act could suffer significant penalties.

The Department of Administration's Information Policy Analysis Division offers assistance with interpreting and meeting the requirements of the MGDPA.[5]

Electronic Records Laws

Uniform Electronic Transactions Act (UETA)[6] and Electronic Signatures in Global and National Commerce (E-SIGN)

UETA (Minnesota Statutes, Chapter 325L) and E-Sign were both enacted in 2000. These laws intend to facilitate the use of information technology in government and business by addressing the legal obstacles that exist in a system that was created for paper records and signatures.

The primary message of the laws is that a court may not determine that an electronic record or signature is untrustworthy simply because it is in an electronic format. A court can, though, reject electronic records and signatures because a government agency is creating, using, or

[3] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf
[4] Minnesota Statutes, Chapter 13; http://www.revisor.leg.state.mn.us/stats/13/
[5] Information Policy Analysis Division (IPAD). Home Page. Minnesota Department of Administration. http://www.ipad.state.mn.us/
[6] Minnesota Statutes, Chapter 325L; http://www.revisor.leg.state.mn.us/stats/325L

managing them in an untrustworthy system or manner. One indicator of untrustworthiness would be an agency's failure to follow the state laws governing records.

Uniform Electronic Legal Materials Act (UELMA) *(a model law)*

The National Conference of Commissioners on the Uniform State Laws (NCCUSL), also known as the Uniform Law Commission (ULC), drafted and approved in 2011 a model law that addresses the authentication and preservation of state electronic legal materials. The Uniform Electronic Legal Material Act[7] (UELMA) establishes an outcomes-based, technology-neutral framework for providing online legal material with the same level of trustworthiness traditionally provided by publication in paper form. The Act requires that official electronic legal material be: (1) authenticated, by providing a method to determine that it is unaltered; (2) preserved, either in electronic or print form; and (3) accessible, for use by the public on a permanent basis. Information and links to the recent activities, including state enactments, can be found on the Uniform Law Commission website.[8]

[7] National Conference of Commissioners on Uniform State Laws. *Uniform Electronic Legal Materials Act.* July 18, 2011. http://www.uniformlaws.org/Shared/Docs/AM2011_Prestyle%20Finals/UELMA_PreStyleFinal_Jul11.pdf
[8] Uniform Law Commission. *Electronic Legal Materials Act Committee Home Page*. 2012.http://www.uniformlaws.org/Committee.aspx?title=Electronic%20Legal%20Material%20Act

Electronic Records Management Strategy

Summary

The arrival of the Information Age means that much of our history is now recorded in electronic format, including your agency's activities. Because of that, you need to develop a strategy for managing electronic records. A government agency's electronic records management strategy must conform to legal mandates, as well as reflect your preferred management practices and technological options.

When you begin to develop your electronic records management strategy, you should aim for a policy that integrates:

- The legal framework as it applies to your agency
- All interested stakeholders (e.g., record creators, the public, information technology staff, records management staff)
- All relevant aspects of your electronic records
- Your preferred management procedures and technologies
- Long-term storage and access needs (both legal and operational)

A sound, integrated strategy reflects the relationship between records management and your operations, and ensures that you manage records in a way that supports your daily work, supports long-term operational needs, and meets your legal requirements.

Because different stakeholders throughout an enterprise have different needs and roles in electronic records management, the development of your electronic records management strategy requires joint planning, communication, and training.

Legal Framework

Your strategy must conform to the legal mandates in such areas as:

- Providing public accountability
- Distinguishing public from not-public records
- Creating records retention schedules and carrying out disposal actions
- Developing and sustaining a trustworthy process for electronic records management

Refer to the *Legal Framework* chapter for more information on legal mandates and legal frameworks in general.

Key Concepts

As you develop an electronic records management strategy, you will need to be familiar with the following key concepts:

- The State of Minnesota's definition of a record
- Records series
- The components of an electronic record
- The records continuum
- Records management goals
- Long-term retention approaches
- General records retention schedules
- Storage options

Definition of a Record

The Records Management Act[9] defines government records as:

> Cards, correspondence, disks, maps, memoranda, microfilms, papers, photographs, recordings, reports, tapes, writings, optical disks, other data, information, or documentary material, regardless of physical form or characteristics, storage media or conditions of use, made or received by an officer or agency of the state and an officer or agency of a county, city, town, school district, municipal subdivision or corporation or other public authority or political entity within the state pursuant to state law or in connection with the transaction of public business by an officer or agency.

In short, an official record includes all information, *regardless of format*, created or used in the course of a government business function or transaction.

The definition *excludes*:

- Library and museum material made or acquired and kept solely for reference or exhibit purposes
- Extra copies of documents maintained only for the convenience of reference

[9] Minnesota Statutes, Chapter 138.17; https://www.revisor.leg.state.mn.us/statutes/?id=138.17

- Stock of publications and processed documents
- Bonds, coupons, or other obligation or evidence of indebtedness, the destruction or other disposition of which is governed by other laws

An electronic record is a record created, generated, sent, communicated, received, or stored by electronic means. Like paper records, electronic records require a long-term records management strategy.

Records Series

Your electronic records are organized into records series. A *records series* is a set of records grouped together because they relate to a particular subject or function, or result from the same activity. All records fall into a records series, and each records series should be managed according to an appropriate records retention schedule.

By managing related records as a group, you can efficiently preserve and dispose of your records. For example, all records (regardless of format) relating to a particular committee's activity on a single issue may constitute a records series that must be preserved for some length of time before disposition.

Your agency will need to organize its own records series based on its unique needs within the legal framework.

Record Components

The components of any record include:

- *Content*. Factual information in the record that documents government business
- *Context*. Information that shows how the record is related to the business of the agency and other records
- *Structure*. Technical characteristics of the record (e.g., file format, data organization, page layout, hyperlinks, headers, footnotes)

Records Continuum

Aside from reflecting your legal requirements, a successful long-term records management strategy reflects the records management continuum.

The records continuum concept is the idea that different stakeholders create, use, manage, and retain records, not in discrete stages, but at different points throughout the record's existence. The continuum concept recognizes that records pass through identifiable stages; however, these stages are reference points, not separate functions. In other words, a record is not simply created, passed to a records manager for short-term storage, and then passed to an archivist for long-term storage. Instead, each person's activities will have an effect on all the others in the continuum. Their roles and responsibilities should be coordinated, not organized autonomously.

The continuum concept outlines four actions that recur throughout the life of a record. These actions are:

- *Identification*. Determining what constitutes a record
- *Intellectual control*. Making decisions about the record
- *Provision of access*. Enabling users to access the records
- *Physical control*. Managing the physical location and format of the record

Each person who touches the record performs one or all of these activities. For example, the records creator, records manager, and archivist all manage the physical location of the record. Therefore, all these people should collaborate on a comprehensive and well-managed electronic records management strategy.

Records Management Goals

Although the specific strategy that your agency develops and implements will be unique, all strategies share common goals. No matter what your final strategy, the records that exist in your agency should be:

- *Trustworthy*. Trustworthy records contain information that is reliable and authentic. A key aspect to trustworthiness is legal admissibility, i.e., whether your records will be accepted as evidence in court.

- *Complete*. Your records should have all the information necessary to ensure their long-term usefulness. You will also need to capture and maintain the necessary metadata about your records. *Metadata* is the "data about the data" that documents the relationship of the record to your agency's activity and to other records. Metadata ensures that you can find and use your records. Metadata includes such elements as the record's creator, the date of creation, and the record series to which the record belongs. (For more information on metadata, refer to the *Metadata* chapter in these guidelines).

- *Accessible*. You should be able to locate and access your records in a way that meets your needs and the needs of all other concerned parties. Some records may need to be immediately accessible, while others may not. As outlined in the MGDPA, records are assumed to be accessible to the public, unless categorized as not-public by the statute.

- *Durable*. You also want to ensure that your records are durable. In other words, they must be accessible for the designated records retention period and stored, as appropriate, "on a physical medium of a quality to ensure permanent records," as stated in the Official Records Act [Minnesota Statutes, Chapter 15.17]. For more information on records storage, refer to the following chapters in these guidelines:

- *Digital Media* for more information about digital media options available for electronic record storage
- *Digital Media Storage* for information about the physical requirements for storing electronic records

Long-Term Retention Approaches

You have two viable, often compatible and complementary, approaches for the long-term retention of your records:

- *Conversion*. When you convert a record, you change its file format. Often, conversion takes place to make the record software independent and available in an open or standard format. For example, you can convert a record created in Microsoft Word by saving it as a Rich Text Format (RTF) file or to PDF/A. (For more information on file formats, refer to the *File Formats* guidelines.)
- *Migration*. When you migrate a record, you move it from one computer platform, storage medium, or physical format to another. For example, you may need to migrate records from old magnetic tapes to new ones or to a different medium entirely to ensure continued accessibility. (For more information on storage media, refer to the *Digital Media* guidelines.)

As you explore conversion and migration options, consider which media are appropriate for long-term retention. You may discover that another medium altogether (e.g., paper or microfilm) is the best option. You may also determine that you want to combine approaches, such as converting all files to an open format and migrating them to a single platform and storage medium. (For more information on migration and conversion, refer to the *Long-Term Preservation* guidelines.)

General Records Retention Schedules

Your electronic records management strategy should include records retention schedules. A records retention schedule is a written document that lists types of records and discusses how long they should be kept and whose purpose it is to serve as an on-going authorization for the management and disposition of records in all forms including paper and electronic records.

Many local government entities, because they have similar office structure and responsibilities, have developed general records retention schedules for all the records commonly created by their members. General records retention schedules exist for state government, cities, townships, school districts, counties, and courts.[10] These general records retention schedules meet the legal requirements for each type of local agency. Your agency can adopt the appropriate general records retention schedules for your type of organization in whole, or in part, or change individual components to create a unique schedule. You may also initially choose to develop a specific schedule for your agency. However, you must submit any proposed changes to the Records Disposition Panel for approval. (For more information on the Records Disposition Panel, refer to the *Legal Framework* chapter of these guidelines as well as to the *Preserving and Disposing of Government Records* booklet, found in the Annotated List of Resources at the end of this chapter.)

Storage Options

Your options for storage include:

- *Online*. Properly designed storage in your computer system may provide full access to appropriate users. Online access means that the record is accessible immediately through your network (e.g., on your network server). This option maintains the greatest functionality.

- *Near-line*. Near-line storage includes storage in a system that is not a direct part of your network, but that can be accessed through your network (e.g., an optical media jukebox). This option maintains a moderate amount of functionality.

- *Offline*. Offline storage refers to storage that is not accessible through your network (e.g., removable media such as magnetic tape). This option retains the least amount of functionality while still maintaining records in an electronic format.

- *Paper or microfilm*. Printing records onto archival-quality paper for storage or outputting them to microfilm may be acceptable as long as the complete record, including all components and metadata, is included.

Key Issues to Consider

Now that you are familiar with some key concepts in electronic records management, you can use the questions below as you develop your own strategy. The careful consideration of these questions will help ensure that:

- All relevant stakeholders agree to the process and are ready to use the procedures outlined in the strategy once it is implemented.

- The strategy meets your legal requirements, such as public accountability, records retention

[10] Example retention schedules for each can be found on the Minnesota State Archives website at http://www.mnhs.org/preserve/records/retentionsched.html

schedules, and trustworthiness.

- You maximize efficiency by working with other agencies and gaining from their experience.

Discussion Questions

- What legal issues do we face? Who will need access to our records (e.g., the public, other government agencies)? Do we have information that *must* be accessible to the public? Do we have information that is not-public as classified by the MGDPA that must *not* be disclosed to the public (e.g., social security numbers, adoption records)?
- Can we adopt one of the general records retention schedules, or do we need to modify or create an agency-specific records retention schedule and seek approval from the Records Disposition Panel?
- What sort of appraisal process will we use to determine which records to keep? How will we ensure that this process identifies all records as defined by the law?
- What are the roles of different groups and individuals in our organization in ensuring a coordinated process? How can we facilitate planning, communication, and cooperation among all individuals who create and use the electronic records? What level of control should different individuals and groups have?
- Can we cooperate with other government agencies to streamline the process and save money or time?
- What best practices can we identify and apply to our own situation?
- What is the life cycle of our data? When should we capture records? How can we describe our records continuum? At which phases along our continuum do we need to actively manage the record? Would we benefit from developing a model of our operational process to aid in this discussion?
- How will we ensure long-term preservation and access? What are our requirements under the law?
- What are our options for long-term retention? What are the advantages of each option? How would each option work in our particular situation? What is our budget?
- What technological resources do we have available? How much of our chosen process can we or should we automate?
- What staff training do we need to ensure the staff complies with the new procedures and policies?
- What elements of the electronic records do we need to keep (e.g., text content only, graphical

appearance, interactivity)?

- What metadata do we need to collect and preserve?

Annotated List of Resources

Minnesota Historical Society, Minnesota State Archives. *Preserving and Disposing of Government Records*. St. Paul: Minnesota Historical Society, May 2008.
http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

Developed for Minnesota government agencies, this overview of the basic principles of records management includes chapters on defining a government record, taking inventory of your records, developing records retention schedules, preserving archival records, disposing of records, and setting up a records storage area. A list of resources for more information is included, as well as information about applicable state law regarding electronic records management. Originally published by the Minnesota Department of Administration in July 2000, the guide was updated jointly by the Minnesota Historical Society and the Minnesota Government Records and Information Network (MNGRIN) in 2008.

Minnesota Department of Administration, Office of Enterprise Technology. *Minnesota Enterprise Technical Architecture*. Version 2.02, 2006.
http://mn.gov/oet/

The Minnesota Office of Enterprise Technology (OET) is charged with establishing and maintaining a state information architecture as specified in Minnesota Statue, Chapter 16E.04 Subdivision 2. According to the OET, "This technical architecture is established to describe technology components of the State's information infrastructure and their individual principles, practices and standards that are to be used to guide the development and delivery of all information systems services. The architecture will provide a reference so that various groups of government IT professionals have a consistent view of the information systems infrastructure and the methods that they employ to develop and deliver information systems services." Chapter 4, Data and Records Management, describes the framework for managing information resources, as well as the standards and guidelines that apply.

ARMA International.
http://www.arma.org/

ARMA is the leading professional records management association. The ARMA website offers practical guidance on a wide range of topics, including electronic records management, E-discovery, information security, and standards and best practices.

Council of State Archivists (CoSA). *Archives Resource Center: Electronic Records*.
http://www.statearchivists.org/arc/states/res_elec.htm

State by state listings of policies and guidelines relating to government electronic records management.

Long-Term Preservation

Summary

During the course of routine business, your agency generates thousands upon thousands of electronic records, from e-mail to web pages to complex e-government transactions. Most are useful for only a short period of time, but some you may need to keep permanently. For those records, you will need to implement a well-considered, well-documented plan for their preservation in order to ensure that they remain trustworthy and useful over time. Tools such as migration, conversion, metadata, and eXtensible Markup Language (XML) will help you not only preserve your records, but also realize their full value.

Legal Framework

For more information on the legal framework to consider when developing a preservation plan for your records, refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[11]. Also consider the Information and Communications Technology Policy[12], which mandates state agency compliance with Minnesota's enterprise technical architecture "to ensure that individual agency information systems complement and do not needlessly duplicate or conflict with the systems of other agencies. . . . [and to] promote the most efficient and cost-effective method of producing and storing data for or sharing data between those agencies." Section 16E.07[13] of this same statute establishes the North Star portal as the state's official online government information service with the idea that "the greatest possible access to certain government information and data is essential to allow citizens to participate fully in a democratic system of government."

Key Concepts

The value of your information justifies your investment in information technology. There is no point to an agency investing large sums in hardware and software if it cannot preserve the use-value of the information it creates, exchanges, and stores. In the short term, this is often not a problem. But, over time, it will be. As technology changes, current hardware and software will become obsolete, and then you might face some hard choices. The challenge is to preserve the usefulness and trustworthiness of your information in an efficient and cost-effective manner.

Any preservation plan for electronic records must take into account the changes in hardware and software, the limitations of storage media, and the potential use-value of your information. As you begin exploring your options, you will need to be familiar with the following:

- Needs Assessment
- Physical Storage Options

[11] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf
[12] Minnesota Statutes, Chapter 16E.04; https://www.revisor.mn.gov/statutes/?id=16E.04
[13] Minnesota Statutes, Chapter 16E.07; https://www.revisor.mn.gov/statutes/?id=16E.07

- File Format Options
- Digital Preservation Techniques
- Preservation Planning
- E-government and Collaboration

Needs Assessment

As a first step in developing your preservation plan, you should do a needs assessment to help guide your decisions. While the complexity of such an analysis will vary from situation to situation, these basic components should always be included.

First, you need to understand the value of your information. The value of your information will justify your investment in technology, over the short- and the long-term. Minnesota's enterprise technical architecture notes that information is the state's most important asset. But all information is not created equal; some has much more value than others. Some of your information, as records, will have legal and evidentiary significance and may well demand special attention. Most of the information you want to preserve will be important to your agency's mission or, increasingly, to the business of other agencies as well. As e-government develops in complexity and sophistication, more and more agencies will be expected to work within the framework of a common technological architecture and to share the information they create.

The practical side of understanding the value of your records is determining their retention requirements. How long do you really need to keep them? Why are you keeping them? Do they have to be kept in electronic format or is there another, more cost-effective option for long-term storage? For instance, a word processing document might be printed and kept as a paper record without losing any of its value. In contrast, printing a web page means a significant loss of information and functionality.

It is also important to ascertain if access to certain data in your records is restricted by statute. The state's Government Data Practices Act and some federal laws, such as the Health Insurance Portability and Accountability Act (HIPAA), will determine if data needs to be protected as confidential or non-public. If it does, then you will need to ensure that your long-term storage and access policies account for those obligations.

In the broadest sense, the demands governing the access and use of your records will determine what preservation options are most appropriate and will dictate the metadata you should create and store along with the records. Metadata is the "data about the data," that allows you to manage, find, and evaluate your information over time. Minnesota's enterprise technical architecture includes metadata standards for GIS data, web content management, and recordkeeping. There are a number of international standards that are pertinent as well. While all-important for the long-term preservation of data, metadata takes on additional significance when you share your information because others must understand the information's structure and content in order to put it to fullest use. For more information about metadata, refer to the

Metadata chapter of these guidelines and the Needs Assessment[14] document created by the Minnesota led National Digital Information Infrastructure and Preservation Program (NDIIPP) project.

Physical Storage Options

As mentioned, choosing the most appropriate storage option for your situation will depend upon your records' access requirements. There are basically three options available to you:

- *Online storage*. Records are kept on a server or hard drive and are immediately available for use over a network. This option is best for records that must be accessed frequently.

- *Near-line storage*. Records are stored on media such as optical disks in jukeboxes or tapes in automated libraries which are attached to a network. Because retrieval is slower than with online storage, this option is most appropriate for records that are accessed occasionally.

- *Offline storage*. Records are stored on removable media and must be manually retrieved. This option provides the slowest access and should be used for records that are only rarely needed.

If you choose near-line or offline storage, you will need to consider what media will best suit your needs. To do this, you should start by analyzing your current and projected volume of stored records, along with the size of the files themselves and any associated metadata. Also take into account any security requirements, such as viewing, use, and modification restrictions.

Different media have different storage characteristics. For more information on media types, storage options and how this may affect your preservation strategy, please review the information in the *Digital Media* and *Digital Media Storage* chapters of these guidelines.

File Format Options

Most records are created using specific, proprietary software applications. Over time, these applications will be upgraded or be phased out altogether. Because upgraded applications may or may not be able to read files created with previous versions, backward compatibility is not a given and cannot be counted on as a preservation tool. Maintaining the software on your own is an option, but over and above the question of costs, that carries the risk the software will fail in time, leaving you with no way to access your records. One common alternative is continually to convert your files from version to version and format to format as your software environment changes.

While non-proprietary formats are the ideal for the long-term preservation of files, they are few

[14] Minnesota Historical Society. *Needs Assessment*. National Digital Information Infrastructure and Preservation Program. February 2012. http://www.mnhs.org/preserve/records/legislativerecords/carol/docs_pdfs/NeedsAssessment022012.pdf

in number and each has its limitations. ASCII or plain text will capture data in the lowest common denominator of formats, losing structure and functions in the process. Rich Text Format (RTF) is a Microsoft format, although it is supported by a variety of vendors and software applications. Portable Document Format (PDF), a popular choice for file sharing and storage, is an Adobe product. Because Adobe makes PDF's specifications publicly available, many believe that it is an open standard when, in fact, the company is under no obligation to continue this practice into the future. Furthermore, PDF has a problem with backward compatibility, with newer versions often incorrectly rendering files created with older ones. To address these problems, an archival version, PDF/A, was developed and became an ISO standard in 2005.[15]

For long-term preservation and use, eXtensible Markup Language (XML) is currently a good choice of formats. An international standard since 1998, XML is both a file format and a text-based, self-describing, human-readable markup language that is independent of hardware and operating systems. Because it is infrastructure-independent, XML is one of the best solutions for re-purposing the content of your records and/or sharing them with others. Proper use of XML requires a certain amount of planning and up-front commitment of money and time, but its structured nature makes it suitable for automation and will allow you to more easily take advantage of whatever new open formats will follow in the future.

For more information on this and additional file types and their associated formats, refer to the *File Formats* guideline in this series.

Digital Preservation Techniques

There are several approaches, some more practical than others, to ensure that electronic records remain useful over time. One is to save all of the hardware, software, and documentation needed to support the records. Known as the "computer museum" approach, it is not very realistic on a large scale because, given how rapidly hardware and software environments change, it means storing and maintaining huge quantities of outdated equipment with no assurance that any of it will work when needed.

Emulation has a similarly antiquarian flavor. Emulator programs simulate the behavior, look, and feel of other programs, thus preserving the functionality of the records in their original format without the necessity of saving the original equipment and software. However, emulation has so far proven more attractive in theory than in practice. There are few examples of success using this approach, and costs have proven high. It has a further limitation in that, at best, emulation simply reproduces earlier, less sophisticated versions of an application. Given all the expenses of technology, it seems problematic to limit the value of information by preserving it in a static framework.

Encapsulation is a third approach to preservation. It involves combining the object to be preserved with all of the necessary details of how to interpret it within a wrapper or package, all

[15] International Organization for Standardization. *ISO 19005-1:2005. Document Management – Electronic document format for long-term preservation – Part 1: Use of PDF 1.4 (PDF.A-1.)* 2011. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=38920

possibly formatted in XML. While appealing in its comprehensiveness, encapsulation has several drawbacks: file sizes are large because of all of the included information; format specifications must be determined; the encapsulated records must somehow be generated, usually separate from the act of record creation; and the encapsulated records must still be migrated over time.

The most common approach to preserving electronic records involves a combination of two other techniques: migration and conversion. Migration is the process of moving files to new media (also known as "refreshing") or computer platforms in order to maintain their value. Conversion entails changing files from one format from one to another and may involve moving from a proprietary format, such as Microsoft Word, to a non-proprietary one such as a plain text file or XML. To avoid losing data in the process, you should perform initial tests and analysis to determine exactly what changes will occur and whether they are acceptable. With both migration and conversion, special attention must be paid to also maintaining the accessibility of any associated metadata. When properly planned and executed, the migration and conversion approach probably represents the easiest and most cost-effective preservation method available today.

Preservation Planning

A preservation plan should address an institution's overall preservation goals and provide a framework that defines the methods used to reach those goals. At a minimum, the plan should define the collections covered by the plan, list the requirements of the records, practices and standards that are being followed, documentation of policies and procedures related to preservation activities, and staff responsibility for each preservation action. It is important to remember that preservation activities are not static and that the preservation plan will need to be reviewed and readdressed on a regular basis to remain viable and useful.

The costs of preservation will be a major factor in the development of your plan. To some degree costs will help determine the level of your preservation efforts. Often there is not enough money available to preserve all electronic records for the long-term. Understanding financial resources allows you to make informed decisions on what to preserve. Without this understanding even the best laid preservation plans will fail.

The Electronic Resource Preservation and Access Network (ERPANet) divides costs into four major categories: technical infrastructure, financial plan, staffing infrastructure, and outsourcing costs.[16] Technical infrastructure costs include equipment purchase, maintenance, and upgrades necessary to keep networks online and adjust to software and hardware obsolesce. A solid financial plan must be backed up with a commitment to long-term funding. Staffing costs include the costs of hiring and training employees. Any services that are outsourced will have a direct effect on your preservation costs. Costs also depend on the record format, level of security required, and the length of time the materials need to be preserved.

[16] ERPANET. Erpa Guidance: Digital Preservation Policy Tool. Electronic Resource Preservation and Access Network. September 2003. http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf

A cost-benefit analysis should be done to analyze each aspect of your workflow to determine the most cost-effective method of preserving identified records. If funding is limited or changes over time, you may need to reanalyze and possibly scale back the amount of materials you are able to preserve. Choices will have to be made. Your retention schedules, needs and risk assessments will be able to help you make these decisions. "Although the costs of preserving digital materials might be high, the cost, consequences and implications of not having a digital preservation policy may be higher and in some cases they could affect the feasibility of the preservation."[17]

When developing a preservation plan, there are many models that can be used as a guideline. A few of these models include:

- Digital Preservation Policies Study Part 1: Final Report October 2008[18]
- Northeast Document Conservation Center Digital Preservation Policy Template[19]
- Digital Preservation Policy Tool[20]
- Preservation Management of Digital Materials: The Handbook[21]
- OCLC Digital Archive Preservation Policy and Supporting Documentation[22]

In addition, specific examples of completed preservation plans can be found online and cover a wide range of topics from how to preserve historical sites and buildings to preserving digital objects.

The following outlines the basic structure of a digital preservation plan and highlights the common points from the resources above.

I. *Purpose Statement:* Why are you writing the preservation plan? Why is digital preservation important to your institution?

II. *Relation Statement:* How does the document relate to others across the institution? Does it complement current records management policies?

III. *Objective Statement:* What are the goals of digital preservation? Are the goals based on a specific project or do they reach the institution level? What are the overarching goals of your preservation program? You may want to include both short- and long-term

[17] ERPANET. *Erpa Guidance: Digital Preservation Policy Tool.* Electronic Resource Preservation and Access Network. September 2003. http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf
[18] Beagrie, Neil et al. *Digital Preservation Policies Study Part 1: Final Report October 2008.* Charles Beagrie Limited for JISC. 2008. http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf
[19] Northeast Document Conservation Center. *NEDCC Digital Preservation Policy Template*. http://www.nedcc.org/resources/soda/downloads/SoDAExerciseToolkit.pdf
[20] ERPANET. *Erpa Guidance: Digital Preservation Policy Tool.* Electronic Resource Preservation and Access Network. September 2003. http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf
[21] Neil Beagrie and Maggie Jones. Preservation *Management of Digital Materials: The Handbook*. Digital Preservation Coalition. November 2008. http://www.dpconline.org/advice/preservationhandbook
[22] Online Computer Library Center. *OCLC Digital Archive Preservation Policy and Supporting Documentation.* Online Computer Library Center. August 8, 2006. http://www.oclc.org/support/documentation/digitalarchive/preservationpolicy.pdf

goals.

IV. *Periodic Review Statement:* How often will the preservation plan be reviewed? Is it based on a schedule or on an event that triggers review? Record what changes are made, who made them and when. This will help ensure your preservation plan is sustainable over time.

V. *Descriptive Statement:* What materials will be preserved? What will not be preserved? What formats are supported? How long will each category of records be preserved? How will access be provided? Who has access to the records? Where are files stored? What hardware and software are being utilized?

VI. Implementation Plans: Documentation that explains the overall details of the preservation plan such as:

a. *Staff responsibilities:* What staff position is responsible for ensuring the preservation plan is carried out?

b. *Financial responsibilities*: Who is responsible for the long-term monetary sustainability of preservation activities including changing technologies, necessary staff, storage space, physical media, and what is required for following set preservation strategies? How will these goals be achieved and sustained?

c. *Record Requirements:* What metadata is required? What metadata standard is being used? What file formats are accepted? What file sizes are accepted? What is the file structure? Include descriptive information about content and context.

d. *Access and Use Restrictions:* Are the records public or non-public? Are there any restrictions based on intellectual property rights? Include specific restrictions in the plan? Provide attribution statements if necessary.

e. *Best Practices, Standards:* What best practices are being followed? For what records? What standards are being followed? For what records? What metadata standard is being followed? What file format standards are being followed? Naming conventions?

f. *Risk Management:* Understand the risks of your digitization plan including what risks your chosen file formats might pose, what risks the acts of migration might pose, and what risks there are to your IT infrastructure. Have a disaster plan that includes disaster recovery procedures in case of a disaster.

g. Stakeholders: Who uses your digital files? Who is dependent on them? Do you work with others? Do others deposit files?

h. *Preservation strategies:* Define what is being preserved; this should include describing the content, structure of the content, and the relationships between documents. Define the type/s of strategies being used to preserve the files such as migration or conversion and how often these strategies are to be carried out? Is there a time schedule or do triggers determine when files are migrated? Do you use outside services? For what? What are their procedures and processes?

i. *Storage Requirements:* What type of storage environment is necessary? What media type/s will be used for storage? Are files backed up? How often? With what process? Where are files stored? If offsite storage is being used, include information on the contracts and the vendor's requirements for file formats etc…

j. *Quality Control/Security Measures:* What methods will be used for quality control? How often will integrity checks be made? How? What processes are you using to ensure the trustworthiness of the files? How is their authenticity and integrity preserved? How will provenance be tracked? How are access and use restrictions controlled? Are audit trails and logging activity processes necessary to ensure trustworthiness? What other security measures are being taken? Who is responsible for maintaining such processes?

Glossary: A glossary of terms may be useful if people who are unfamiliar with digital preservation will be accessing the preservation plan.

E-Government and Collaboration

The State of Minnesota's e-government framework should influence your preservation plans. The long-term preservation of records will demand a variety of investments and decisions that will involve time, staff, technology, and specialized expertise. Practically speaking, the state probably cannot afford to have every agency make all those investments independently. Similarly, no agency, even with the best of intentions, can consistently make all those decisions correctly. Finding effective and economic solutions means working together.

The state's enterprise technical architecture reflects the idea of collaboration. In order to facilitate the development of e-government, the architecture identifies a series of issues, approaches, and standards that will make your agency's investments in information technology more likely to succeed. At the same time, these also will facilitate the long-term preservation of digital resources through sharing services and solutions. In developing your preservation strategy, start by looking at what other agencies are doing and what you can learn from their experiences.

Key Issues to Consider

Long-term retention requires long-term preservation. To ensure long-term preservation, a preservation plan and associated policies should be developed. The foundation of your preservation plan should be your needs assessment, as well as an analysis of the costs, benefits, and risks involved with each of the options you are studying. Your records management, information technology, and legal staff should all be involved in the process to make sure your plan meets your business requirements and fits in with your general electronic records management strategy. Be sure to document your decision-making process in addition to your choices and plans for implementation.

At the minimum, your preservation plan should include the following items:

- Rationale and requirements for your preservation program.
- List of relevant records series and their retention and access requirements.
- Explanation of the selected preservation technique(s), including schedules for preservation actions, quality assurance testing, backups, etc. and instructions for documentation.
- Pointer to a business continuity or disaster recovery plan.

Once completed, your preservation plan should not gather dust on a shelf. Rather, it should be a reference document for all preservation activities, and it should be kept up to date as your situation changes (e.g., changes in use needs, hardware, software, media, security/access requirements, retention periods, legal mandates).

Discussion Questions

When beginning to develop a preservation plan, you will be faced with many choices. These are just a few of the questions you should ask during the process.

- How long do we need to keep these records? What will be the costs associated with such preservation tasks as migration and conversion over time?
- What best practices can we identify and apply to our situation? Can we cooperate with other agencies or organizations to share expertise or save money?
- Do we need to keep the records in electronic format or is another format, such as paper or microfilm, more appropriate? How much functionality do we need to retain over time?
- How often are these records accessed? What is the best storage solution (e.g., online, near-line, offline)?
- What is the most appropriate storage media for the records? How will we ensure that we retain the hardware necessary to handle the media? What documentation should we collect and maintain regarding the media and hardware?
- How will we ensure that the record content is accessible and readable over time? Is the format and necessary software proprietary or non-proprietary? What documentation should we collect and maintain regarding format and software?

- How will we perform periodic quality assurance checks to ensure accessibility and trustworthiness over time? How will we document these checks?

- What indexing and metadata schemes should we employ to ensure that the records can be easily located and evaluated for use?

- How will these records be used? Will they be shared with others inside our organization? Outside? Would XML enhance the use-value of the records?

- Have the records been compressed or encrypted? If so, how does this fit into our management plan?

- Are there data access issues that require special security measures?

- What hardware and software configurations are we moving to in the foreseeable future? How do these records fit in with that plan?

- What staff training is necessary to ensure compliance with the preservation plan?

Annotated List of Resources

Primary Resources

Cornell University. *Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems*. ICPSR: Inter-University Consortium for Political and Social Research. 2007.
http://www.dpworkshop.org/dpm-eng/eng_index.html

> A resource on digital preservation that provides a general background on preservation issues. Various preservation strategies are defined, including bit-stream copying, refreshing, back-ups, migration, replication, standards, normalization, and emulation as well as others.

Digital Curation Centre. *Resources for Digital Curators*. May 5, 2009.
http://www.dcc.ac.uk/resources

> The Digital Curation Centre in the UK is a central location for scientists, researchers and scholars to learn more about digital curation and preservation. The Resource Centre provides access to information about standards, legal issues, and current research. The Digital Curation Manual provides information about appraisal and selection, preservation metadata, archival metadata, and preservation strategies as well as other related topics. A life-cycle model provides a visual interpretation of the life-cycle of a digital object.

Beagrie, Neil and Brett Scillitoe. *The Handbook. [online version]* The Digital Preservation Coalition. Updated December 2008.
http://www.dpconline.org/advice/preservationhandbook

> This handbook covers many aspects of digital records including preservation. The section on preservation provides a strategic overview of digital preservation and covers preservation issues of digital files including technological, organizational, and legal issues. A decision-making tree addresses the selection of digital materials for long-term retention based on rights and responsibilities, technical issues, documentation and metadata and overall cost.

Government Record Branch of North Carolina. Electronic Records.1/25/2012.
http://www.records.ncdcr.gov/erecords/default.htm#guide

> A resource created specifically for government records in North Carolina that provides information on electronic record best practices, data transfer guidelines (including a user guide and video tutorials for BagIt), and email, web site, and digital imaging project guidelines.

Lefurgy, Bill. *Digital Preservation Tools and Services*. Agogified: Technology, Information and Culture, Mashed and Served. 2010.
http://agogified.com/tools-and-services

> A resource that lists some of the available digital repository services and systems, file format management utilities and references, file integrity utilities, and file transfer specifications and utilities.

Information Standards Quarterly (ISQ). *Special Issue: Digital Preservation.* National Information Standards Organization. Spring 2010, Vol. 22. Issue 2.
http://www.loc.gov/standards/premis/FE_Dappert_Enders_MetadataStds_isqv22no2.pdf

> A discussion of digital preservation metadata, trustworthy digital repositories, audio-visual digitization guidelines, and digital preservation education.

Lavoie, Brian and Lorcan Dempsey. "Thirteen Ways of Looking at… Digital Preservation". *D-Lib Magazine*. Vol. 10, No. 7/8. July/August 2004.
http://www.dlib.org/dlib/july04/lavoie/07lavoie.html

> Digital preservation means many things, and this article shows the many different ways that people look at the idea of digital preservation which provides insight into the complexity of the idea of preservation. Topics include digital preservation as an ongoing activity, a set of agreed outcomes, an understood responsibility, a selection process, an economically sustainable activity, a cooperative effort, an innocuous activity, an aggregated or disaggregated service, a complement to other library services, a well-understood process, an arm's length transaction, one of many options, and as a public good.

New York State Archives. *Archival Needs Assessment Guide and Template*. March 2001.
http://www.archives.nysed.gov/a/records/mr_pub59_accessible.html

> Background information on what a needs assessment is for and what is addressed during a needs assessment. Resources for more information can be found under each topic. The template helps to organize your own needs assessment.

Lawrence, Gregory W., William R. Kehoe, Oya Y. Rieger, William H. Walters, and Anne R. Kenney. *Risk Management of Digital Information: A File Format Investigation*. Washington, D.C.: Council on Library and Information Resources. June 2000.
http://www.clir.org/pubs/reports/pub93/pub93.pdf

> This publication offers detailed guidance on migration (which is defined to include conversion) as a preservation technique through a risk assessment process. A useful

workbook is provided to assist users in applying quantitative risk assessment measurements to their own environment.

World Wide Web Consortium (W3C). Extensible Markup Language (XML).
http://www.w3.org/XML/

The W3C is the international body responsible for the development and ongoing refinement of the XML family of standards. This site provides links to the specification itself, as well as pointers to working groups and other resources.

Additional Resources

Library of Congress. *Digital Preservation*.
http://www.digitalpreservation.gov/

This is the current home page for the Library of Congress's National Digital Stewardship Alliance, Digital Preservation Outreach and Education program, and the National Digital Information Infrastructure and Preservation Program. It includes links to publications, tools and services, program initiatives, and other related resources that are updated as individual projects develop.

Northeast Document Conservation Center (NEDCC). *Digital Preservation Readiness Webliography*. Preservation Leaflets. 2007.
http://www.nedcc.org/resources/leaflets/6Reformatting/08DigitalPreservationReadiness.php

A list of resources on digital preservation covering general preservation topics, cost and business models, selection decisions, metadata, standards, copyright, and sample plans and policies.

Minnesota Historical Society, Minnesota State Archives. *Preserving and Disposing of Government Records.* St. Paul: Minnesota Historical Society, May 2008.
http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

Developed for Minnesota government agencies, this overview of the basic principles of records management includes chapters on defining a government record, taking inventory of your records, developing records retention schedules, preserving archival records, disposing of records, and setting up a records storage area. A list of resources for more information is included, as well as information about applicable state law regarding electronic records management. Originally published by the Minnesota Department of Administration in July 2000, the guide was updated jointly by the Minnesota Historical Society and the Minnesota Government Records and Information Network (MNGRIN) in 2008.

Business Case for Digital Preservation

Summary

As more and more records with long-term value are 'born digital' the need for digital preservation business cases has multiplied. A business case is 'a package of information, analysis, and recommendations"[23] that covers a particular issue. Business cases must be tailored to meet specific needs and address concerns of all stakeholders. A handful of state governments have put together a business case for digital preservation that can be used as a model. Reviewing these business cases can be helpful; however variables will undoubtedly require you to modify the business case to address your particular situation and needs.

This document does not try to build a business case for you, but introduces the main elements of a strong business case, both general and specific issues to consider, and background information on what you may need to think about when developing your own. This is followed by a resource list that includes links to sample business cases. Together these pieces provide you with a starting point and a path to follow when developing your own business case.

Legal Framework

For more information on the legal framework you must consider when developing a business case, refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[24].

Key Concepts

As you develop your business case, you will need to be familiar with:

- Essential Elements of a Business Case

Essential Elements of a Business Case

[23] Dawes, Sharon S. et al. *Making Smart IT Choices: Understanding Value and Risk in Government IT Investments.* Center for Technology in Government. April 2004. http://www.ctg.albany.edu/publications/guides/smartit2/smartit2.pdf

[24] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

A business case provides a roadmap to understanding a current situation and where you want to go. The Center for Technology in Government lists the following items as essential elements of a strong business case[25]:

1. A brief, compelling, service-oriented problem statement.
2. A mission statement or vision of the future that addresses the problem.
3. A description of the specific objectives to be achieved.
4. A description and rationale for your preferred approach.
5. A statement of the benefits that address the concerns of all relevant stakeholders.
6. Measures for gauging improved performance or progress toward each objective.
7. A statement of the likely risks of your initiative and how they will be addressed.
8. A basic plan of work with a timeline and key milestones.
9. A project management plan and names and roles of key managers.
10. Alternatives considered and how they would or would not work.
11. Cost estimates and potential sources of funding.
12. Opposing arguments and your responses to them.

These elements should all be addressed and answered with your own situation in mind. Chapter Three of "Making Smart IT Choices: Understanding Value and Risk in Government IT Investments"[26] by the Center for Technology in Government provides suggestions of data sources for each element as well as additional details, context, and examples for each. Chapter 4 in the same resource goes into more detail on how to present your business case to stakeholders and various audiences.

Key Issues to Consider

Business cases are used to show the value of an initiative. Listed below are some of the things to think about when developing a business case for digital preservation.

> *Analyze, and then prioritize:* When developing a business case, analyze your current environment, wants and needs, and then begin to set priorities. If you set priorities first, you might not be able to see the big picture. Priorities are allowed to change. Set your goals accordingly.
>
> *Appropriate solution to task and environment:* Make sure that the solution you are promoting is appropriate for the task at hand and is a good fit for the current and future environment. You will be able to delve into this when you explore element number four (your chosen solution) and element number ten (alternative solutions).

[25] Dawes, Sharon S. et al. *Making Smart IT Choices: Understanding Value and Risk in Government IT Investments.* Center for Technology in Government. April 2004. http://www.ctg.albany.edu/publications/guides/smartit2/smartit2.pdf

[26] Dawes, Sharon S. et al. *Making Smart IT Choices: Understanding Value and Risk in Government IT Investments.* Center for Technology in Government. April 2004. http://www.ctg.albany.edu/publications/guides/smartit2/smartit2.pdf

Integration into routine: The easier the proposed solution is able to be integrated into current routines, the more likely it will be accepted. Don't make major changes unless you already have buy-in from stakeholders. Think about timeframes. Think about who will be affected both internally and externally. Make sure to emphasize how changes will benefit everyone affected.

Cost control: Often business case proposals include an argument for controlling costs. How will the proposed idea prove to be a cost benefit over time? How will it save money or make the organization more efficient? Remember costs can be measured not only in dollars but in staff time and other resources. It is important to be able to make cost estimates as well as point to other sources of funding if available or necessary.

Use values: Use value is often a driver in digital preservation. As a content provider/creator, you must ensure that what people want is accessible now as well in the future. Often, the more people use something, the more value it has. Use value can make a strong case for long-term digital preservation.

Risk analysis: What are the risks involved in the proposal? How will these be addressed? (elements seven and twelve) In addition to possible risks in moving forward, what are the risks if this proposal is not adopted? What future challenges will be introduced? Who will be affected? How? What will the associated costs be?

Collaboration and partnership: Is there any opportunity to collaborate or partner with agencies or organizations that are in a similar position or share similar goals? Is there anyone that you can learn from? Who can you use as a successful example or model to follow?

Specific issues to consider for digital preservation are listed below. (Many of these might also serve as the catalyst of your digital preservation business case.)

Authentication: State government records are often consulted for official or legal purposes. Any digital preservation solution must also be able to address the authentication of the documents. Being able to prove that a digital record is authentic is essential for not only the public good but to avoid legal troubles.

Disaster recovery/Continuing of Operations Plan (COOP): Consider how digital preservation fits in with the organization's disaster recovery and continuing of operations plans. Are your digital records included in these plans? Will you need to access any digital records immediately after a disaster? Will you be able to?

Preservation equals access and use over time: Preservation ensures that records are available for access and use over a long period of time. Who uses your records, both internally and externally? Who would be affected if records were not preserved? Continued access reinforces the value of records.

Retention/disposition policies: The time period that records should be accessible depends on retention and disposition schedules. Holding on to materials unnecessarily costs money, takes up valuable space, and may cause legal troubles.

Data sharing/interoperability: If looking to design or implement a new system, you need to see how the proposed system will fit in with current systems. Are there any data architecture requirements for your organization or agency? Does a centralized IT structure exist? At what level (state, county, local, or institution)? Is a goal to be able to share data across agencies or organizations? If so, you must have an understanding of all involved systems before you implement anything new. The more a system can interact with other systems, the more value you can get from the data. Sharing data also often increases value.

Public access/protection of non-public info: If sharing or providing 'public' access to data or records, how will you handle the protection of non-public information or data? What legal mandates or statutory requirements affect you? You need to be able to address these. There must be controls in place to restrict access to non-public data while providing access to public data.

Accountability/audit-ability: Making records available often increases accountability and audit-ability. As a government agency, being accountable may be a high priority.

Going green: Moving from paper to digital records is sometime part of a 'going green' initiative.

In addition to exploring the above issues, the article "Digital Archiving From Fragmentation to Collaboration" is recommended as it addresses not only sample digital preservation case studies, but also the technological, social and political issues surrounding digital archiving in general. It provides a good framework for putting many different types of issues in context.

Discussion Questions

- Have we set priorities for preserving and providing access to digital files?
- Is the solution we are proposing appropriate? Can it be integrated into current routines?
- What risks are involved in our proposal? How can we address these?
- Are there opportunities to collaborate with others?
- Have we considered all relevant legal requirements?

Annotated List of Resources

Blue Ribbon Task Force on Sustainable Digital Preservation and Access.
http://brtf.sdsc.edu/

> The final report from the Blue Ribbon Task Force "Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information" published in February 2010 discusses issues surrounding the sustainability of digital preservation. (http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf)

Dawes, Sharon S. et al. *Making Smart IT Choices: Understanding Value and Risk in Government IT Investments.* Center for Technology in Government. April 2004.
http://www.ctg.albany.edu/publications/guides/smartit2/smartit2.pdf

> This document sets the stage with defining and describing the essential elements of a business case. The descriptions are followed by further explanations, and specific examples making it easy to understand and follow along. Practical advice for working with diverse audiences is given as well as stressing the importance of identifying and understanding your various audiences and stakeholders. Focus on chapters three and four of this resource as these are the ones that discuss how to prepare and present a business case.

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

> Based in the United Kingdom, the Coalition has published reports and case studies on long-term digital preservation that address issues business cases cover.

Erpanet. *Business Models related to Digital Preservation.* Amsterdam. September 20-22, 2004.
http://www.erpanet.org/events/2004/amsterdam/Amsterdam_Report.pdf

> Provides information on preservation issues and business models within the public and cultural heritage sectors. Financial issues are discussed as well as relating practical experience to specific business models.

GeoMAPP: Geospatial Multistate Archive and Preservation Partnership. *Geospatial Archiving Business Planning.* 2008-2011.
http://www.geomapp.net/publications_categories.htm

> GeoMAPP is a National Digital Information and Infrastructure Preservation Program Project (NDIIPP) focused on the preservation of geospatial data. State partners have explored and created business cases specifically for GIS data. Currently the project

websites have resources that include a template for completing a systems inventory, a PDF of a PowerPoint about business planning and getting stakeholder buy-in, a poster on building a business plan, and links to Utah's business cases. By the end of December 2011 they will also have some additional tools and resources available.

Business Case Examples

Dollar, Charles. *Digital Preservation Readiness Capability Maturity Model and Digital Preservation Readiness Balanced Scorecard; Digital Preservation Planning Project.* Delaware Division of Libraries and the Delaware Digital Preservation Steering Committee. August 3, 2007.
http://state.lib.de.us/For_Libraries/Planning/Digitization/Digitization%20Deliverable%202_final%20Draft8-10-07_XX.pdf

Discusses the development of a digital preservation framework for the state of Delaware. This paper covers the topics of roles and responsibilities, policy, strategy, collaborative awareness, technological expertise, adoption of neutral open standard formats, storage management, planned media renewal, digital object integrity, digital object security, preservation metadata, and access of digital materials. Cost estimates are also discussed.

Office of Secretary of State, Georgia Archives. *Business Case for the Digital Archives of Georgia (DAG).* August 2007.
http://sos.ga.gov/archives/who_are_we/rims/digital_History/documents/DAG_%20Business_Case_v2%201.pdf

Georgia's business case. Includes sections for an executive summary, business need/problem, project objectives, recommended solution, timeline, challenges, assumptions, benefits, and for risks and critical dependencies.

State of Utah. *Business Plan for Archival Preservation of Geospatial Data Resources*. December 30, 2008.
http://www.geomapp.net/docs/Utah_Business_Plan_Geospatial_%20Archive_2008.pdf

An example of a business case that specifically addresses geospatial data. The business case includes the following sections: an introduction, goals, program benefits, program requirements and costs, organizational approach, and implementation plan.

Utah Department of Administrative Services, Division of State Archives. *Electronic Records Management Business Case.*
http://www.geomapp.net/docs/ut_ERMBusinessCase.pdf

Includes goals, business drivers, challenges and issues, risks, recommendations and next steps with managing state electronic records.

Metadata

Summary

Metadata, usually defined as "data about data," is used to describe an object (digital or otherwise), its relationships with other objects, and how the object has been and should be treated over time. Metadata allows users to locate and evaluate data without each person having to discover it anew with every use. A structured format and a controlled vocabulary, which together allow for a precise and comprehensible description of content, location, and value, are its basic elements.

Anyone who has suffered the exercise in irrelevance offered by an Internet search engine will appreciate the value of precise metadata. Because information in a digital format is only legible through the use of intermediary hardware and software, the role of metadata in information technology is fundamentally important.

Whatever you want to do with the information contained within a record (e.g., protect its confidentiality, present it as evidence, provide citizens access to it, broadcast it, share it, preserve it, destroy it) will be feasible only if you and your users can understand and utilize the metadata associated with it. To use metadata effectively you must understand and apply standards that are appropriate to your needs.

Key Concepts

To understand, create, and use metadata effectively, you will need to know more about:

- Legal needs and statutory mandates
- Metadata categories and functions
- Metadata standards
- Metadata and technology

Legal Needs and Statutory Mandates

As part of a government entity, you need to pay particular attention to metadata in order to help you meet basic legal needs and statutory mandates.

For example, Minnesota’s Records Management Act mandates that government agencies cannot dispose of records without the approval of the state’s Records Disposition Panel. Before approval is granted the Records Disposition Panel must understand what the records are, what their significance is, and the proposed method of disposal. In the records management process, metadata is usually structured as a records retention schedule or an Application for Authority to Dispose of Records form (PR-1).

Similarly, the Minnesota Government Data Practices Act classifies data under nine different categories which specify how, when, or if the public may gain access to government data. You cannot guess what level of access or security to provide just by looking at the data itself. You need some additional information – some metadata – in order to follow the law.

The Official Records Act, the Records Management Act, and the Minnesota Government Data Practices Act (MGDPA) are some of the laws most pertinent to the use of metadata. These laws mandate that government agencies must create and keep records in order to be accountable for their actions and decisions and that these records should be accessible to the public unless categorized as not-public by the state legislature. These laws also help establish the records management process for government records.

The metadata requirements of all of these statutes are encompassed in the state's Recordkeeping Metadata Standard, discussed below. For more information on the specifics of these laws and the legal framework you must consider when dealing with government records refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[27].

Metadata Categories and Functions

Metadata is generally categorized into four or five groupings based on the information the metadata captures, as seen in the chart below.

Descriptive Metadata	Metadata that describes the intellectual content of a resource and used for the indexing, discovery and identification of a digital resource.
Administrative Metadata	Metadata that includes management information about the digital resource, such as ownership and rights management.
Structural Metadata	Metadata that is used to display and navigate digital resources and describes relationships between multiple digital files, such as page order in a digitized book.
Technical Metadata	Metadata that describes the features of the digital file, such as resolution, pixel dimension and hardware. The information is critical for migration and long-term sustainability of the digital resource.

[27] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008**.** http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

Preservation Metadata	Metadata that specifically captures information that helps facilitate management and access to digital files over time. This inherently includes descriptive, administrative, structural, and technical metadata elements that focus on the provenance, authenticity, preservation activity, technical environment, and rights management of an object.

Recording these various types of metadata may support a variety of functions for government agencies, but the primary uses are for:

- Legal and statutory reasons (e.g., to satisfy records management laws and the rules of evidence)

- Technological reasons (e.g., to design and document systems)

- Operational or administrative reasons (e.g., to document decisions and establish accountability)

- Service to citizens, agency staff, and others (e.g., to locate and share information)

In all of these cases, utilizing metadata will be most effective if the metadata uses a structured format with a controlled vocabulary when appropriate. "Structured format" means the metadata is defined in terms of specific, standardized elements or fields, based on document type. For example, metadata elements for a library catalog entry for a book include author, title, subject(s), and location, among other things. Unless all the elements are there, users will not be able to evaluate the metadata; they won't be able to answer the question "Is this the book I want?" This structure could be created in-house by institutional practices or more formal by following a specific metadata standard.

"Controlled vocabulary" means that there is an approved of or standard set of terms that can be used as content for each metadata element. Using a controlled vocabulary ensures consistency across a collection and allows items to be found easier and to be compared easier. For example, using a controlled vocabulary in the metadata element 'subject' in a library catalog entry may restrict entries to the Library of Congress's list of Subject Headings rather than allowing any keyword being entered in that field. Controlled vocabulary assists with record comparison across various collections or objects.

Metadata Standards

To work effectively, metadata has to be precise and comprehensible. The entire community of creators and users has to understand what it means. There are a variety of metadata standards in use around the world, with three principal standards in general use in Minnesota government today. Minnesota's Office of Enterprise Technology[28] recommends the following standards in its enterprise architecture:

- Minnesota Recordkeeping Metadata Standard
- Minnesota Web Metadata Standard
- Minnesota Geographic Metadata Guidelines

Minnesota Recordkeeping Metadata Standard (RKMS)

The *Minnesota Recordkeeping Metadata Standard* [29] is designed to support the accountability of government and the proper use of government records as mandated by law. It is based on Dublin Core, but includes additional metadata elements that help support legal mandates over time. [*Dublin Core*[30] is an official international standard (NISO Standard Z39.85; ISO Standard 15836) with fifteen metadata elements.]

The standard consists of twenty elements, ten of which are mandatory and ten optional. In addition, many of these elements contain a number of sub-elements, some mandatory and some optional. To ensure compatibility across metadata sets, six of the ten mandatory elements have direct counterparts both in the Dublin Core and the geographic metadata standards. Overall, the recordkeeping metadata elements are:

- *Agent*. An agency or organizational unit that is responsible for some action on or usage of a record, or an individual who performs some action on a record, or who uses a record in some way. *(mandatory)*

- *Rights Management*. Legislation, policies, and caveats that govern or restrict access to or use of records. *(mandatory)*

- *Title*. The names given to the record. *(mandatory)*

- *Subject*. The subject matter or topic of a record. *(mandatory)*

- *Description*. An account, in free text prose, of the content and/or purpose of the record. *(optional)*

- *Language*. The language of the content of the record. *(optional)*

- *Relation*. A link between one record item and another, between various aggregations of records, or a link between a record and another information resource. *(optional)*

[28] Minnesota Office of Enterprise Technology. *Policies, Standards, and Guidelines.* 2011. http://mn.gov/oet/policies-and-standards/data-management/#

[29] Minnesota State Archives. *Minnesota Recordkeeping Metadata Standard IRM-20.* 2009.http://www.mnhs.org/preserve/records/metadatastandard.html

[30] Dublin Core Metadata Initiative. *Home Page*. 2012. http://dublincore.org/

- *Coverage*. The jurisdictional, spatial, and/or temporal characteristics of the content of the record. *(optional)*

- *Function.* The general or agency-specific business function(s) and activities that are documented by the record. *(optional)*

- *Date*. The dates and times at which such fundamental recordkeeping actions as of the record's or records series' creation and transaction occur. *(mandatory)*

- *Type*. The recognized form or genre a record takes, which governs its internal structure. *(optional)*

- *Aggregation Level*. The level at which the record(s) is/are being described and controlled or the level of aggregation of the unit of description. *(mandatory)*

- *Format*. The logical form (content medium and data format) and physical form (storage medium and extent) of the record. *(optional)*

- *Record Identifier*. A unique code for the record. *(mandatory)*

- *Management History*. The dates and descriptions of all records management actions performed on a record from its registration into a recordkeeping system until its disposal. *(mandatory)*

- *Use History*. The dates and descriptions of both legal and illegal attempts to access and use a record, from the time of its registration into a recordkeeping system until its disposal. *(optional)*

- *Preservation History*. The dates and descriptions of all actions performed on a record after its registration into a recordkeeping system which ensures that the record remains readable and accessible for as long as it has value to the agency and to the community at large. *(optional)*

- *Location.* The current (physical or system) location of the record or details about where the record usually resides. *(mandatory)*

- *Disposal*. Information about policies and conditions that pertain to or control the authorized disposal of records or information about the current retention schedule and disposal actions to which the record is subject. *(mandatory)*

- *Mandate*. A source of recordkeeping requirements. For example, a piece of legislation, formal directive, policy, standard, guideline, set of procedures, or community expectation which (explicitly or implicitly) imposes a requirement to create, keep, dispose of, or control access to and use of a record. *(optional)*

Minnesota Web Metadata Standard

The Web Metadata Standard also uses the Dublin Core metadata standard. For example, when you use the search engine on Minnesota's North Star site, it is the standard Dublin Core metadata elements used in the Minnesota Web Metadata Standard that helps you find exactly what you're looking for by organizing the contents for easy access and retrieval.

The Web Metadata Standard[31] set includes these elements:

- *Title*. The name of the resource given by the creator or publisher.
- *Creator*. The name of the person who created the resource.
- *Subject*. The topic of the resource.
- *Description*. A short, text description of the resource's contents.
- *Publisher*. The name of the entity that published the resource. Note that the publisher is not the person who posted the resource to the web site, but the entity responsible for the publication of the resource, such as your agency.
- *Contributor*. Someone aside from the creator who made a significant contribution to the resource.
- *Date*. Either the creation date or the publication date. Your agency will need to determine which date to use.
- *Resource Type*. The category the resource belongs to, such as committee minutes, press release, or report.
- *Format*. The file format of the resource. For more information on file formats, refer to the *File Formats* guidelines.
- *Identifier*. A text string or number unique to the resource, such as a URL or other formal name. See the *File Naming* chapter in these guidelines for more information on naming web site files for longevity and ease of use.
- *Source*. Information about the source from which the current resource is derived (e.g., an abstract of a report).
- *Language*. The language used in the resource (e.g., English, Spanish).
- *Relation*. An element that refers to related resources.
- *Coverage*. Either geographic (e.g., Minnesota) or temporal (e.g., the years 2000–2001).

[31] The Web Metadata Standard (http://mn.gov/oet/policies-and-standards/data-management/#) uses Dublin Core elements as described here (http://www.bridges.state.mn.us/dcore.html).

- *Rights Management*. A text statement regarding copyright and use permission.

Minnesota Geographic Metadata Guidelines (MGMG)

The *Minnesota Geographic Metadata Guidelines* (MGMG) provide a common approach for documenting all types of geographic data. They have been designed to be straightforward, intuitive, and complete. The guidelines are based on a standard developed by the Federal Geographic Data Committee in 1993: *The Content Standards for Digital Geospatial Metadata*. In developing the MGMG, the Standards Committee of the Minnesota Governor's Council on Geographic Information created a streamlined implementation of the federal standard, while retaining the essence of its original content. Information about the guidelines is available on their website[32].

The Minnesota Geographic Metadata Guidelines includes a number of metadata elements, arranged in seven sections:

- *Identification Information*
- *Data Quality Information*
- *Spatial Data Organization Information*
- *Spatial Reference Information*
- *Entity and Attribute Information*
- *Distribution Information*
- *Metadata Reference Information*

Metadata and Information Technology

Metadata is useful for the management of information in any storage format, paper or digital, but it is critically important for information in a digital format because the information is only discoverable through the use of intermediary hardware and software. We can open up a book or hold microfilm up to a light to determine what it says; but we can't just look at a CD and say what's on it. We cannot possibly hope to locate, evaluate, or use all the files on a single computer or network, let alone the Internet, without metadata.

Databases often store and provide access to metadata separately from the digital files. Metadata can also be stored with or embedded in a digital file. Most software applications automatically

[32] Minnesota Geospatial Information Office. *Minnesota Geographic Guidelines, version 1.2,* October 7, 1998. http://www.mngeo.state.mn.us/committee/standards/mgmg/metadata.htm

create metadata and associate it with files, generally making the standardization of metadata simpler. One example of automatic and standardized metadata is the header and routing information that accompany an e-mail message. Another is the set of properties created with every Microsoft Word document; certain elements such as the title, author, file size, etc., are automatically created, but other elements can be customized and created manually. However, the more manually entered metadata, the harder it becomes to enforce standards. If a lot of manually entered metadata is necessary, policies and procedures for metadata entry are a necessity. By standardizing the process it will be easier to manage, access, and preserve the files long-term. Normally, some combination of automatically and manually created information is best for precise and practical metadata.

If information technology makes metadata necessary, it's also information technology that makes metadata useful. Useful metadata can inform business rules and software code that transforms it into "executable knowledge." For example, metadata can be used for batch processing of files. A date element is critical to records management, as most record retention schedules are keyed to record date of creation. Metadata in more sophisticated data formats, such as eXtensible Markup Language (XML), allow for extraction, use, and calculation based on specific components of a metadata record.

Key Issues to Consider

Now that you are familiar with some of the basic concepts and types of metadata, you can consider some of the issues that have to be addressed in order to use metadata effectively. The most important are:

- *Audiences:* Most people who rely on metadata are unaware they're using it or even that it exists. Nevertheless, when you create metadata, you have to be aware of the audiences for your information in order to determine the appropriate standards and approaches. To make your decisions, you should know which information resources your audiences use, which questions they ask, and what their level of expertise is.

- *Partnerships*: To increase the value of both metadata and the information it describes, you need to work with other creators, custodians, and users of information. If you agree on metadata standards, tools, and practices in collaboration with others, you will create a much more beneficial information management program for your whole organization.

- *Implementation*: Selecting a standard is a good first step. Putting it into practice is a more useful and difficult one. Creating and maintaining metadata over time will demand attention, resources, and staff. You will get a good return on that investment if you keep in mind your legal mandates, your business processes, and your customers as you choose what standards and practices are most appropriate for you.

- *Education*: One critical element of a practical metadata program to keep in mind is education. You will need to know about what others are doing with the standards, the tools, and the uses of metadata. Over time these may change and you will need to keep up with recent developments.

- *Promotion*: To promote the understanding, use, and creation of metadata, as well as to ensure that there are enough resources to support a metadata program, it is important to draw people's attention to metadata and its importance.

Discussion Questions

- What metadata do we need to manage the collection?
- Who is the intended audience or user group for our collection? What do they expect or need?
- Are there any institutional, technological or legal demands we need to consider? What are our legal needs? Does our agency have a records management or data practices office?
- What business functions is the metadata supposed to fulfill?
- What metadata already exists?
- Does our agency have an information and/or technical architecture? What metadata standards does it recommend?
- Are our software applications creating metadata?
- What are the metadata standards pertinent to our profession or business functions?
- Are the offices or departments of our agency already creating metadata? Are they using different standards?
- How much time/money can be spent on creating metadata?
- Do the managers and resource allocators in our agency support a metadata program? Have we made a business case to them?

Annotated List of Resources

Minnesota Recordkeeping Metadata Standard (Minnesota Office of Enterprise Technology Standard IRM 20)
http://www.mnhs.org/preserve/records/metadatastandard.html

The Minnesota Recordkeeping Metadata Standard was developed to facilitate records management by government entities at any level of government. It shares many of its elements with other metadata standards, such as the Dublin Core and the Minnesota Geographic Metadata Guidelines set, but goes further to address such issues as access restrictions, data practices, and records retention and disposition, thereby enabling the practical implementation of statutory mandates for records management. The standard is comprised of twenty elements, ten of which are mandatory.

Dublin Core Metadata Initiative
http://dublincore.org/

The Dublin Core Metadata Initiative is the official site for the Dublin Core (DC) project. The fifteen-element metadata standard is the product of a number of workshops that began in 1995 and is now an official international standard (NISO Standard Z39.85; ISO Standard 15836). Intended to serve users in a flexible manner, the elements are all optional, repeatable, and labeled with descriptive names. Metadata generated from this scheme may be represented in a number of ways (e.g., HTML, RDF) for use on the Internet.

Minnesota Department of Natural Resources. *Best Practice Guidelines for Web Metadata*.
http://www.bridges.state.mn.us/bestprac/index.html

Multiple documents and downloads are available on this web site, including guidelines on how to use the Dublin Core Metadata Element Set as part of the process of archiving web content and a description of each element's purpose and method of creation. The site also offers a bibliography, a training manual on applying the Dublin Core metadata set, and background reports.

Minnesota Geospatial Information Office (MnGeo)
http://www.mngeo.state.mn.us/

MnGeo "coordinates that development, implementation, support, and use of geospatial technology." The MnGeo website offers access to a variety of resources, including standards and the Minnesota Geographic Data Clearinghouse.

Federal Geographic Data Committee (FGDC). *Metadata.*
http://www.fgdc.gov/metadata

> This site is sponsored by the FGDC, which is made up of several federal agencies. Working with such partners as state and local governments, the academic community, and industry, the FGDC is supervising the development of the National Spatial Data Infrastructure (NSDI) with the goal of sharing geographic data through standards, policies, and procedures. Through subcommittees and working groups, the FGDC has several geospatial data standards completed or in some stage of development. These include the Cadastral Data Content Standard, the Spatial Data Transfer Standard, the Spatial Data Accuracy Standard, the Address Content Standard, and the Government Unit Boundary Data Content Standard.
>
> The FGDC has developed the Content Standard for Digital Geospatial Metadata (CSDGM) (http://www.fgdc.gov/metadata/csdgm) to be used by all federal agencies. This metadata standard is composed of 334 different elements (119 of which only contain sub-elements). The FGDC also coordinates the National Geospatial Data Clearinghouse for participants worldwide interested in sharing digital geospatial data that conforms to the CSDGM. In the future, the CSDGM is expected to be modified to be made compliant with an emerging international metadata standard, ISO 19115.

Minnesota Department of Administration, Office of Enterprise Technology. *Minnesota Enterprise Technical Architecture.* Version 2.02, 2006.
http://mn.gov/oet/

> The Minnesota Office of Enterprise Technology (OET) is charged with establishing and maintaining a state information architecture as specified in Minnesota Statue, Chapter 16E.04 Subdivision 2. According to the OET, “This technical architecture is established to describe technology components of the State’s information infrastructure and their individual principles, practices and standards that are to be used to guide the development and delivery of all information systems services. The architecture will provide a reference so that various groups of government IT professionals have a consistent view of the information systems infrastructure and the methods that they employ to develop and deliver information systems services.” Chapter 4, Data and Records Management, describes the framework for managing information resources, as well as the standards and guidelines that apply.

File Naming

Summary

A file name is the chief identifier for a record. In the world of electronic records, the record's file name provides metadata that places the record in context with other records, records series, and records retention schedules. In most organizations, the policy for naming a file (and hence a record) is left to individuals or to groups of individuals (e.g., departments, committees). Consider establishing an agency-wide file naming policy that complements your electronic records management strategy.

Consistently named records foster collaboration based on mutual understanding of how to name files and use file names (including the file name metadata). Consistently named records also help you to meet your legal requirements. Legally, your records must be trustworthy, complete, accessible, legally admissible in court, and durable for as long as your approved records retention schedules require. Records that are consistently and logically named are easier to manage to meet these requirements.

In other words, with each staff member consistently naming electronic records, other staff members will be able to look at a record's file name and use that information to recognize the contents and characteristics of the record and to make decisions about it. For example, a staff member could see that "HF0035broch96/97P.pdf" is a brochure about a House bill (HF0035) in the 1996/1997 session that is available to the public.

Legal Framework

For more information on the legal framework you must consider when developing a file naming policy, refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[33].

Key Concepts

As you develop your file naming policy, you will need to be familiar with the following:

- Differences among file names, file paths, and addresses
- Common file name elements
- Planning for a File Name Policy

Differences Among File Names, File Paths, and Addresses

[33] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

A *file name* is the name of the file as it stands alone. The *file path* shows the location of the file. For example, the file named "CommitteeAMinutes021401.doc" might be stored in a series of nested directories (its file path) for all committees as: "X:Committees/CommitteeA/Minutes/2001/February/CommitteeAMinutes021401.doc." An *address* describes the location of a file delivered on the Internet. For example, a map of a public park named Smith Park might have the following address: "http://www.parks.org/smith.html."

Common File Name Elements

When developing your file naming policy, you may wish to include some of the following common elements in the file name:

- Version number (e.g., version 1 [v1, vers1])
- Date of creation (e.g., February 24, 2001 [022401, 02_24_01])
- Name of creator (e.g., Rupert B. Smith [RBSmith, RBS])
- Description of content (e.g., media kit [medkit, mk])
- Name of intended audience (e.g., general public [pub])
- Name of group associated with the record (e.g., Committee ABC [CommABC])
- Release date (e.g., released on June 11, 2001 at 8:00 a.m. central time [61101_0800CT])
- Publication date (e.g., published on December 24, 2003 [pub122403])
- Project number (e.g., project number 739 [PN739])
- Department number (e.g., Department 140 [Dept140])
- Records series (e.g., SeriesX)

Planning for a File Name Policy

Having a file name policy in place will assist with managing your electronic records. When creating a file name policy the following suggestions[34] should be considered:

[34] Information about file naming conventions gathered from the following resources: Alberta Government. *Naming Conventions for Electronic Records*. August 2005. http://www.im.gov.ab.ca/publications/pdf/DocumentNamingConventions.pdf ; Bibliographic Center for Research's (BCR) *BCR's CDP [Collaborative Digitization Program] Digital Imaging Best Practices Version 2.0* June 2008. http://mwdl.org/docs/digital-imaging-bp_2.0.pdf ; Brookhaven National Laboratory (BNL) Web Communication

- Create unique file names. Duplicate file names will cause problems.

- File names should be simple and easy to understand.

- Use only alpha-numeric characters. Avoid using special characters such as: ? / $ % & ^ # . \ : < >. Special characters are often reserved for use by the operating system.

- Use underscores (_) and dashes (-) to represent spaces. Spaces are often reserved for operating system functions and might be misread.

- Use leading zeros with the numbers 0-9 to facilitate proper sorting and file management.

- Dates should follow the ISO 8601 standard of YYYY_MM_DD or YYYYMMDD[35]. Variations include YYYY, YYYY-MM, YYYY-YYYY. This maintains chronological order. If dates of creation are used, these can make following retention schedules easier.

- Keep the file name as short as possible and always include the three character file extension preceded with a period (Ex: .jpg or .doc).

- Include the version number in the file name by using 'v' or 'V' and the version number at the end of the document. (Ex: 2004_Notes_v01.doc) Avoid using the word version or draft and the beginning of the file name for access purposes (Ex: Version1_2004_Notes.doc).

- Order the pieces of information or elements being used to create the file name in the most logical order based on retrieval methods. For example, use the date first on events that are time specific or reoccurring, and use the name of the event for events that are infrequent and will be easier to find by name rather than date.

As you develop your policy, you will also need to address the following:

- *Persistence over time*. File names should outlast the records creator who originally named the file. With good stakeholder and staff input, and training, you should be able to develop file names that make sense to staff members once the file creators are no longer available.

Standards. *File Naming Conventions and Directory Structure*. February 5, 2008. http://www.bnl.gov/webstandards/fileNaming.asp; Digital Projects Advisory Group, University Libraries at the University of Colorado at Boulder. *File Naming Conventions for Digital Collections*. March 4, 2008. http://ucblibraries.colorado.edu/systems/digitalinitiatives/docs/filenameguidelines.pdf ; JISC Digital Media. *Choosing a File Name*. November 2008. http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/choosing-a-file-name/; North Carolina Department of Cultural Resources. *Best Practices for File-Naming*. May 7, 2008. http://www.records.ncdcr.gov/erecords/filenaming_20080508_final.pdf.

[35] North Carolina Department of Cultural Resources. *Best Practices for File-Naming.* May 7, 2008. http://www.records.ncdcr.gov/erecords/filenaming_20080508_final.pdf

- *Access and ease of use*. The policy should be simple and straightforward. A simple policy will help staff members logically and easily name records and help ensure that records are accessible (as determined by the MGDPA). A simple policy will be more consistently used, resulting in records that are consistently named, and thus easier to organize and access.

- *Ease of administration*. The policy should work with your computer infrastructure, so that you can monitor policy compliance, manage records and records series, gather metadata, and perform other administrative tasks easily and in compliance with all legal requirements. For example, if all the records in a specific records series are easily identifiable by file name, they will be easier to gather and manage.

- *Scalability*. Consider how scalable your file naming policy needs to be. For example, if you want to include the project number, don't limit your project numbers to two digits, or you can only have ninety-nine projects.

- *Determining what metadata to collect*. You will need to decide what metadata to collect and include in file names. This will help ensure the long-term usefulness of your records and help you to meet legal requirements for accessibility (for public records) and accountability, as well as protect not-public records.

- *Universal retrieval*. Ensure that the staff and the public (as appropriate) can access your files. Legally, public records must be accessible. Standard file names allow users to find records efficiently.

- *Determining the official copy*. Determine which file is the "official" copy. As part of your web content management (see the *Web Content Management* chapter of these guidelines), you should include in your policy which web site files are official records, and which version of the electronic file is the official record. Including an indicator of official record status in a file name may be useful for this purpose. The inclusion of this parameter in your policy will help you meet your legal requirements to capture records as set forth by the Official Records Act. The inclusion of this designation may also make administration of your web site records easier.

- *Determining file naming boundaries*. Pay close attention to the freedom you give staff members (and outside vendors) in naming files. Provide guidelines and training on file naming. You will not be able to manage every electronic record's file name, so you will need to rely on staff members and vendors to name files in compliance with your policy. By providing guidelines and training, you can maximize policy compliance in a way that meets your operational and legal requirements.

- *Relationship to and connection with paper records*. Determine how the names of your electronic records relate to the names of paper files you have stored. Because electronic records may be part of records series that include paper records, the file naming policy for electronic records should fit logically with your paper records naming. For example, a letter published on a web site might be part of a records series that includes additional paper documents in a file folder. By ensuring that the electronic records' and the paper records' file names mesh, you can more easily manage the records series.

Key Issues to Consider

Now that you are familiar with some of the basic concepts of file naming, you can use the questions below to discuss how they relate to your agency.

Pay special attention to the questions posed by the legal framework, including the need for public accessibility, as appropriate. Consider your current and future activities and records to help determine the components of a file naming policy that will work now and in the future. For example, you may currently publish official statements or press releases on paper, but in the future, you may publish such records on the web.

Discussion Questions

- Do we have a file naming policy? Does it cover various type of documents in various environments? Is it extensible into the future? Does it make sense to all user groups?

- How will people be accessing the files? How will people “think of” this record (e.g., “I need to find a copy of XYZ.doc.” or “I need information about legislation passed in 2002/2003.”)?

- What information is most important to capture in the file name itself?

- Are there limitations on the length of the file name? By the software, computer system, or storage device?

- Will the records move location (e.g., from one server to another, from a server to long-term storage)? Could these changes affect the file naming strategy?

- How will staff members and the public access and open files in the short-term and long-term? What limitations do these systems have for file naming?

Annotated List of Resources

Cool URIs Don't Change. In: *Style Guide for Online Hypertext*. Cambridge, MA: World Wide Web Consortium (W3C), 1998.
http://www.w3.org/Provider/Style/URI

> This section of the complete style guide discusses the file naming concepts for the World Wide Web to ensure the accuracy of links and the longevity of the names.

Digital Preservation Education for North Carolina State Government Employees. *Digital Preservation Tutorial: File Naming.* December 2011.
http://digitalpreservation.ncdcr.gov/tutorials.html

> A four-part video tutorial describing why file naming is important, how to change a file name, what not to do when changing a file name, and best practices for file naming. These videos show how deliberate file naming can lead to responsible file management and ongoing digital preservation.

Naming and Addressing: URIs, URLs,….
http://www.w3.org/Addressing

> These web pages describe the relationship of URIs, URLs, and URNs. The pages also provide links and other information about other file naming topics for the web, such as metadata, markup languages, events, and history.

PURL
http://purl.oclc.org/docs/index.html

> The OCLC PURL Service provides a comprehensive introduction to the subject of PURLs. Available from this web site are Frequently Asked Questions on PURLs, introductions to the subject, and the opportunity to create and modify a PURL.

Wikipedia. File Naming. Last update 29 February 2012.
http://en.wikipedia.org/wiki/Filename

> An article that describes ‘filename’ in association with computer systems. More information is also provided about reserved characters and words within file names as well as a comparison of file name limitations for various systems.

File Formats

Summary

Rapid changes in technology mean that file formats can become obsolete quickly and cause problems for your records management strategy. A long-term view and careful planning can overcome this risk and ensure that you can meet your legal and operational requirements.

Legally, your records must be trustworthy, complete, accessible, legally admissible in court, and durable for as long as your approved records retention schedules require. For example, you can convert a record to another, more durable format (e.g., from a nearly obsolete software program to a text file) and that copy, as long as it is created in a trustworthy manner, is legally acceptable.

The software in which a file is created usually uses a default format when the file is saved. This is indicated by the file name suffix (e.g., .PDF for portable document format). However, most software allows authors to select from a variety of formats when they save a file. For example, Microsoft Word allows the author to select document [DOC], Rich Text Format [RTF], or text [TXT], as well as other format options. Some software, such as Adobe Acrobat, is designed to convert files from one format to another. The format you choose will affect your long-term records management abilities.

Legal Framework

For more information on the legal framework you must consider when developing a file format policy refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[36].

Key Concepts

As you consider the file format options available to you, you will need to be familiar with the following concepts:

- Proprietary, Non-Proprietary, Open-Source, and Open Standard File Formats
- File Types and their Associated Formats
- Preservation Options: Conversion and Migration
- Compression
- Importance of Planning

[36] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

Proprietary, Non-proprietary, Open Source, and Open Standard File Formats

- *Proprietary formats*. Proprietary file formats are controlled and supported by just one software developer. Microsoft Word (.DOC) format is on example.

- *Non-proprietary formats*. These formats are supported by more than one developer and can be accessed with different software systems. eXtensible Markup Language (XML) is a popular non-proprietary format for government records.

- *Open Source formats*. In general, open source refers to any program whose source code is made available for use or modification as users or other developers see fit. Open source formats are published publicly available specifications for storing data which are often maintained by a standards organization. Open formats can be used by proprietary and open source software alike.

- *Open Standard formats*. Open standard software formats are created using publicly available specifications (open source formats). Although software source codes remain proprietary, the availability of the standard increases compatibility by allowing other developers to create hardware and software solutions that interact with, or substitute for, other software.

When choosing a file format to use for your electronic records management purposes, it is important to understand how proprietary, non-proprietary, open formats, and open standards may affect the accessibility and accountability of your records over the long term.

File Types and their Associated Formats

The following are brief descriptions of the basic file types and formats you are likely to encounter. Additional information can be found in the *Digital Imaging* chapter of these guidelines while resources in the Annotated List of Resources provide more detailed information on individual file formats.

- ***Text files***. Text files are most often created in word processing software programs. Common file formats for text files include:

  - *Proprietary formats*, such as Microsoft Word files, which carry the extension of the software in which they were created.

  - *RTF or Rich Text Format files*, are supported by a variety of applications and saved with formatting instructions (such as page layout).

  - *Portable Document Format (PDF) files*, which contain an image of the page, including text and graphics. PDF files are widely used for read-only file sharing. Adobe Acrobat is, by far, the most popular PDF file application, although others are available.

  - *Portable Document Format* (PDF/A) files. PDF/A, as standard file format for long-term archiving of electronic documents, is a subset of PDF. Files are 100% self-contained, and do not rely on outside sources for document information. ISO standard: ISO 19005-1:2005.

- ***Graphics files***. Graphics files store an image (e.g., photograph, drawing) and are divided into two basic types; vector-based and raster-based.

  Vector-based files store an image as mathematical formulas. Vector image programs use this mathematical formula to display and scale the image without distortion. Common types of vector-based file formats include

  - Drawing Interchange Format (DXF) files, which are widely used in computer-aided design software programs, such as those used by engineers and architects.
  - Encapsulated PostScript (EPS) files, which are widely used in desktop publishing software programs.
  - Computer Graphics Metafile (CGM) files, which are widely used in many image-oriented software programs (e.g., Photoshop) and offer a high degree of durability.
  - Shapefiles (SHP), ESRI GIS applications use vector coordinates to store non-topological geometry and attribute information for features.

  Raster-based files store the image as a collection of pixels. Raster graphics are also referred to as bitmapped images. Raster graphics cannot be scaled without distortion. Common types of raster-based file formats include:

  - Bitmap (BMP) files are relatively low-quality files used most often in word processing applications. Uncompressed.
  - Tagged Image File Format (TIFF) files are usable with many different software programs and are often the format of choice for a high-quality master image. Uncompressed or lossless compression.
  - Joint Photographic Experts Group (JPEG) files are very common format for digital photography. JPEGs are also the preferred format for Internet delivery and file sharing of photographs. Lossy compression.
  - Joint Photographic Experts Group (JPEG2000). An evolving format with multiple compression techniques based on wavelet technology. Lossless compression.
  - Graphics Interchange Format (GIF) files were widely used on the Internet for graphics and logos with areas of solid color. Due to color limitations, photographs are not accurately represented with this format. GIF can also be used for low

resolution animations. (PNG has improved on the color limitations of GIF.) Lossless compression.

- Portable Network Graphic (PNG) files, designed to replace GIF, are patent and license free and produce higher quality files than GIF. PNG format is preferred for images that contain text or line art, especially on the Internet. Lossless compression.

- ***Data files***. Data files are created in database software programs and are therefore often represented proprietary formats. Data files are divided into fields and tables that contain discrete elements of information. The software builds the relationships between these discrete elements. For example, a customer service database may contain customer name, address, and billing history fields. These fields may be organized into separate tables (e.g., one table for all customer name fields). You may convert data files to a text format, but you will lose the relationships among the fields and tables. For example, if you convert the information in the customer database to text, you may end up with ten pages of names, ten pages of addresses, and a thousand pages of billing information, with no indication of which information is related.

- ***Spreadsheet files***. Spreadsheet files store the value of the numbers in their cells, as well as the relationships of those numbers. For example, one cell may contain the formula that sums two other cells. Like data files, spreadsheet files are most often in the proprietary format of the software program in which they were created. Data can be shared between different spreadsheet programs by saving individual spreadsheets as a text file in the Data Interchange Format (DIF), however the value and relationship of the numbers may be lost.

- ***Video and audio files***. These files contain moving images (e.g., digitized video, animation) and sound data. They are most often created and viewed in proprietary software programs and stored in proprietary formats. Common files formats in use include QuickTime (.MOV), Windows Media Video (.WMV), and Motion Picture Experts Group (MPEG) formats (.MP3); others include .AVI and .WAV files.

- ***Markup languages***. Markup languages, also called *markup formats*, contain embedded instructions for displaying or understanding the content of the file. They provide the means to transmit and share information over the web. The following markup language file formats are supported by the World Wide Web Consortium (W3C)[37] as standards

  - *Standard Generalized Markup Language (SGML)*, a common markup language used in government offices worldwide, is an international standard. HTML and XML are derived from SGML.

  - *Hypertext Markup Language (HTML)* is used to display most of the information on the World Wide Web. Because presentation is combined with content trough the use of pre-defined tags, HTML is simple to use but limited in scope. Other markup languages such as XHTML and XML offer greater flexibility.

[37] World Wide Web Consortium (W3C). *Home Page.* 2012. http://www.w3.org/

- *eXtensible Markup Language (XML)* is a relatively simple language based on SGML that is gaining popularity for managing and sharing information. XML provides even greater flexibility and control than XHTML while avoiding the complexities associated with SGML.

- *eXtensible Hypertext Markup Language (XHTML)* combines the flexibility found in XML with the ease of use associated with HTML. Strict XHTML rules improve consistency and provide the ability to create your own markup tags. Because they share similar rules, converting XHTML into XML is easier that converting HTML into XML.

The table below summarizes the most common file formats.

File Format Type	Common Formats	Example Applications	Description
Text	PDF, RTF, TXT, DOC	Letters, reports, memos, e-mail messages saved as text	Created or saved as text (may include graphics)
Vector Graphics	DXF, EPS, CGM, SHP	Architectural plans, complex illustrations, GIS	Store the image as geometric shapes in a mathematical formula for undistorted scaling
Raster Graphics	TIFF, BMP, GIF, JPEG, PNG	Web page graphics, simple illustrations, photographs	Store the image as a collection of pixels which cannot be scaled without distortion
Data File	Proprietary to software program	Human resources files, mailing lists	Created in database software programs
Spreadsheet File	Proprietary to software program, DIF	Financial analyses, statistical calculations	Store numerical values and calculations
Video and Audio Files	QuickTime (MOV), MPEG, Real Networks (RM), WMV, WAV, MP3, AVI	Short video to be shown on a web site.	Contain moving images and sound
Markup Languages	SGML, HTML, XHTML, XML	Text and graphics to be displayed on a web site	Contain embedded instructions for displaying and understanding the content of a file or multiple files

Preservation Options: Conversion and Migration

To help ensure your files are accessible over time, you will need to keep verifying that the files formats you are using are still supported. When formats are no longer supported, you will need to decide if you are going to convert and/or migrate your file formats. If you convert your records, you will change their formats, perhaps to a software-independent format. If you migrate your records, you will move them to another platform or storage medium, without changing the file format. However, you may need to convert records in order to migrate them to ensure that they remain accessible. For example, if you migrate records from an Apple operating system to a Microsoft Windows operating system, you may need to convert the records to a file format that is accessible in a Windows operating system (e.g., RTF, Word 2000). For more information on conversion and migration, refer to the *Electronic Records Management Strategy* and *Long-Term Preservation* chapters of these guidelines.

You will be faced three basic types of loss when converting or migrating files that will need to be considered before finalizing your plan. The amount and type of loss needs to be analyzed to determine the best course of action. The three types of loss are:

- *Data*. If you lose data or if it becomes corrupted, you lose, to a varying degree, the content of the record. Bear in mind that, legally, your records must be complete and trustworthy. Metadata may also be altered or lost.

- *Appearance*. If you convert all word processing documents to RTF, you risk loss of the structure of the record; you may lose some of the page layout. You must determine if this loss affects the completeness of the record. If the structure is essential to understanding the record, this loss may be unacceptable.

- *Relationships*. Another risk is the loss of the relationships of the data within the file or between files (e.g., spreadsheet cell formulas, database file fields). Again, this loss may affect the legal requirement for complete records.

Compression

As part of your records management strategy, you may choose to compress your files. A few of the pros and cons are summarized below.

- *Pros*
  - Saves storage space
  - Files are more quickly and easily transmittable
- *Cons*
  - May result in data loss
  - Introduces an additional layer of software dependency (the compression software)

Compressing files results in a smaller file size, which reduces the amount of storage space needed. However, to create a smaller file size, information is often removed from the file. For example, when an image file is compressed, pixels that the software determines will not be missed are removed, relying on the human eye to fill in the absent details. When an audio file is compressed, sounds often unnoticeable to the human ear are removed, resulting in a smaller file size that, to most people, sounds the same as the uncompressed file. Compression options vary in their degree of data loss. Some are intentionally "lossy," such as the ones described above while others are designed to be "lossless." Lossless compression results in a smaller file size, but allows for exact reconstruction of the original file from the compressed data, unlike lossy compression which only approximates the original data. Because of these issues, you may choose to compress some files and not others.

Importance of Planning

Many of the challenges associated with records management can be overcome with good planning. When trying to determine the most appropriate file format/s to use for long-term access, there are many things to consider. Weighing the pros and cons of each of the suggestions below will assist with planning efforts.

- *Accessibility*. The file format must enable staff members and the public (as appropriate under the MGDPA) to find and view the record. In other words, you cannot convert the record to a format that is highly compressed and easy to store, but inaccessible.

- *Longevity*. Developers should support the file format long-term. If the file format will not be supported long-term, you risk having records that are not durable, because the software to read or modify the file may be not be available. Records should be migrated or converted if you determine a file format is no longer supported. Open source, open standard and non-proprietary formats are preferable to completely proprietary ones.

- *Accuracy*. If you convert your records, the file format you convert to should result in records that have an acceptable level of data, appearance, and relationship loss, if any.

- *Completeness*. If you convert your records, the file format you convert to should meet your operational and legal objectives for acceptable degree of data, appearance, and relationship loss, if any.

- *Flexibility*. The file format needs to meet your objectives for sharing and using records. For example, you may need to frequently share copies of the records with another agency, use the records in your daily work, or convert and/or migrate the records later. If the file format can only be read by specialized hardware and/or software, your ability to share, use, and manipulate the records is limited.

Key Issues to Consider

Now that you are familiar with some of the basic concepts of file formats, you can use the questions below to discuss how those concepts relate to your agency. Pay special attention to the questions posed by the legal framework, including the need for public accessibility as appropriate, completeness, trustworthiness, durability, and legal admissibility. Consider the degree of acceptable data, appearance, and relationship loss. Take a long-term approach so that your file formats will meet your operational and legal requirements now and in the future.

Discussion Questions

- What are our goals for electronic records management?
- How is our agency affected by the legal requirements?
- What current file formats do we use? Is it anticipated that these will be supported long-term?
- Are we planning on converting and/or migrating our records? If so, when? How often?
- How will we find and document loss and/or changes?
- What levels of data, appearance, and relationship loss are acceptable?
- How will our decisions affect other groups that may need current and future access to our records (e.g., other government agencies, the public)?

Annotated List of Resources

The National Archives (NARA). Frequently Asked Questions (FAQs) About Selecting Sustainable Formats for Electronic Records.
http://www.archives.gov/records-mgmt/initiatives/sustainable-faq.html

> Answers to frequently asked questions about sustainable formats including the characteristics of a sustainable format, the importance of using a sustainable format, and contacts for more information.

Library of Congress. Sustainability of Digital Formats.
http://www.digitalpreservation.gov/formats/

> Defines file format for still images, sound, text, moving images and web archives. Discusses what factors make a format sustainable and how to evaluate file formats based on a set of goals.

Lawrence, G.W., W.R. Kehoe, O.Y. Rieger, et al. *Risk Management of Digital Information: A File Format Investigation*. Washington, D.C.: Council on Library and Information Resources, 2000.
http://www.clir.org/pubs/abstract/pub93abst.html

> This publication provides an overview of file format issues related to records management strategies. The publication also provides a comprehensive workbook for users to help them develop a records management strategy.

Clausen, Lars R. *Handling File Formats*. Denmark: The State and University Library, The Royal Library, May 2004.
http://netarchive.dk/publikationer/FileFormats-2004.pdf

> This report is a publication of the Netarchive.dk project, which seeks strategies for archiving the Danish part of the World Wide Web. The report offers a succinct and intelligent analysis of the issues surrounding file format preservation, including the categorization of formats, aspects of preservation quality, assessment criteria for future usability, and preservation strategies.

JISC Digital Media; Still Images, Moving Images and Sound Advice.
http://www.jiscdigitalmedia.ac.uk

> “JISC Digital Media is hosted at the Institute for Learning and Research Technology (ILRT) at the University of Bristol” and provides information about electronic media. The website features technical and project management advice on still images, moving

images and sound media including explanations of file formats. The site has also compiled a glossary of terms.

ANSI/ARMA Standard 16-2007: The Digital Records Conversion Process: Program Planning, Requirements, Procedures
http://www.arma.org/standards/DigitalConversion.cfm

"This standard provides information on keeping electronic records authentic during conversion, planning programs for records conversion, and understanding recordkeeping requirements during conversion… Part I of the standard addresses decisions relating to program planning and recordkeeping issues. Part II discusses the actual conversion process."

PRONOM: The File Format Registry.
http://www.nationalarchives.gov.uk/PRONOM/Default.aspx

PRONOM is maintained by the Digital Preservation Department of the UK National Archives. Visitors to the site can search within five areas (File Format, Product, Vendor, Support Period, and Release Date), each of which offer more options. Choosing "File Format," for instance, allows visitors to search just by extension to get a straightforward list of associated software or by compatible products, which returns a list of products, versions, release dates, vendors, read/write capabilities, and invariance. Links on vendor and product lead to a wealth of additional detail. Reports can be easily printed or exported into XML or CSV (Comma Separated Value file) for further use.

Wotsit's Format: The Programmer's Resource.
http://www.wotsit.org

This online catalog of file formats is broken down into categories such as "Graphics Files," "Text Files/Documents," and "Spreadsheet/Database." Visitors can browse each section or can use the provided search engine to zero in on their mark. Each format carries a one-line description and a link to further information either online or in a download file.

World Wide Web Consortium (W3C)
http://www.w3.org

W3C is a consortium of organizations around the world that develops and promotes common web protocols. The site contains news, specifications, guidelines, software, and tools for web development on a wide variety of topics, including markup languages and transfer protocols.

Digital Media

Summary

On-going and rapid advances in technology dictate that you store your electronic records on media that enable you to meet your long-term operational and legal requirements. Legally, your records must be trustworthy, complete, accessible, legally admissible in court, and durable for as long as you need them. Because every digital storage option will eventually become obsolete, consider digital storage options that will enable you to maintain records by migrating and/or converting them during their required retention period.

Legal Framework

For more information on the legal framework you must consider when selecting digital storage media refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[38].

Key Concepts

Before you determine which digital media will meet your long-term legal and operational needs, familiarize yourself with the following key concepts:

- Digital Media
- Magnetic Media
- Optical Media
- Solid State Media
- Digital Media Capacity
- Media Life Expectancy
- Care and Handling of Digital Media
- Storage Options
- Performance Issues to Consider

Digital Media

[38] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

Digital data is stored on digital media. Digital media can be divided into three main types:

- *Magnetic*. On magnetic media, the digital data is encoded as microscopic magnetized needles on the surface of the medium (e.g., tape).

- *Optical*. On optical media, the digital data is encoded by creating microscopic holes in the surface of the medium (e.g., CD or DVD).

- *Solid State*. Containing no moving parts, solid state media encode digital data by applying small voltages of electricity that temporarily induced a group of transistors either on or off. (e.g., flash memory cards, USB removable media).

Based on the characteristics of the different types of media, access to the information is divided into two categories, sequential and random access.

- *Sequential*. Sequential access requires the user to access specific information by accessing the preceding information on the medium. For example, if you want to view a specific portion of magnetic tape, you must first fast-forward through the preceding portion of the tape.

- *Random*. Some digital media allow users to access the stored information from any physical place on the media. For example, when you connect a flash drive to your computer or insert a DVD, you can access any single file stored on the media without having to first access all the files that precede it.

When choosing a media one must also consider the purpose for storing the data. How long will the records need to be accessible? Who will need to access the files? Are there are legal requirements associated with ensuring the authenticity of the records? If so, Write-Once Read-Many (WORM) technology should be considered. WORM technology originally was an optical media option. It required a special WORM disk drive to enable the user to read or write to specific WORM disks. Today, WORM technology is also available to use with magnetic media, for a lower price with a higher capacity. As long as the machine can recognized WORM tapes, there is no need for a separate tape drive.

Magnetic Media

Magnetic media includes:

- *Magnetic Disk*. Magnetic disks include the hard disk found in your computer that stores the programs and files you work with daily. Magnetic disks provide random access. Also included are:

  - *External Hard Drive.* External hard drives are encased in housing and connected via cable to a computer port.

  - *Network Environment.* Multiple hard drives are connected to each other in a way that shares resources and information, creating a network.

- *Magnetic Tape*. Magnetic tapes come in reel-to-reel, as well as cartridge format (encased in housing for ease of use). The two main advantages of magnetic tapes are their relatively low cost and their large storage capacities. Magnetic tapes provide sequential access to stored information, which is slower than the random access of magnetic disks. Magnetic tapes are a common choice for long-term storage or the transport of large volumes of information. WORM technology is available with many of these tape formats. The only requirement is that the machine is capable of recognizing the special WORM tape. As an example, Linear Tape-Open (LTO) is an open-standard magnetic take system that allows interoperability between tapes and tape drives made by different manufacturers.

Optical Media

Optical media options include:

- *Compact Disk (CD)*. Compact disks come in a variety of formats. These formats include CD-ROMs that are read-only, CD-Rs that you can write to once and are then read-only, and CD-RWs that you can read and write to in multiple sessions. CD-RWs have less life expectancy than non-rewritable disks. CDs are relatively stable and with proper error checking suitable for data storage of five years before refreshing.

- *Standard Definition Digital Versatile Disk (SD-DVD/DVD)*. These disks are also called digital video disks, but do not necessarily include video. DVD disks have more storage capacity than CD-ROMs. DVDs come in various types +/- and may or may not be compatible with each other (see list below). When DVDs players and recorders were first developed they would only play the + or – formats, not both. Today, however, most DVD recorders and players will accept either the +/- format. The life expectancy of DVDs are similar to that of CDs but with the ever changing technology, may not be the most reliable storage medium for long-term files. Common types of DVDs include:

  – *DVD+R and DVD-R*. DVD+R and DVD-Rs can be written to once and then are read only. (4.7 GB per layer) DVD-Rs are more commonly compatible with older machines; DVD+Rs were only officially recognized as an official DVD format in 2008.

  – *DVD-RAM*. These DVDs are rewritable disks with exceptional storage capacity. They come in one- or two-sided formats. Rewritable disks have less life expectancy that non-rewritable ones.

  – *DVD+RW and DVD-RW*. These are direct competitors to DVD-RAM with similar functionality, are rewritable and have slightly greater storage capacity.

- *Write-Once, Read-Many (WORM) disk*. WORM disks require a specific WORM disk drive to enable the user to write or read the disk. WORM disks function the same as CD-R and DVD-R disks.

- *High-Definition DVD.* Envisioned to be the successor of standard definition DVDs. High-definition DVDs have higher storage capacity than a standard definition DVD. A single sided HD-DVD can store 25 GB rather than the 4.7 of a SD-DVD. The optical technology uses a blue ray (rather than a red ray) which has a shorter wavelength which increases the storage capacity of the disks. Blu-Ray and HD-DVD were two competing formats. Support for HD-DVD was discontinued in early 2008.

  – *Blu-ray Disc.* Developed by the Blu-Ray Disc Association, the main uses for this disk are for video, computer games and data storage. BD-R and BD-RE are the read-only and rewriteable formats of the Blu-ray disc.

  – *HD-DVD.* Supported by Toshiba. A competitor of Blu-Ray, that also offered higher storage capacity than previous DVD formats. Support for HD –DVD +/- Rwas discontinued in early 2008, making Blu-ray the format of choice for higher capacity optical storage.

- *Optical cards.* Optical cards, also known as “smart cards,” are the size of a credit card. They come in read-only and read-write formats. They are not in widespread use except for limited applications, such as automatic teller machines, personal identification for security systems, and airline reservations.

- *Optical tape*. Optical tape is tape coated with optical recording material. Optical tape is not widely used.

Solid State Media

Solid state media is used in various removable devices utilizing flash memory including digital cameras, cell phones, computer games, music players, and video recorders. Small cards and “memory sticks” store images, games, music, data, programs, and video. Storage capacities of these cards or sticks are ever increasing; when they were originally introduced their size was between 32 MB and 512 MB; there are now models that can store over a terabyte. Memory cards and flash drives (memory sticks) can be connected to a computer via card reader or USB port to assist with data transfer between devices. The small size, portability, and no moving parts make solid state media attractive. Data is accessed randomly. All formats are re-writable. Long-term storage capabilities of solid state technology are still being studied. Some examples of solid state media are listed below.

- *Flash Memory Cards.* Memory cards are made in a variety of sizes and range from around an inch square to around a centimeter long. Larger cards are often used in digital cameras and smaller cards in cell phones. There is a great variety of card types in use as well as the storage size of the cards.

- *USB Flash Drives.* Connected to a computer via a USB port, these ‘drives’ (storage devices) are a very portable option for data transfer.

- *Solid State Hard Drives.* Uses solid state memory to store data, with more capacity than memory cards and flash drives. The drives have no moving parts which are an advantage, but the cost per GB is currently more than with an external hard drive.

Digital Media Capacity

As indicated above, various types of media store different amounts of data. Storage capacity and file size is measured in bytes, the basic unit of measurement.

- 1,024 bytes make a kilobyte (KB)
- 1,024 KBs make a megabyte (MB)
- 1,024 MBs make a gigabyte (GB)
- 1,024 GBs make a terabyte (TB)
- 1,024 TBs make a petabyte (PB)

To put things in perspective, a one page Word document may be only 27 KB; a 75 page document is 425 KB; and a Word document with images may be 20 MB. A single photograph, depending on size and quality, can range from less than 5 KB to 30 MB and higher. To best determine what media you would like to store files on, you will need to understand the amount of storage you need now as well as in the future. This will help determine which storage option is the best for you.

Media Life Expectancy

All digital media has finite life spans which are dependent on a number of factors, including manufacturing quality, age and condition before recording, handling and maintenance, frequency of access, and storage conditions. Studies have indicated that under optimal conditions, the life expectancy of magnetic media ranges from 10 to 20 years for different types, while optical media may last as long as 30 years. However, in real life situations, most media life expectancies are significantly less.

Care and Handling of Digital Media

To help make your digital media last as long as possible, follow the guidelines below. This list is not complete, as each media type will have its own requirements for proper handling.

All Media

- Purchase and use high quality storage media. Batch test new media to validate manufacturing quality.

- Read a statistical sample (3% minimum) of recorded media annually to identify and correct any loss of data. Re-copy batch if errors appear.

- Prohibit smoking and eating in areas where digital media are stored and also in media test or evaluation areas.

- Maintain media in storage areas that are dust-free and controlled for temperature and humidity.

- Open a recordable media package only when ready to record.

Magnetic Media

- Wind and rewind magnetic media before recording.

- Every three to four years — or more frequently if you read them often — rewind each tape under controlled tension.

- Before they are five years old, re-copy tapes onto new and/or updated tapes.

- Minimize handling and avoid touching the media surface or edges.

Optical Media

- Do not touch or mark the data side of the disk surface. Handle disks by the outer edge or center hole.

- Be careful not to damage the label side of the disk. Do not apply or attempt to reposition adhesive labels. Do not write on disk with pen, pencil, or fine tip marker. Use only non-solvent based felt tip permanent marker.

- Check disk for damage or contamination after each use. Clean only when surface contamination is visible by wiping the disk from the center out in a radial motion with an anti-static cloth.

- Depending on use and storage conditions, CDs and DVDs should be re-copied every five years or sooner.

Solid State Media

- Be careful not to get the media wet.
- Because of the small and portable nature of many flash drives and memory cards, be careful to not lose or misplace the media.
- Be careful with inserting the card into a card reader or flash drive into the USB port. Use your computers 'Disconnect Safely' tool when finished.
- Do not expose to great temperature fluctuations.

Storage Options

As part of a records management plan for electronic records, you will need to determine where and how these records will be stored. This decision will be based on the likelihood of access of those resources versus the overall cost in maintaining them. Your options for storage include online, near-line, and offline.

Online: Records stored online are continually accessible via a network. Records are located on a hard drive or networked server. This option maintains the greatest functionality but requires more expensive network storage.

Near-line: Records stored in a near-line environment are stored on removable media such as on an optical disk or magnetic tape. Files can be accessed via the network, but are not physically on the network (e.g., an optical media jukebox). This option maintains a moderate amount of functionality. While the storage space is cheaper than online storage, near-line storage requires that the user take time to manipulate both the files and media of choice to access the records.

Offline: Records stored offline are stored on removable media that is not accessible through a network. Files must be physically retrieved from the digital media itself, such as on an external hard drive or magnetic tape. This option trades functionality for stability, but maintains records in a digital format.

For more information on storage procedures and facility requirements please see the *Digital Media Storage* chapter of these guidelines.

Performance Issues to Consider

As you discuss your digital media options, consider each option's performance characteristics in terms of your records management needs.

- *Planning.* In addition to choosing a storage medium, you should establish procedures to refresh your digital stored records periodically. Refreshing digital media occurs when you copy stored data from old to new digital media. To verify that there was no loss of content or that the content was not changed during the transfer, perform integrity checks (such as a checksum) on the content both before and after the process.

- *Speed of access*. When selecting a digital storage medium, consider how quickly you or authorized members of the public may need to access your records. You may find that some types of records require fast access, while others do not. For example, you may need fast access to key policy decisions, but not to employee records.

- *Capacity*. The volume of records that you can store on the medium will be a key consideration. Examine the volume of the records you now store, and try to determine what your needs may be in the future. Consider the official definition of a record and whether that definition will affect the records volume that you need to manage.

- *Longevity*. Research how long the industry will support various media options and compare those figures with the time period that you need to keep your records according to the approved records retention schedule. You may find a medium that meets all your needs, but is not widely used or has a high risk of becoming obsolete, thereby limiting its usefulness in the future.

- *Durability*. Research how easily a given medium can be damaged or will deteriorate. You may find that a medium that deteriorates after three years will still be a suitable option for records that need to be retained for only one year. Be sure to review your records retention periods.

- *Portability*. Determine how portable your stored records should be. Some media, such as flash drives and DVDs are very portable. Consider who will be accessing your records. For example, will the public, the press, or other agencies frequently access your records? You should also consider whether you or anyone accessing the records will need special devices to read the records. What equipment will be necessary to view the files?

- *Compatibility*. Assess the backward and forward compatibility of the digital media you are considering. For example, DVD drives are backward-compatible for CDs, but a CD drive is not forward-compatible for DVDs. This discussion will help you to determine how often you may need to upgrade supporting computer systems, migrate records, and/or convert records.

- *Cost.* Assess the costs and benefits of each medium you consider. What will the medium itself cost as well as the equipment required to create and view it? Be sure to discuss the costs of converting and/or migrating records, as well as the basic costs of the system.

Key Issues to Consider

Now that you are familiar with some of the basic concepts and options of digital storage media, you can use the questions below to discuss how those concepts relate to your agency.

Pay special attention to the questions posed by the legal framework, including the required records retention periods. Examine your current and future records series. Some records series may require large storage capacities, may need to be retained for a long time, or may be frequently accessed by the public, other agencies, or other groups. Prioritizing your needs in light of the legal requirements will help you narrow your discussion and focus your research.

The point is to determine the best option for your agency that meets your legal and operational needs, not merely to automatically upgrade technology. For example, if you are currently using magnetic tape, you may discover that magnetic tape remains your best choice.

Discussion Questions

- What types of records do we need to store (e.g., graphics, text, database text)? What file formats? How large are our record files?
- Which performance issues are most important in our situation?
- How long do we need to retain the records?
- How often will we need to access the records?
- Will all records or specific records series be frequently accessed by the public or other groups?
- How well does our current media meet our needs? What costs would be incurred for supplies, equipment, and training that would be required if we were to switch to or add a new storage medium?
- Are any of the media that we use or are considering expected to become obsolete in the near future? Will the medium, as well as the necessary hardware and software, still be available from a number of suppliers for as long as we need? Has the developer defined a migration path for improved versions of the medium?

Annotated List of Resources

Byers, Fred R. *Information Technology: Care and Handling for the Preservation of CDs and DVDs – A Guide for Librarians and Archivists*. NIST Special Publication 500-252. Gaithersburg, MD: National Institute of Standards and Technology; Washington, D.C.: Council on Library and Information Resources. October 2003.
http://www.foray.com/images/pdfs/CDandDVDCareandHandlingGuide.pdf

> This guide discusses the physical characteristics of various optical media, as well as methods for their proper care and handling to ensure longest possible use in any given environment. A useful glossary is included.

National Archives Records Administration. *Frequently Asked Questions about Optical Media*. August 6, 2007.
http://www.archives.gov/records-mgmt/faqs/optical.html

> These frequently asked questions provide a brief overview of optical media from a records management perspective including answering how information is recorded onto optical media and addressing issues of stability.

National Archives Records Administration. *Frequently Asked Questions (FAQs) about Optical Media: Storing Temporary Records on CDs and DVDs*. August 6, 2007.
http://www.archives.gov/records-mgmt/initiatives/temp-opmedia-faq.html

> These frequently asked questions provide general information on using optical media as storage media.

The PC Technology Guide
http://www.pctechguide.com

> This site is a comprehensive resource on all aspects of the personal computer. Topics include hardware, software, computer use, and digital media.

Webopedia
http://www.webopedia.com/

> This comprehensive online encyclopedia for the information technology community provides an easy-to-understand, searchable database of terms.

COOL (Conservation OnLine): Electronic Storage Media
http://cool.conservation-us.org/bytopic/electronic-records/electronic-storage-media/

These pages are part of the Conservation OnLine, Resources for Conservation Professionals web site operated by the Foundation of the American Institute for Conservation. This web page is a collection of materials from other sources about electronic conservation, including resources on disaster recovery, electronic media, electronic formats, and storage environments.

Digital Media Storage

Summary

State and local governments use computers to create, capture, or maintain public records. To be accountable to the citizens of Minnesota, government agencies are required by law to keep records documenting their activities. Many of these records, because they will be of long-term or enduring value, must remain accessible over time. Some will be needed to continue critical government operations, some to document programs, and some to provide legal evidence. While the law does not require that the records be kept permanently, the approved retention schedule may dictate that they are kept for an extended period of time (10+ years). For practical reasons, you may want to remove the records that you do not refer to frequently from an online system to a lower-cost off-line storage facility until their disposal date. To ensure timely access to automated information, users must be able to identify and retrieve records online, near-line, or off-line. If stored off-line, records of enduring value will require special maintenance due to the basic instability of the media as well as system conversions that may jeopardize their long-term safety.

Legal Framework

For more information on the legal framework to consider when considering storage options for your records, refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[39]. Keep in mind the legal requirements you may have for providing access to records over time.

Key Concepts

As you discuss and develop a plan for digital storage, you will need to consider:

- Storage Options
- Planning and Maintaining Digital Media Storage
- Storage Environment Guidelines
- File Storage with a Third-Party
- Storage Facility Details

Storage Options

[39] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

As part of a records management plan for electronic records, you will need to determine where and how your records will be stored. This decision will be based on the likelihood of access of those resources versus the overall cost in maintaining them. Your general options for storage include online, near-line, and offline. Records that are accessed most should be stored online; records accessed the least off-line. After it is determined where the records will be stored you will need to determine how to provide various users with access. This will depend on the records themselves as well as the type of media chosen.

Make sure to select appropriate media and systems for maintaining your records for the required period of time. Regardless of your choice, files may need to be refreshed (transferred to new media) or migrated (to a different format) possibly within a pre-determined period of time. For more information about the types of digital media please review the *Digital Media* section of these guidelines. Online, near-line, and offline storage are described below.

Online Storage

Online storage allows immediate access to records to anyone on the system's network. Properly designed storage will provide access to appropriate users only. Online storage maintains the greatest functionality but requires more expensive network storage.

Examples of online storage include:

- *Storage Area Networks (SAN).* SANs allow access to remote drives with the same convenience of internal hard drives. A SAN is a networked system.
- *Just a Bunch of Disks (JBOD).* JBOD is a collection of disks that are set up to look like one large disk to users. If one disk goes bad, only the information on that disk is lost; the information on the others is still readable. This is sometimes used to store backup files.
- *Redundant Array of Inexpensive/Independent Disks (RAID*). RAID uses a group of disks to back up data on a daily basis. There are many different configurations of RAID arrays with their own features. RAID is frequently used, but depending on the size of your institution, may or may not be cost effective. Although RAID is used to backup local files, it is not intended to be used as a main backup method.
- *Cloud Computing*. More recently, online storage also refers to files that can be accessed online via the Internet. With cloud computing files are stored by a third-party and accessed through a web service. This technology continues to be explored. It may not be appropriate for confidential files or vital records that need to have 24 hour access. As with any outside provider, questions need to be asked about their policies and procedures for storing, preserving, and providing access to records.

Near-line Storage

Near-line storage is storage in a system that is not a direct part of your network, but that can be accessed through your network (e.g., an optical media jukebox). Access to these files is done with an automated process that selects the correct disk/tape from a disk/tape library and makes it accessible. This option maintains a moderate amount of functionality. While the storage space is cheaper than online storage, near-line storage requires extra time to manipulate both the files and media of choice to access the records. Near-line storage is often used for backups as large quantities of data can be managed quickly.

Offline Storage
Files that cannot be accessed immediately are said to be stored offline (e.g., files not accessible through your network such as on removable media like external hard drives or magnetic tape). Older records or records that do not need to be accessed frequently are often stored online. This option trades functionality for stability. There is a focus on data accuracy, protection, and security due to the long-term storage necessity. The longer records need to be maintained the more important preservation methods, back-up procedures, storage conditions, handling procedures, and security become. Offline storage can be stored in-house, off-site, or outsourced.

Examples of offline storage include:

- Removable magnetic or optical media (tape, DVD). For integrity purposes, read-only media is preferred.
- Flash media (solid state media has no moving parts)
- External hard drives (with moving parts)

Planning and Maintaining Digital Media Storage

Even when properly cared for, all digital media and hardware have limited life expectancy. Media life spans are dependent on a number of factors, including manufacturing quality, age and condition before recording, handling and maintenance, frequency of access, and storage conditions. Hardware and software may be supplanted by rapid advances in technology. Therefore, storage of digital media demands greater planning and attention than the traditional formats such as paper or microfilm. The suggestions in this chapter provide basic information on the design and management of a digital storage facility. For more information on choosing and caring for digital media, refer to the *Digital Media* chapter of these guidelines.

Record Maintenance

For records of long-term or enduring value stored on electronic media, agencies should consider the following:

Access: Maintain your records in a usable format and keep up-to-date materials needed to access them, including indexes and other documentation, until they are scheduled for disposal. In instances where you maintain non-confidential public records permanently in your agency, you will need to create a plan that provides easy access to those records upon request.

Backups: Maintain backup copies of records and all materials required to access them in an off-site, preferably geographically different, location that does not share the same disaster threat. Create policies and procedures for backing up records.

Labeling: Develop procedures for labeling storage media. Each external label should carry information unique to the medium it identifies. At minimum, it should display the name of the organizational unit responsible for the data, the system title, the file title, the disposition date or permanent status of the record, and its security classification, if applicable. Larger storage systems (such as an external hard drive) should have at the minimum a printed inventory of the files it contains.

Inventories: Develop procedures to maintain an accurate and up-to date inventory of records stored off-line. If using tapes or other portable media, a useful inventory will contain the following information about each item: item ID; file title(s); system title; dates covered by file(s); date moved off-line; the recording density; type of internal labels; volume serial number, if applicable; number of tracks; character code/ software dependency; information about block size; and the number of the item if part of a multi-item set. Where applicable, it will also give the number of records for each set of data, the format of the record, and logical record length.

Storage Environment Guidelines

Each type of storage location has specific requirements to function properly and protect your digital records. Requirements for online, near-line, and offline storage environments are discussed below.

Online

All types of online storage for a local network depend on computer equipment and servers. Networked computer equipment is generally housed in a ‘server room’.

Some requirements for a well designed server room include:

- *Location and Accessibility*. If possible, centrally locate your server room. This will make it easier to connect computers to the network. A sever room should be easily accessible to authorized personnel to facilitate monitoring and maintenance when needed. Security systems should be in place to keep unauthorized personnel out and monitor who accesses the room.

- *Size*. If designing a networked system, make sure the space you dedicate to the server has enough room to allow for growth over time. Using a space to its capacity at the outset will be detrimental in the future.

- *Storage*. Depending on the size and number of your servers, you can store them on wall mounted racks or on floor racks. These racks separate the individual servers and keep them from touching each other, which could cause overheating and other physical damage.

- *Temperature*. To keep equipment from overheating and being damaged, the server room should be temperature controlled, and set at about 65-75 degrees. It is a good idea to have separate temperature controls for the server room; relying on central air conditioning that cools an entire building is not a good idea, as temperatures can fluctuate drastically throughout the building.

- *Power*. Make sure there is an appropriate amount of power being delivered to the server room. Make sure that more power is available for expansion as needed. If the amount of power is not adequate, the servers will overheat and fail.

- *Cable Management*. Many cables will be attached to the servers. It is important to make sure these cables are well organized. Make sure cables are not twisted, bent, resting on the floor, or under any pressure. Damaging the cables will damage the network connections and possibly the data on the servers. Labeling cables is also a good idea.

- *Environment*. Keep the server room clean; dust can damage the servers. Protect the servers from water damage (sprinklers, leaky pipes). Keep magnets away, since magnets can damage digital data on magnetic storage media.

- *Documentation*. Overtime, hardware and software updates will need to be performed on the servers. Keep a running log of all updates. This can help problem solve technology issues that may arise in the future.

Near-line

Near-line storage equipment requires the same general storage requirements as online storage equipment. The physical media the records are stored on will also have their own care and handling procedures. Please review the 'Care and Handling' section in the *Digital Media* chapter of these guidelines to understand the care and handling of the media itself.

Offline

Offline storage consists of removable media including magnetic tape, optical disks, or external hard drives housed in a storage facility. For care and handling of the removable media itself review the 'Care and Handling' section of the *Digital Media* chapter of these guidelines. The desirable qualities of a storage facility for this media include:

- *Adequate floor space.* You will need to consider:
  - The current volume of media you need to store.
  - The projected volume of media you will need to store in the future based upon your records retention requirements.

- *Security.* Allow only approved people access to the storage facility. You will want to consider, among other things:
  - A controlled auditable entrance (e.g., security code keypad, smart-card swipe).
  - An alarm system that sounds if an unauthorized person attempts to enter the storage facility.

- *Convenient location.* Consider how often you will need to access the records in your offline storage facility to help determine how conveniently located your storage facility needs to be.

- *Adjustable Lighting.* Your storage facility will need to have adequate lighting available for people using the facility.

- *Ventilation*. Good ventilation will help prevent dampness, mold, and pest infiltration.

- *Temperature and humidity control.* Proper temperature and humidity are essential for preserving the electronic records on digital media. Temperatures and humidity levels that are above or below the recommended range can deteriorate electronic (and paper) records. Above all, you should strive for a consistent environment, without sudden or drastic changes in temperature or relative humidity. A good temperature and humidity requirement for storage facilities is as follows: Temperature between 60 and 69 F. Relative humidity between 35-45%.

- *Clean air quality*. The air in the storage facility should be free from pollutants (e.g., strong cleaning solution fumes). Dust can also be particularly damaging to digital media.

- *Damage prevention*. Protect your storage facility from:
  - Pest infestation (e.g., mice, cockroaches, silverfish)
  - Fire, smoke, and sprinkler damage
  - Water damage, either from leaky pipes and leaky foundations, or from trapped moisture in walls, floors, and ceilings.
  - Damage from magnets, since magnets can damage digital data on magnetic storage media and thereby damage your electronic records.

File Storage with a Third Party

You may also consider using a third-party storage facility that can store, access, and deliver records to you. Third-party services include management of offsite storage facilities as well as cloud computing technologies, both of which are contracted out. Cost-benefit analysis should be done to determine if working with a third party will be beneficial to your institution. Be certain that the third-party policies, procedures, and facility can meet your operational needs and legal requirements.

Storage Facility Details

In addition to the above requirements, if you use a storage facility, you will need to:

- *Establish a policy*. Your storage facility and procedures policy should mesh with your overall records management strategy. Address both operational and legal requirements to ensure that you store and handle your records in accordance with state laws, while also meeting your operational needs.

- *Evaluate the physical storage space.* Storing your electronic records in a space designed for that purpose will help you maintain your records as long as legally and operationally necessary.

- *Develop access procedures.* Procedures for access and use of the storage facility must detail who physically and electronically may access the facility, retrieve records, add records, and dispose of records.

Determine your needs, priorities, and budget for the following components of a storage facility:

- *Storage Aids.* Appropriate storage aids for the media may include shelving, file cabinets, and storage boxes. You may also need special cleaning supplies (e.g., lint-free dusting cloths, cotton gloves for handling sensitive media).

- *Facility map.* Consider creating a map of the storage facility so that you know which digital media are stored in each area.

- *Access and use training*. Provide instruction and training for staff members who will be submitting items for storage, accessing stored records, and retrieving records. Established guidelines and training will enable you to provide service, stay organized, and protect your records.

- *Circulation control.* Develop a circulation log or other method for tracking facility access and records circulation. For a reliable circulation control system, you will need to develop an indexing system that accounts for all the digital media stored in the facility. A central authority should manage the index's content. Options include a paper list, card file, or database. You should be able to look at the circulation control index and determine the exact status of each stored media (e.g., if checked out, with whom and when due; if disposed of, when destroyed or disposed of; date of final disposition).

- *Acceptance system.* Develop a process that allows agency members to place records into the facility. Items submitted for storage should have, at minimum, the:
  - Name of the record series
  - Security classification; open or restricted
  - Record series inclusive dates
  - Unique locator number or identifier
  - Name of the agency and/or department submitting the item
  - Records disposal date

- *Special consideration for vital records.* Your vital records should have the best storage facility you can devise and afford. Be certain that your facility map shows the location of digital media containing your vital records; so that you can locate them immediately should a disaster occur. An off-site storage location for back-up copies is best.

- *On-going maintenance schedule*. Establish an ongoing system for maintaining the storage facility, including:
  - Regular cleaning, using chemicals that will not leave harmful residue or fumes
  - Procedures for checking deterioration of physical storage media (e.g., warped compact disks, cracked disks, moldy boxes)

  - Procedures for checking deterioration of electronic content (e.g., unreadable disks, inaccurately read records, missing or scrambled information on records)
  - On-going maintenance program (e.g., reading samples, spinning tapes to tighten them)
  - Regular maintenance of storage facility equipment (e.g., furnaces, air conditioners, dehumidifiers)

- *Disaster recovery (Continuity of Operation) plan.* As part of your records management policy, include a disaster recovery plan that provides a series of detailed actions (including who is responsible for executing each step of the disaster plan) if a disaster should occur at the storage facility. Include the response procedures for multiple types of disasters (e.g., flood, fire, smoke, explosion). The goal of the plan should be to have the facility operational and the greatest number of records recovered in the least amount of time. Train staff members and practice the disaster recovery plan. For more information on disaster recovery, refer to the Disaster Preparedness guidelines[40] on the Minnesota State Archives' web site.

Key Issues to Consider

Now that you are familiar with some of the basic concepts of storage facilities, you can use the questions below to discuss how those concepts relate to your agency. Pay special attention to the questions posed by the legal framework, including the need for public accessibility and protection of not-public records as set forth in the Minnesota Government Data Practices Act (MGDPA). Consider your current and future activities and records to help determine your requirements for a storage facility. The answers to these questions will guide your development of a storage facility that meets your agency's needs and legal requirements.

Discussion Questions

- What are our goals for storage and access? How do we prioritize these goals? How does this prioritization affect our budget?

- Are there other government agencies that are able to share resources? Collaborations?

- How long do we need to retain our records according to applicable retention schedules?

- Will we be storing an increasing volume of electronic records over time?

- How frequently will the records need to be accessed? How strictly must access to the records be monitored? Will the public access our records directly, or will we access records on behalf of the public? How will we protect not-public records as defined under the MGDPA?

- What are our needs for floor space, storage aids, location, and security systems?

[40] Minnesota Historical Society. *Disaster Preparedness*. Minnesota State Archives. March 2003. http://www.mnhs.org/preserve/records/disaster.html

- Will the storage area be located in our daily work space or in a separate location? What are the cost differences of our options?

- Are we considering a third-party storage facility? How will we be sure that the third-party can meet all of our legal and operational requirements?

- Who is responsible for enforcing the storage system policy and procedures? Who will maintain the map and index?

- How will we accept and process records into the storage facility?

Annotated List of Resources

Byers, Fred R. *Information Technology: Care and Handling for the Preservation of CDs and DVDs – A Guide for Librarians and Archivists*. NIST Special Publication 500-252. Gaithersburg, MD: National Institute of Standards and Technology; Washington, D.C.: Council on Library and Information Resources. October 2003.
http://www.foray.com/images/pdfs/CDandDVDCareandHandlingGuide.pdf

> This guide discusses the physical characteristics of various optical media, as well as methods for their proper care and handling to ensure longest possible use in any given environment. A useful glossary is included.

Minnesota Historical Society, Minnesota State Archives. *Preserving and Disposing of Government Records.* St. Paul: Minnesota Historical Society, May 2008.
http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

> Developed for Minnesota government agencies, this overview of the basic principles of records management includes chapters on defining a government record, taking inventory of your records, developing records retention schedules, preserving archival records, disposing of records, and setting up a records storage area. A list of resources for more information is included, as well as information about applicable state law regarding electronic records management. Originally published by the Minnesota Department of Administration in July 2000, the guide was updated jointly by the Minnesota Historical Society and the Minnesota Government Records and Information Network (MNGRIN) in 2008.

Minnesota Historical Society, Minnesota State Archives. *Disaster Preparedness*. St. Paul: Minnesota Historical Society, 2000.
http://www.mnhs.org/preserve/records/disaster.html

> Also available as a downloadable file, the information on these web pages summarizes the basic concepts of disaster preparedness, including disaster prevention, disaster planning, disaster recovery, and disaster preparedness resources.

National Archives and Records Administration (NARA). *Frequently Asked Questions About Managing Federal Records in Cloud Computing Environments*.
http://www.archives.gov/records-mgmt/faqs/cloud.html

> Includes explanation of cloud computing including the benefits and concerns of the technology.

Wyld, David C. *Moving to the Cloud: An Introduction to Cloud Computing in Government*. IBM Center for the Business of Government. 2009. http://www.businessofgovernment.org/article/moving-cloud-introduction-cloud-computing-government

> An introduction to cloud computing including its current use in government and challenges with implementing it.

The PC Technology Guide
http://www.pctechguide.com/pc-data-storage

> This site is a comprehensive resource on all aspects of the personal computer. Topics include hardware, software, computer use, and digital media.

COOL (Conservation OnLine): Electronic Storage Media
http://cool.conservation-us.org/bytopic/electronic-records/electronic-storage-media/

> These pages are part of the Conservation OnLine, Resources for Conservation Professionals web site operated by the Foundation of the American Institute for Conservation. This web page is a collection of materials from other sources about electronic conservation, including resources on disaster recovery, electronic media, electronic formats, and storage environments.

Digital Imaging

Summary

Government agencies use digital imaging to enhance productivity, provide greater access to certain types of information, and as a preservation option. Digital imaging offers many advantages, including: improved distribution and publication, increased access, streamlined workflows, and a greatly reduced need for physical storage space. Digital files made available over the web allow government agencies to provide information to partners or the public quickly and efficiently. In addition, through the application of optical character recognition (OCR) software, digital images can be used to create text-searchable files which increase access and use.

While digital imaging is a popular option for access and long-term preservation, it is an investment with potentially very high up-front costs. Continuing investments in all aspects of an imaging process are also required on a routine and frequent basis. Digital imaging should make financial sense for your agency. To assure your digitized records are fully admissible in court, they must be trustworthy, complete, and durable for as long as your approved records retention schedules require.

Legal Framework

Imaging is, by Minnesota state law, a recognized and legitimate form of record reproduction. Therefore, if the images replace the originals, they are subject to the same legal requirements as the originals. To ensure such digital records are fully admissible in court, you must be able to demonstrate that the records are complete and were created in a trustworthy manner. Following record keeping regulations will also assist you in managing your records appropriately. Legislation varies from state to state, so it is important to research the requirements that may affect your particular digitization project, such as access to the records and disposition of the originals. Federal, state, local and organizational policies also apply.

Government records generally have more requirements to follow than business records. Laws relating to the “collection, creation, storage, maintenance, dissemination, and access”[41] of government records are common. Regulations specifically affecting government records usually address such issues as privacy and security, retention and disposition, and public access to information. To assure that your imaged records are fully admissible and meet all evidentiary standards, you should review the requirements for each law in the *Legal Framework* chapter of these guidelines as well as the Uniform Photographic Copies of Business and Public Records as Evidence Act[42]. This Act establishes that an accurate reproduction of a record is as admissible in evidence as the original in any judicial or administrative proceeding. It further stipulates that if

[41] Minnesota Office of the Revisor of Statutes. *Minnesota Statutes: 13.01 Government Data*. https://www.revisor.leg.state.mn.us/statutes/?id=13.01

[42] Minnesota Office of the Revisor of Statutes. *Minnesota Statutes, Chapter 600.135: Photographic Copies of Business and Public Records.* https://www.revisor.mn.gov/statutes/?id=600.135

an accurate and durable reproduction is made, the original record may be destroyed in the regular course of business unless its preservation is required by law.

Key Concepts

Before you determine whether digital imaging will meet your long-term legal and operational needs, acquaint yourself with the following key concepts:

- Imaging Terms
- Cost Justification
- In-House vs. Outsourcing
- File Formats
- Metadata
- Image Storage
- Preservation Strategies
- Retention Schedules and Disposition of Originals
- Providing Access
- Implementation Strategy

Imaging Terms

Digital imaging is a process by which a document or photo is scanned by computer and converted from analog format to a computer-readable digital format. After scanning, the original document or photo is represented by a series of pixels arranged in a two-dimensional matrix called a bitmap or raster image. This image can then be kept on a network or transferred onto a variety of electronic storage media, such as DVD, for storage and use.

For a better understanding of imaging you should be familiar with the following terms:

- Pixel Bit Depth: The number of bits used to define each pixel. The higher the bit depth, the greater the number of tones (color or grayscale) that can be represented. Digital images can be bi-tonal, grayscale, or color. In general, higher bit depths are recommended for master images to accurately represent the original document.

  Table 1: Common pixel bit-depths

Bit-depth	Displays	Recommended for
1-bit or “bi-tonal”	black and white	Typewritten documents
8-bit grayscale	256 shades of gray	Black and white photographs, half-tone illustrations, handwriting
24-bit color	Approximately 16 million colors	Color graphics and text, color photographs, art, drawings, maps

- Resolution: The quality of a digital image is dependent on the initial scanning resolution. Resolution is expressed in the number of dots, or pixels, used to represent an image, expressed commonly as “dpi,” dots per inch. You may also see “ppi” (pixels per inch) and “lpi” (lines per inch) used. As the dpi value increases, image quality increases but so does the file size. [“’Lines’ or rows of pixels is a term used within the photographic industry as a common shorthand for the number of pixels across the long dimension of digital images of photographs.”[43]]

To determine the scanning resolution you need, you must first determine the desired quality of your images and the storage capacity of your computer system. You will also need to consider the desired speed of delivery of the images, especially if they will be accessed over the Internet. You may want to scan high-resolution masters of your images and then create lower resolution copies for web delivery. General recommendations for master files are listed in the table below; however, there are many other factors that need to be considered before selecting a scanning resolution such as size and quality of original document and desired results.

Table 2: Common scanning resolutions for master files[44]

Material	Recommended resolution (8-bit grayscale and 24-bit color)
Textual records	400-600 dpi
Photographs, negatives, slides	4000-8000 lpi.
Prints, paintings, drawings	600 dpi.

With the large variety of sizes for photographs and photographic material, in order to consistently produce high-quality images, the resolution of photographs is sometimes expressed in the number of pixels across the long-dimension of an image. When creating standard-sized images from photographs or negatives of differing sizes (e.g., 35mm, 4”x 5”),

[43] Maxine K. Sitts, *Handbook for Digital Projects: A Management Tool for Preservation and Access* (Andover, Massachusetts: Northeast Document Conservation Center, 2000). http://www.nedcc.org/resources/digitalhandbook/dman.pdf

[44] This is a very simplified chart of common scanning resolutions. For more details please review: Federal Agencies Digitization Initiative Still Image Working Group (FADGI). *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files.* August 2010. http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf

the scanning resolution in dpi varies. In such cases, it is often easier to measure resolution as the number of pixels across an image's long dimension. For example, each of the following files measures 3000 pixels in the long-dimension, although they have varying values of dpi. Some experimentation may be required to find the best resolution for different materials being digitized for a project.

Table 3: Resolution as the number of pixels across the long-dimension of an image[45]

Original photo size	Digital image size	Scanning resolution
8"x10"	2400 x 3000 pixels	300dpi
4"x5"	2400 x 3000 pixels	600dpi
35mm negative	2400 x 3000 pixels	2100dpi

- Compression: Compression is the process of reducing the file size of an electronic file, which saves file space. There are two types of compression, lossless and lossy. Under lossless compression the file is compressed without the loss of data. In the process of lossy compression, data is lost as lossy compression attempts to eliminate redundant or unnecessary information. Depending upon the degree of compression, this information loss may be unnoticeable to the human eye. For example, it is possible for a JPEG file (a lossy compression) and a TIFF file (lossless) to appear exactly the same, although the JPEG file is missing data, making it significantly smaller. These file formats, and others, are discussed in the following section.

- Master Images and Access Images. Images created with the intent of replacing an original document will be considered a master image or master copy. Master copies should be of high quality and follow recommended standards that ensure complete and trustworthy records. Master copies are not used on a regular basis. Access images are generally copies of master files whose main purpose is to provide access to users on a regular basis. There are standards and best practices for access copies as well; however access copies are generally lower in quality resulting in smaller file sizes which allows for easier access.

- Optical Character Recognition (OCR). A process that translates words in a digital image into machine readable text. These words can then be used by machines in various ways, including for full-text searching or editing.

Cost Justification

While digital imaging is popular and commonplace, you must remember it is an investment with potentially very high up-front costs. You also need to keep in mind that, because of the rapid

[45] Information taken from: Maxine K. Sitts, *Handbook for Digital Projects: A Management Tool for Preservation and Access* (Andover, Massachusetts: Northeast Document Conservation Center, 2000), page 97. http://www.nedcc.org/resources/digitalhandbook/dman.pdf

pace of technological obsolescence, you will need to make continuing investments in all aspects of an imaging process on a routine and frequent basis.

Digital imaging, as an investment, should make financial sense for your organization or agency. A comprehensive analysis will help estimate costs and evaluate possible benefits for your agency. Costs of digital imaging occur during project development, the digitization process, and continue as the digital collection is maintained over time and all must be taken into consideration. Examples of such costs are listed below.

> *Project Development*
> Project development includes the costs to select, prepare, and catalog the documents that are to be digitized. Cataloging includes creating or linking any necessary metadata to the original object. Selection of appropriate hardware and software is also part of this phase.
>
> *Digitization Process*
> The digitization process includes creation of a digital image, entering of the metadata, and developing and implementing a system to store the images. Providing access to the images could be considered part of this process or part of the ongoing costs, as could database creation.
>
> *Ongoing Costs*
> Ongoing costs include the salary and benefits of current and new staff, money for external technical support if needed, additional training costs, maintenance of hardware and software, replacement costs for failed or obsolete equipment or software, vendor contracts, and Internet connections.[46] These costs continue to occur after the collection has been digitized.

Benefits of digital imaging include better customer service, higher office productivity, lower storage costs, and the option of using the Web to make digitized information easily accessible. However, due to the expense associated with imaging, justifying imaging systems based only on potential cost savings is not recommended.

In-House vs. Outsourcing

One of the first decisions to be made is who will be doing the digital imaging. Most agencies and businesses do not have the appropriate scanning equipment, software, or staff expertise to execute a large digitizing project. Evaluation of your resources will help determine if your digitization process should be done in-house or outsourced to a vendor who specializes in digital imaging.

Vendors provide digitizing services, technical advice, and sometimes the long-term maintenance of the resulting files. Before talking to vendors, be familiar with digitizing technology, the terms used by the industry, and have a clear idea of your project and its goals.

[46] Roderick, Elizabeth. *More Than Just Pretty Pictures: A Cost/Benefit Analysis of Digital Library Holdings*. CAUSE98 EDCAUSE Conference. December 9, 1998. http://net.educause.edu/ir/library/html/cnc9804/cnc9804.html

Some questions to ask include:

- How much material will be digitized? What type of materials will be digitized? Textual documents? Photographs? Maps?

- Is there non-public information included in the materials to be digitized? If so, can the materials leave your site? What precautions are necessary to ensure the security of the materials?

- How much time do you have to devote to digital imaging? What resources are currently available? Scanning equipment? Computers? Software? Staff expertise?

- What is the physical condition of the materials? Do they need to be prepared for scanning (removing staples and paperclips)? Do they have any special handling requirements that would keep them from being outsourced? Can they be transported easily?

- What is the required quality of the digital images? High or low resolution? Black and white or color?

- What is the desired end product? A document management system? A searchable online collection? Who is the intended audience? Staff members? Researchers? The general public?

- Why are you digitizing the materials? What file format(s) fit your requirements? Do you need both master and access copies? How will each be created? And when? Do the access copies need to be watermarked?

- What will happen to the original paper documents that were imaged? Do they need to be kept for any reason? Local access? Retention schedules? If not, how will they be properly disposed of?

When answering these questions you may find that some of the process is best done in-house, while outsourcing other portions is a better choice. For example, you might choose to do most of the project yourself while outsourcing a few tasks or do a little prep work and outsource most of the project. As you start evaluating your choices keep in mind the possibility of business failure and the inevitability of product obsolescence. The best way to protect yourself is to insist on an open-system architecture, using non-proprietary hardware and software. Non-proprietary means that the chosen hardware and software is not specific to that vendor. If proprietary software is unavoidable, it should be licensed beyond the length of the contract. As there will inevitably be some bugs in the system, a contract should completely spell out the provisions for implementation, service, upgrades, and repair.

The Northeast Document Conservation Center highlights issues relating to working with vendors in the preservation leaflet *Outsourcing and Vendor Relations* [47] including details on how to find a vendor, how to interview vendors, and how to work with vendors.

File Formats

In any digital imaging project, choosing the file formats you will use is important. Like scanning resolution, the file format directly affects the quality and file size of your images. Choosing the best file format for your needs requires knowing the type of materials you will be imaging (e.g., test, art, graphics, photos), how long your images will need to be retained, and how they will be used (e.g., archival or display functions which effects the necessary quality of image and desired speed of delivery of the images). Master images, in formats such as TIFF, have large file sizes, making their delivery cumbersome for some web and document management system applications. To enhance the speed of delivery, you can create copy images from the master images. Derivative images have smaller file sizes, are of lower quality, and typically use a lossy compression. The JPEG file format is commonly used for copy images.

Choosing formats that provide access to the greatest number of people over a long period of time is ideal. General guidelines to follow include using open-source or non-proprietary formats, choosing formats that are widely available and accepted, and using formats that have become standards in the industry. The formats must be stable, well-supported, and well-documented. The most important concept to remember is overall readability and use[48]. People must be able to read and use the digital records over time. Non-proprietary is not always the best solution. TIFF, for example, is considered an archival standard by many even though the specifications for the file type are copyrighted by Adobe[49]. However, Adobe has made the comprehensive specifications for TIFF 6.0 public and states that “the goal is that TIFF files should never become obsolete and that TIFF software should not have to be revised more frequently than absolutely necessary.

Compression is another issue that must be considered; you must decide if file compression is acceptable or not. There are different types of compression, each with its own intended use. Lossy, loses information during the compression process, such as with a JPEG, while the lossless technique looks identical to the uncompressed file, as with a TIFF file. In general, for archive or master copies compression is not acceptable as information is lost during the compression process. Other new methods of compression have recently been developed, including using fractal and wavelet compression. JPEG 2000 uses wavelet compression and is a method that may allow ‘compressed’ files to be used as archival masters.[50]

[47] Dale, Robin L. *Outsourcing and Vendor Relations*. Northeast Document Conservation Center (NEDCC). 2007. http://www.nedcc.org/resources/leaflets/6Reformatting/07OutsourcingAndVendorRelations.php

[48] JISC Digital Media. *File Formats and Compression: Open –v Proprietary- Formats and Compressions*. March 2009. http://www.jiscdigitalmedia.ac.uk/stillimages/advice/file-formats-and-compression/

[49] Joint Information Systems Committee. *JISC Standards Catalogue*. October 2006. http://standards.jisc.ac.uk/catalogue/TIFF.phtml

[50] JISC Digital Media. *File Formats and Compression.* March 2009. http://www.jiscdigitalmedia.ac.uk/stillimages/advice/file-formats-and-compression/

Before a decision is made, you must also determine if there are any enterprise/agency/state guidelines that must be followed. The Minnesota Office of Enterprise Technology has produced the *Enterprise Technical Architecture* for the state of Minnesota, a guide that discusses the ideas behind the practices and standards for the state of Minnesota. "Data architecture describes how the State's electronic data should be defined, stored, maintained and retained to facilitate processing, accessing, sharing, and analyzing from any part of the enterprise for appropriate constituencies according to existing federal and state laws".[51] As is the case with any choice, choosing a file format is just one piece of the puzzle, and you must look at the entire project to see how all the pieces will fit together. The choice becomes an "attempt to balance the requirements for quality, stability, potential longevity and industry acceptance."[52]

Common types of digital image file formats include:

- *Tagged Image File Format* (TIFF) files, which are widely usable in many different software programs. TIFF files utilize lossless compression and are commonly used for master copies. TIFF graphics can be any resolution, and they can be black and white, grayscale, or color. TIFF is a very extensible format, allowing variations to be created for specific applications. Files in TIFF format end with a .tif extension.

- *Graphics Interchange Format* (GIF) files. GIF supports color and grayscale. Limited to 256 colors, GIFs are more effective for images such as logos and graphics rather than color photos or art. It should be noted that although the GIF format is widely used, it is technically proprietary. A lossless compression, files in GIF format end with a .gif extension.

- *Joint Photographic Experts Group* (JPEG) files. JPEG is a lossy compression technique for color and grayscale images. Depending upon the degree of compression, the loss of detail may or may not be visible to the human eye. Files in JPEG format end with a .jpg extension.

- *Joint Photographic Experts Group* (JPEG2000) files. Uses wavelet-based image compression to produce both lossy and lossless digital images. Lossless images may compete with TIFF files for archival quality masters. Files in JPEG2000 format use .jp2, .jpf and other file extensions.

- *Bitmap* (BMP) files. BMP files are relatively low quality and used most often in word processing applications. BMP format creates a lossless compression. Files end with a .bmp extension.

- *Portable Network Graphics* (PNG) files. A lossless compression designed to replace GIF files, PNG files can be ten to thirty percent more compressed than GIFs. PNG is completely

[51] Office of Enterprise Technology. *Enterprise Technical Architecture 2.02*. September 8, 2006. http://mn.gov/oet/images/EA_R_Enterprise_Technical_Architecture_Reference_2002-02.pdf
[52] InterPARES 2 Project. *General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation*. March 2007. www.interpares.org/display_file.cfm?doc=ip2_file_formats(complete).pdf

patent and license free and is of higher quality than GIF. Files in PNG format end with a .png extension.

- *Portable Document Format* (PDF) files. PDFs are useful for viewing and printing multiple documents and images. Commonly used to capture, distribute, and store electronic documents, PDF preserves the fonts, images, graphics, and overall "look" of the original digital files. As with the GIF format, the PDF format is proprietary, although widely used. Files in PDF often end with a .pdf extension.

- *Portable Document Format* (PDF/A) files. PDF/A, as standard file format for long-term archiving of electronic documents, is a subset of PDF. Files are 100% self-contained, and do not rely on outside sources for document information. ISO standard: ISO 19005-1:2005.

For a more in-depth discussion of file formats and their properties, refer to the *File Formats* chapter of these guidelines.

Metadata

Metadata, usually defined as "data about data" is used to describe an object (digital or otherwise), its relationships with other objects, and how the object has been and should be treated over time. A structured format and a controlled vocabulary, which together allow for a precise and comprehensible description of content, location, and value, are its basic elements. Metadata often includes items like file type, file name, creator name, date of creation, and the record's classification under the Minnesota Data Practices Act.

Metadata is crucial to any digital imaging project, enabling proper data creation, storage, retrieval, use, modification, and retention of your digitized records. In addition, standardized metadata helps validate the trustworthiness of your system and the legal admissibility of your digitized records in court.

Metadata is especially important in facilitating retrieval of digital images. Digital images are stored as graphic files. Unless you plan to use OCR, the only way to locate specific information will be through its metadata. Metadata makes it possible to locate, use, and evaluate information through standard search criteria such as subject heading, numerical identifier, or keyword.

In addition to descriptive metadata assisting with access, preservation metadata captures information that helps facilitate management and access to digital records over time. Preservation metadata focuses on documenting the provenance, authenticity, preservation activity, technical environment, and rights management of an object.

For more information concerning metadata, refer to the *Metadata* chapter of these guidelines.

Image Storage

Digital images can be stored online, near-line and offline. Online storage includes storage area networks (SANS); near-line includes optical jukeboxes; and offline includes removable magnetic, optical, and flash memory media and devices. Where and how you store your images will depend on access needs. However, it is highly recommended that you store master digital images on media that assure the stored records are tamper-proof, increasing the level of security for the data.

When determining the best storage method, one must consider your institution's current storage capacity and digital file management system. Is there enough space for the project at hand, for future projects? Does your current system work for you? Are you able to produce, manage, and store back-up copies of the files or will you need outside help? Are all your files stored in one place or do you have backups offsite? How often are backups done? Who is in charge of them? How are they documented? Do you have a disaster recovery plan?

Looking at all of the available storage choices, their benefits and potential problems, it may be hard to determine best practices for long-term storage. You must study your digital imaging project, determine what is important to your institution, and understand your current and future resources before determining the best storage method.[53]

Due to the limited life expectancy of digital media, no digital storage medium is adequate for the long-term or archival preservation of records. The most generous estimate of physical obsolescence is thirty years. Technological obsolescence, though, will probably come within five to ten years. As a result, you should assume the need to migrate all your files to a new storage medium on a regular basis. In the meantime, you will need to protect your stored data with a comprehensive back-up system.

For more information on digital storage and storage media, refer to the *Digital Media* and *Digital Media Storage* chapters of these guidelines.

Preservation Strategies

Once you have decided on a file format and a storage plan, the challenge will be to keep that file accessible and viable. Digital files are not able to sit on a shelf for decades like paper files could in the proper environment. The storage medium, file type, and software and hardware used to create and store the file all affects the file shelf life. Files must be preserved over time to ensure accessibility and use. The action of preserving digital files must be addressed in any digitization plan and should "involve a number of organized tasks associated with a variety of technical

[53] To highlight the lack of standards, the South Carolina Department of Archives and History recommended using magnetic tape for long-term storage of electronic records while the North Carolina ECHO (Exploring Cultural Heritage Online) program specifically recommends NOT using magnetic media for long-term storage. South Carolina Department of Archives and History. *Electronic Records Management Guidelines: Digital Media Storage - Facilities and Procedures Version 2.* March 2008. http://arm.scdah.sc.gov/NR/rdonlyres/E03AB5A2-2B90-490B-96BF-D3E838FABCF7/0/ermDMSFP.pdf and North Carolina ECHO (Exploring Cultural Heritage Online). *Digitization Guidelines: Chapter 6 – Digital Preservation*. 2007. http://www.ncecho.org/dig/guide_6preservation.shtml#6.4

approaches or strategies that ensure digital resources are not only stored appropriately, but also adequately maintained and thus consistently useable over time."[54]

There are three common methods for preservation of digital files: migration/conversion, technology emulation, and technology preservation, of which the first one focuses on keeping the digital material immediately accessible and the last two focus on the technology used to create the digital file.

Establishing a plan for preserving your data is required for any digital imaging project. The procedures should address many of the issues discussed in these guidelines as well as periodic checks that can help to identify any data loss that may occur over time for quality control and authenticity purposes. One must also keep up to date on new technologies and standards as they are developed, as they may become useful for future preservation activities.

For more information on preservation strategies, refer to the *Long-Term Preservation* chapter of these guidelines.

Retention Schedules and Disposition of Originals

Depending on the purpose of your digital imaging project, you will need to consider what to do with the original files. Do you preserve them or destroy them? Was your purpose for imaging to create greater access while preserving the originals or was your purpose to eliminate the paper records altogether? What are the retention requirements for the records? Can you legally dispose of them?

Before disposing of scanned materials, there are several things to first consider. If you are working with government records, for example, you must first determine if the records have a retention schedule. All records must be on an approved retention schedule, and you must have the authority to dispose of them before doing so. If your goal of imaging was to make the digital files the official version of the records and the current retention schedule specifies the paper copy as the official version, you must amend the retention schedule and have it approved by the appropriate authority before the paper copy can be disposed of. Internal policies should also be updated to reflect this change. If the records are not listed on a retention schedule, one must be created for them and approved before the records can be disposed of.

You must also determine the classification of the records you wish to discard. Is the content of the records public, or is there a level of privacy or security attached to them? This information should also be covered in the retention policy. If records contain non-public information, they must be disposed of properly. At a minimum, sensitive paper records must be shredded; using the cross-cut method is more secure than the strip method of shredding. For highly sensitive materials, methods of disposal include being "pulverized (rendered into a powder by grinding), macerated (rendered into a pulp by chemicals) or incinerated (burned)."[55] Secure disposal is

[54] UKOLN at the University of Bath. *Good Practice Guide for Developers of Cultural Heritage Web Services: Digital Preservation*. April 11, 2006. http://www.ukoln.ac.uk/interop-focus/gpg/Preservation/
[55] University of Miami Leonard M. Miller School of Medicine. *Secure Data Disposal Methods*. http://it.med.miami.edu/x677.xml

usually required for any document that contains private information. The process of disposal should be written into a digitization plan or policy and should address record types, responsibilities (who is in charge), schedule (time frame) and location (in-house or outside vendor) when the records are to be disposed of. Risk prevention policies and a paper trail documenting the steps taken will assist you in the worst case-scenario of sensitive documents being improperly disposed of and misused.

If you do choose to dispose of the original files, do not be in a hurry to do so. You will want to be extremely confident that you have the legal authority to dispose of the records. In addition, you will want to make sure that you have no reason to go back to the paper records. Complete all of your cataloging, quality checks and indexing on the digital files. You may find that some of the records are unreadable, or were skipped, in which case you will need to access the paper records to correct this problem.

If you choose to keep the originals, you should evaluate the storage conditions. Follow best practices developed for storage environments, including temperature, humidity, and security for each record format.

Providing Access

Now that you have taken the time, money, and energy to digitize your records, how do you plan on providing access to them? Who will you provide access to? What are your terms and conditions for use? How will you ensure or verify the authenticity or trustworthiness of the records? How will you secure any non-public content?

Before determining a method for providing access, it might be helpful to determine how users currently access your records. Analyze the current process, determine if it is a good one, and modify your new access procedures as necessary, especially if providing access to digital materials is new to your institution. There are many options available, including creating your own custom web interface or using a digital asset management system.

In some cases it might be necessary to restrict access to documents with private or confidential information. There are a number of ways to do this, including linking access privileges to log-in type (e.g., the public cannot see the records, but staff can), flagging confidential records and having the system filter them out of search results or browsing options, or segregating confidential records so that access is through a separate process.

How will people find the information they are looking for? Will you create an index of the records? Will only certain fields in your database be keyword searchable or will a search be performed over the entire database? Will the documents themselves be searched? If so, do you need to employ optical character recognition (OCR) on them? Again, there are a number of commercial search tools available, or you may wish to develop one specific to your needs and records. Look at what others with similar records have done to get ideas.[56]

[56] The variety of interfaces range from simple point and click navigation to being able to search various fields within a record. In 2008, the Minnesota Office of the Revisor of Statutes Office began offering online digitized versions of the state's session laws dating back to 1849 (https://www.revisor.leg.state.mn.us/laws/). Users click through years

After people find the information they are looking for from your database or online interface, how do they know that the information they are receiving is authentic, accurate, or trustworthy? A report written about information assurance issues and requirements for the National Archives and Records Administration states that in order to have an authentic and secure environment, policies must address the availability, integrity, authentication, and confidentially of data.[57] Participants of the Minnesota State Archives led National Digital Information Infrastructure and Preservation Program project created a web page and wrote three white papers that addresses such issues.[58] There are many tools available that help protect your information and computer systems including use of firewalls, intrusion detection systems, file integrity checks, and secure computing technologies such as HTTPS. Completing a risk assessment of your records and computing environment will allow you to make appropriate, cost-effective decisions.

After the necessary steps have been taken to authenticate the records and establish yourself as a trusted source, it is important to inform people about the terms and conditions of use. Conditions of use inform users of your policies, provenance of the collections, copyright holders, and permissions of use specific to the records of interest. Contact information should also be included if the user requires more assistance. Informing the user of such things shows that you are a responsible curator for the information entrusted to you. This in turn fosters further authenticity of the information and a greater trust in the repository.

Implementation Strategy

To successfully implement a digitizing project in a timely manner, you must create an implementation strategy which manages workflow. A digitizing project incorporates a myriad of tasks, the successful management of which can save time and money. While a vendor may be contracted for the project, you will still need to manage an assortment of activities, including the:

- Selection of materials to be digitized
- Preparation of materials, including sorting files, removing staples and paperclips, weeding out unnecessary materials, and conservation of any deteriorating documents.
- Creation of standardized metadata
- Quality control of source materials and digital images

and chapters to reach PDFs of images of individual pages. In contrast, the Minnesota Historical Society has digitized the state’s birth records from 1900-1934 (http://people.mnhs.org/bci/). Users can search for full names, partial names, and even misspelled names by entering the information into the search fields. Entering information in additional fields, such as year of birth and county, narrows the search results.

[57] Nguyen, Binh Q. *Information Assurance Issues and Requirements for Distributed Electronic Records Archives*. Army Research Laboratory. April 2003. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA413692&Location=U2&doc=GetTRDoc.pdf

[58] This website provides resources on authentication including three white papers that introduce the issues and address methods and their associated costs. Minnesota Historical Society. *Preserving State Government Digital Information: Resources on Authentication*. December 2011. http://www.mnhs.org/preserve/records/legislativerecords/authentic.htm

- Staff training on new hardware and/or software

- Advertising, promotion, and user evaluation

- Long-term maintenance of resulting electronic files

You may also want to…

- Conduct a cost-benefit analysis to determine the cost justification of a system purchase and to determine the possible benefits to the agency with its implementation. Get upper management support. Your cost benefit analysis must include an annual expense of fifteen to twenty percent of the purchase price for training, upgrades, maintenance, and storage.

- Conduct a records and workflow analysis to determine and document existing and planned agency information needs.

- Provide specific plans for an ongoing process of migrating long-term and archival records from older to newer hardware and software platforms.

- Assign a permanent staff member as systems administrator and require the vendor to provide a project director during the installation and training periods.

When researching hardware, software, and other technologies…

- Consider data storage requirements, document scanning throughput rates, and the accurate reproduction of the image. Select systems that provide enough scanning resolution to produce a high-quality image that is at least as legible as the original record. Validate the quality of the image by testing with actual documents.

- Use an indexing database that provides for efficient retrieval, ease of use, and up-to-date information about the digital images stored in the system. Incorporate metadata to facilitate records management.

- Seek vendors who use standard rather than proprietary compression algorithms and file headers to make migrations of data more certain and reliable. If vendors use proprietary algorithms, they must be able to demonstrate their capacity to bridge to standard compressions and file headers.

When it comes time for implementation…

- Establish operational practices and provide technical and administrative documentation to ensure the future usability of the system, continued access to long-term records, and a sound foundation for assuring the system’s legal integrity. Documentation should include information on hardware and software, including brand names, version numbers, dates of installation, upgrades, replacements, and conversions; operating procedures, including methods for scanning or entering data; revising, updating, indexing, backing up; testing the readability of records; applying safeguards to prevent tampering and unauthorized access to protected information; and carrying out the disposition of original records.

- Determine which records you want to capture and manage digitally and if back-file records will be included. Review general and agency specific retention and disposal schedules, and dispose of documents that the agency is not required to retain.

- Determine if the records are adequately organized. Make certain that the records were properly filed and correct all mis-filings before imaging.

- Digitize in phases, beginning with the most highly used records first.

- Institute procedures to ensure quality and integrity of scanned images. Include visual inspection in your operational procedures to verify the completeness and accuracy of the scanning process both in the initial digitization process to magnetic media and when the image is converted to the records storage medium.

- Incorporate retention and disposal of electronic images and electronic records into agency retention schedules.

- Use non-rewritable recording media to preserve record integrity. Provide adequate environmental conditions for digital storage media. Label digital media, tapes, and other storage containers with particular care since it is impossible to determine content merely by looking at the storage medium.

- To retrieve information in records that will be held for many years, you must develop and document indexes with both today’s and tomorrow’s users in mind. Design backup procedures to create security copies of digitized images and their related index records.

- Verify that a disaster preparedness plan is in place to facilitate image and data backup, storage and recovery. For more information on disaster preparedness see the Minnesota State Archive’s online resources for Disaster Preparedness[59].

- Annually sample 3% of both the working and security copies of the digital records and indexes to make sure the data are still readable. Prepare an appropriate plan for “refreshing” data and migrating and converting images and corollary indexes to new storage media as needed.

Your implementation strategy may include setting up a pilot project. A pilot project will allow you to test the technology, examine the effectiveness of your digital images in providing and managing information for patrons or employees, and help determine how you can better

[59] Minnesota Historical Society. *Disaster Preparedness*. Minnesota State Archives. March 2003. http://www.mnhs.org/preserve/records/disaster.html

implement a digital imaging system. A pilot project is especially necessary to study the impact and effectiveness of imaging before undergoing a large digitizing project for a whole department or organization.

Phases are an effective approach to implementing large digitizing projects. Rolling out the system in phases enforces an organized and careful approach to implementation. This allows small errors to be caught and corrected before they snowball into large and costly issues. Phases can be applied in several ways depending upon the structure of your organization and scope of your project. For example, you may want to phase in the system by departments or by function. If your project will be implemented over a lengthy time period, you may want to phase in your system beginning with your organization's highest priorities.

Key Issues to Consider

Now that you are familiar with some of the basic concepts of digital imaging, you can use the questions below to discuss how those concepts relate to your agency. Pay special attention to the questions posed by the legal framework, including the need for public accessibility as appropriate, completeness, trustworthiness, and legal admissibility. Consider the resolution and delivery requirements of your digital images, and choose the file formats and digital storage media that will best fit your needs.

The goal is to determine the best option for your agency that meets your legal and operational needs, not merely to automatically upgrade technology. If you cannot justify the costs of digital imaging, keeping your records in their original form may be the best option.

Discussion Questions

- What are our goals for digital imaging?
- How is our agency affected by the legal requirements? Are the records public? Are there security requirements?
- What is the desired end product? A document management system? A searchable online collection?
- What type of materials will be digitized? Textual documents? Photographs? Maps?
- What is the required quality of the digital images? High or low resolution? Black and white or color?
- What file formats and digital storage media will best fit our needs?
- What metadata is necessary for each file? Is it readily available or do we need to spend time gathering that information?

- What are some strategies for implementing our digital imaging project?
- Are we able to do this in-house? Do we need to outsource some or all of the processes?
- How will we provide access to the digital records?
- How long do we need to keep the digital files? What is our long-term preservation plan? What do we do with the original documents?
- Can we justify the costs of digital imaging?

Annotated List of Resources

JISC Digital Media; Still Images, Moving Images and Sound Advice.
http://www.jiscdigitalmedia.ac.uk

> "JISC Digital Media is hosted at the Institute for Learning and Research Technology (ILRT) at the University of Bristol" and provides information about electronic media. The website features technical and project management advice on still images, moving images and sound media and also a glossary of terms.

Federal Agencies Digitization Initiative Still Image Working Group (FADGI). Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files. August 2010.
http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf

> Written by FADGI working group, the document represents shared best practices for digitizing materials. Updates reflect the current recommendations of the working group, specifically on the sections covering equipment and image performance metrics, quality management, and metadata.

Cornell University, Department of Conservation and Preservation. *Moving Theory Into Practice: Digital Imaging Tutorial.*
http://www.library.cornell.edu/preservation/tutorial/

> Produced by the Digital Imaging and Preservation Policy Research (DIPPR) team at Cornell University's Department of Conservation and Preservation, this web tutorial provides an overview of technical and project management issues regarding digital imaging. Tutorials in English, French and Spanish use examples of actual digital images to demonstrate variations in image quality.

Sitts, Maxine K. *Handbook for Digital Projects: A Management Tool for Preservation and Access*. Andover, Massachusetts: Northeast Document Conservation Center, 2000.
http://www.nedcc.org/resources/digitalhandbook/dman.pdf

> This handbook, published by the Northeast Document Conservation Center, is geared towards librarians, archivists, and other cultural or natural resource managers. Provides a basic technical overview of digital imaging and emphasizes project management, cost justification, vendor relations, and related issues.

Minnesota Historical Society, Minnesota State Archives. *Retrospective Digitization of Government Records*. March 2009.
http://www.mnhs.org/preserve/records/legislativerecords/digitization.htm

> Digitizing government records is a process that needs to be thought out and planned in detail before undertaking. This paper summarizes major concepts that need to be considered before starting a digitization project for textual documents. Topics covered include legal requirements, cost justification, file formats, file naming guidelines, resolution requirements, metadata and indexing, and storage options.

California Digital Library. *CDL Digital File Format Recommendations: Master Production Files.* V.1. August 2011.
http://www.cdlib.org/gateways/docs/cdl_dffr.pdf

> These standards, published by the California Digital Library at the University of California, provide recommendations for image quality, file formats, and storage media for digital image collections.

Electronic Document Management Systems

Summary

An electronic document management system (EDMS) is a software program that manages the creation, storage and control of documents electronically. The primary function of an EDMS is to manage electronic information within an organization's workflow. A basic EDMS should include document management, workflow, text retrieval, and imaging. Not all EDMSs have records management capability. To qualify as a records management system, an EDMS must be capable of providing secure access, maintaining the context, and executing disposition instructions for all records in the system. Before implementing a system you must determine how it fits into your overall records management strategy. EDMS functionality is often integrated into Content Management (CM) systems. These systems combine additional functionality such as website management with workflow tools, standard templates and access rights.

Legal Framework

If you choose to use an EDMS, your selection requires a careful, considered balance between your legal requirements and your technological options. Use of an EDMS is not a panacea for implementing your electronic records management strategy. You should not assume that the requirements for a government agency are built into an EDMS. In fact, the use of an EDMS can lead to records management problems, especially for government agencies with specific legal requirements. For example, an EDMS may improve collaboration during document development. However, the EDMS also may create multiple copies of a document and may not provide the access security you need to protect not-public records as defined by the Minnesota Government Data Practices Act (MGDPA)[60]. The decision to use an EDMS requires significant planning and analysis.

Examine the advantages offered by an EDMS in light of your legal requirements as a government agency. For more information on applicable rules and statutes refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[61].

Key Concepts

As you discuss the merits of an EDMS for your agency, you will need to be familiar with the following key concepts:

- Government Standards
- Document workflow integration

[60] Minnesota Statutes, Chapter 13. http://www.revisor.leg.state.mn.us/stats/13/

[61] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

- Basic functions
- Optional functions
- Basic process for selecting an EDMS

Government Standards

Government agencies are subject to government regulations and guidelines in the selection of an EDMS. Federal guidelines are set forth in the *Department of Defense 5015.2-STD*[62]*, Design Criteria Standard for Electronic Records Management Software Applications*. Bear in mind that even though an EDMS may meet all the Department of Defense guidelines, it may not meet all the requirements for the State of Minnesota and your agency. You must carefully examine if the EDMS supports:

- Adequate security for the protection of not-public records
- Adequate access to public records
- Ability to capture and manage electronic records (if your EDMS has this function) in a way that meets legal requirements for parameters such as trustworthiness, completeness, accessibility, legal admissibility, and durability
- All electronic formats included in the official definition of a government record

Each vendor's EDMS has different degrees of functionality. In an EDMS designed for the private sector, the functions available may not allow you to meet your legal requirements. For example, an EDMS designed for the private sector may be unable to:

- Manage all the required file formats that constitute government records
- Preserve the record's required metadata
- Ensure trustworthiness
- Provide adequate security of not-public information and records

Document Workflow Integration

You should look for an EDMS that will help you integrate and automate document management and records management at each point in your agency's records continuum. As discussed in the *Electronic Records Management Strategy* chapter of these guidelines, records should be managed as part of a continuum, rather than as having discrete stages in a life cycle. The right EDMS may increase the ease of this integrated management.

[62] Department of Defense. *Electronic Records Management Software Applications Design Criteria Standard*. April 25, 2007. http://www.dtic.mil/whs/directives/corres/pdf/501502std.pdf

Consider your agency's document workflow. An EDMS should support your workflow needs and enable you to capture and manage records as part of your daily work (one of the requirements for records to be accepted as evidence under the law).

To learn more about which documents are records, refer to the *Electronic Records Management Strategy* chapter of these guidelines.

Basic Functions

At a minimum, look for an EDMS that provides:

- *Security control*. This function controls which users have access to which information. Any system that you use must be able to protect not-public records as defined by the MGDPA.

- *Addition, designation, and version control*. The EDMS should allow users to add documents to the system and designate a document as an official government record. It should also automatically assign the correct version designation.

- *Metadata capture and use*. The EDMS should allow you to capture and use the metadata appropriate for your agency.

Optional Functions

You may also want an EDMS that can provide:

- *Records management*. EDMS systems do not always include the ability to perform records management functions. Those that offer records management functionality are sometimes referred to as Electronic Document and Records Management Systems (EDRMS). In addition to these systems, stand-alone records management software, referred to as Records Management Applications (RMA), are available. A records management system must be able to provide secure access, maintain the record's context within a record series, and automate the execution of disposition instructions for all records in the system. EDRMs and RMAs often require individual users to make decisions as to which documents qualify as records, thereby adding a layer of complexity to the work process. As a result, suitable training for all users is of utmost importance to a successful implementation. Federal guidelines are set forth in the *Department of Defense 5015.2-STD*[63], *Design Criteria Standard for Electronic Records Management Software Applications*. Bear in mind that even though an EDMS may meet all the Department of Defense guidelines, it may not meet all the requirements for your agency. Therefore, you must also consider any legal requirements applicable to your agency. Due to the impact on users and the additional expense associated with successful implementation, successful incorporation of records management software into your agency will require patience, ongoing management support, and consistent availability of resources.

- *Storage*. This function will allow you to store documents within the EDMS or to centrally manage your adjunct storage system.

[63] Department of Defense. *Electronic Records Management Software Applications Design Criteria Standard*. April 25, 2007. http://www.dtic.mil/whs/directives/corres/pdf/501502std.pdf

- *Free-text search*. This function allows users to search every word in an entire document or a specified group of documents. Other systems search only metadata.

- *Hypertext links*. Some EDMSs will provide hypertext links from one document to another to facilitate navigating and browsing among related documents.

- *Automatic conversion*. Some EDMSs will automatically convert one file format to another when the file is designated as a record (or at another specific point in the workflow). (For more information on conversion, refer to the *Electronic Records Management Strategy* and *Long-Term Preservation* chapters of these guidelines.)

- *Compound document management*. Some EDMSs manage compound documents better than others. Compound documents are single documents that contain multiple elements (e.g., text, photographs, video, hypertext links).

With so many developers and systems currently on the market, the list above describes only a few of the optional features that your agency may be interested in.

Basic Process for Selecting an EDMS

The following basic process for selecting, implementing, and managing an EDMS should serve as a baseline for you to develop a more specific process for your agency. The basic process includes:

- *Needs assessment*. The first stage is to work with internal stakeholders and understand your legal obligations to determine your unique needs. If you wish to use the EDMS for records management, be sure that you identify trustworthiness, completeness, accessibility, legal admissibility, and durability as needs (as discussed in the *Electronic Records Management Strategy* chapter of these guidelines). Be sure to think of not only your immediate needs, but also your long-term requirements.

- *Vendor selection*. You will need to carefully select an EDMS vendor. You may need to issue a request for proposals that sets forth your legal requirements and vendor selection criteria. You may also contact other Minnesota government agencies with similar systems. In short, you will want to gather as much information as you can about potential EDMSs as they are used in government agencies.

- *Implementation plan*. You will need to work with the vendor and internal stakeholders to develop a comprehensive implementation plan. The plan should include a:

    – Technological implementation plan that outlines how and when the system will be installed and tested

    – User implementation plan that includes training and system rollout

- *Deployment*. As detailed in your implementation plan, you will need to install and test the system, and train users.

- *Management*. As you use the system, you will need to continue to manage and refine your use of the system.

Throughout each of these stages, you will need to document the entire process, including needs assessment, implementation, management, and refinement. You will also need to document the system itself, including hardware, software, operational procedures, and security measures to ensure records in the system remain trustworthy over time.

Key Issues to Consider

You should consider your operational and records management requirements, including the legal framework you operate in as a government agency, as well as your desired product features and agency-specific workflow in order to select an appropriate EDMS.   To help with these decisions, form a team that includes representatives from your agency's upper management, information technology group, records management team, and legal department as well as users and content creators.

Use the questions below to consider whether to pursue an EDMS, as well as how to select a vendor. Take a long-term approach in discussing these questions. Consider the types of documents and records you create now and which types you may create in the future. Remember to think of your records as needing to be managed along a continuum, rather than in discrete stages.

Discussion Questions

- What are our current and future needs? What are the current and future needs of all involved stakeholders?
- Do we want to use the EDMS just for workflow management or do we want to use it for records management as well?
- Which records do we want to capture and manage using our EDMS? Will back-file records be included? Review general and agency specific retention schedules and dispose of documents the agency is not required to retain.
- Which formats do we use now and which formats are we likely to use in the future?
- What metadata do we need to include? Who will manage it?
- How does the legal framework affect our discussion and decision? Think about how document acceptability issues affect future interaction with the legal community.
- How do we use records now? How will we use records in the future? What records do we need to share and store?

- Are the records are adequately organized and indexed to facilitate retrieval? Ensure that the records are filed properly and correct all mis-filings before system implementation.

- How do our records fit into our current workflow? How may we need to modify our workflow to accommodate an EDMS? At which points in our workflow do we need to capture records? Consider how automation adds value to your current process.

- How will we dispose of records in the EDMS? Will the system enable us to transfer, convert, and/or migrate records easily?

- What are the roles and responsibilities of groups and individuals in terms of electronic records management?

- What features are essential to us in a document management system? What features might be the most useful, but nonessential, elements of a document management system? What is our budget?

- How will we mesh a new system with systems currently in place (e.g., e-mail systems, databases, word processing systems)?

Annotated List of Resources

Primary Resources

Association for Information and Image Management International. *Analysis, Selection, and Implementation of Electronic Document Management Systems (EDMS)*. Silver Spring, Md.: Association for Information and Image Management International, June 2009. http://www.aiim.org/documents/standards/ARP1-2009.pdf

> As recommended practice, the information in this document provides readers with a set of procedures and activities to be considered and/or practiced during the analysis, selection, and implementation of a document management system.

Association for Information and Image Management International. *Implementation Guidelines and Standards Associated with Web-based Document Management Technologies*. Silver Spring, Md.: Association for Information and Image Management International, 2002. http://www.project-consult.net/Files/AIIM+ARP1+2002.pdf

> This document contains a set of recommended practices for the implementation of selected web-based document management technologies. The document provides specific recommended activities for each phase of implementing such technologies.

National Archives and Records Administration (NARA). *Recommended Practice: Developing and Implementing an Enterprise-wide Electronic Records (ERM) Proof of Concept Pilot*. March 2006.
http://www.archives.gov/records-mgmt/policy/pilot-guidance.html

> Information on how to plan, develop, and evaluate a proof of concept pilot application as a method for exploring new technologies.

National Archives and Records Administration (NARA). *Guidance for Building an Effective Enterprise-wide Electronic Records Management (ERM) Governance Structure.* December 2005.
http://www.archives.gov/records-mgmt/policy/governance-guidance.html

> “This document defines governance and its importance to the success of IT projects, the purpose and function of that governance, how project-specific governance (such as those instituted for enterprise-wide ERM) fits within and alongside other established governance structures and the risks attendant in the absence of good governance.”

National Archives and Records Administration (NARA*). Recommended Practice: Evaluating Commercial Off-the-Shelf (COTS) Electronic Records Management (ERM) Applications.* November 2005.

http://www.archives.gov/records-mgmt/policy/cots-eval-guidance.html

> To be used as a case study as agencies examine system requirements, this document summarizes the "Environmental Protection Agency's (EPA) experience identifying the COTS products that would best meet the needs of agency staff for both Electronic Document Management (EDM) and Electronic Records Management (ERM) functionality".

Additional Resources

AIIM International
http://www.aiim.org/

> This web site is published by the Association for Information and Image Management (AIIM). AIIM is an international professional organization for "users and suppliers of the content, document and process management technologies that drive e-business." The site includes information about events, articles, industry studies, and white papers. The web site also includes a products and services vendor directory.

ARMA International
http://www.arma.org/

> Published by ARMA International, this site focuses on strategic information management issues for records and information managers, information technology professionals, imaging specialists, archivists, librarians, and others. The site includes a buyer's guide and virtual trade show of industry vendors, as well as publications, a bookstore, white papers, industry news, legislative updates, and information on industry standards.

Records Management Application Compliance Testing
http://jitc.fhu.disa.mil/recmgt/

> This site lists vendors with EDMS products that have been tested and approved by the federal government. The site provides links to the vendor's web sites. The site also provides access to a number of federal guidelines for records management, including the DOD Standard 5015.2 Design Criteria Standard for Electronic Records Management Software Applications.

Email Management

Summary

An electronic mail message or “email” consists of a digitally created, transmitted, and stored message and any attached digital documents. State and local governments use email for a variety of tasks such as sending and receiving internal and external correspondence, distributing memos, circulating drafts, disseminating directives, transferring official documents, and supporting various business processes of the organization. As such, email messages are potentially official government records, and as both state statutes and case laws make clear, email must be included in your overall records management strategy.

Email documents that hold information about the day-to-day operations of state and local government must be easy to locate; those that hold information of long-term or permanent value must be adequately protected; and those with transitory value must be deleted when no longer needed. Allowing email to be managed by personal preference or routine system back-ups and administrative procedures that treat all email alike can result in serious legal, operational, and public relations risk. By establishing policies, applying records management procedures, and training users, you can create an environment that promotes successful management of email records.

Legal Framework

For more information on the legal framework you must consider when developing an email records management policy refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives’ *Preserving and Disposing of Government Records*[64]. Also review the requirements of the:

- *Official Records Act* [Minnesota Statutes, Chapter 15.17[65]] which:
  - – Mandates that government agencies must keep records to fulfill the obligations of accountability and stipulates that the medium must enable the records to be permanent.
  - – Stipulates that you can copy a record and that the copy, if trustworthy, will be legally admissible in court. This stipulation means that you can copy your email messages to paper or to text files, as long as the record’s content, context, and structure are intact.
  - – Does *not* differentiate among media. The *content* of the email message determines whether the message is a record.
- *Records Management Act* [Minnesota Statutes, Chapter 138.17[66]] which establishes the

[64] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

[65] Minnesota Office of the Revisor of Statutes. *2008 Minnesota Statutes: Chapter 15.17 Official Records*. https://www.revisor.leg.state.mn.us/statutes/?id=15.17

[66] Minnesota Office of the Revisor of Statutes. *2008 Minnesota Statutes: Chapter 138.17 Government Records; Administration*. https://www.revisor.leg.state.mn.us/statutes/?id=138.17

Records Disposition Panel to oversee the orderly disposition of records, including email records, using approved records retention schedules.

- *Minnesota Government Data Practices Act (MGDPA)* [Minnesota Statutes, Chapter 13[67]] which mandates that government records should be accessible to the public unless categorized as not-public by the state legislature. Managing access to public versus not-public email records is especially important because email is so easily forwarded, misdirected, and sent to groups of people.

- *Uniform Electronic Transactions Act (UETA)* [Minnesota Statutes, Chapter 325L[68]] and Electronic Signatures in Global and National Commerce (E-Sign)[69] [a federal law]. Both UETA and E-Sign address the issues of the legal admissibility of electronic records created in a trustworthy manner and the application of the paper-oriented legal system to electronic records.

Additional Legal Considerations

Within the context of the above laws, you should also consider:

- *The ramifications of the Armstrong litigation*. One of the first legal cases utilizing e-discovery, was Armstrong v. Executive Office of the President (1 F.3d 1274 [DC Cir 1993]). In this case a federal court found in favor of a group of researchers and nonprofit organizations who wanted to prevent the destruction of email records created during the Reagan administration. The court determined that federal government agency email messages, depending on content, *are* public records and that complete metadata must be captured and retained with the email record. Although a federal decision, this litigation strongly influenced government agencies at all levels, and agencies began paying closer attention to their email records management practices, including the capture of metadata.

- *The ramifications of Zubulake v. UBS Warburg LLC*. The 2003 case of Zubulake v. UBS Warburg LLC has also been a major influence on the courts when determining what electronic records need to be produced during litigation. The five decisions from this case help provide a baseline standard of what needs to be available for litigation purposes; including ensuring that all relevant documents are able to be discovered, retained, and produced when necessary.[70] The Zubulake decisions also prompted the idea of a 'litigation hold' on electronic records. This hold ensures that documents, if they relate to current or future litigation, must be retained as long as necessary, including past their retention period if it has already past.

- *Legal discovery*. When developing your electronic records policy, balance your legal and

[67] Minnesota Office of the Revisor of Statutes. *2008 Minnesota Statutes: Chapter 13: Government Data Practices*. https://www.revisor.leg.state.mn.us/statutes/?id=13
[68] Minnesota Office of the Revisor of Statutes. *2008 Minnesota Statutes: Chapter 325L: Uniform Electronic Transaction Act*. https://www.revisor.leg.state.mn.us/statutes/?id=325L
[69] Thomas. *Search Results for Bill Number S.761 for the 106th Congress*. Library of Congress. http://thomas.loc.gov/cgi-bin/query/z?c106:S.761:
[70] Kroll OnTrack. *Zubulake v. USB Warburg*. http://www.krollontrack.co.uk/zubulake/

operational requirements with the risk of being engaged in legal discovery. You must meet all government requirements for managing your email records, but you should also be able to respond to discovery in an affordable, efficient, and practical way. In 2006, the Federal Rules of Civil Procedure (FRCP)[71] were amended to specifically address discovery issues for "electronically stored information", including email.

Key Concepts

As you develop your email records management policy, you will need to be familiar with the following key concepts:

- What is Electronic Mail?
- Other Electronic Messaging Systems
- Determining Value of Electronic Messages
- Retention and Disposition of Electronic Messages
- Managing Documents and Metadata
- Developing an Email Policy
- Training for Staff Members
- Processes for Preserving Email

What is Electronic Mail?

Email can be a confusing term because it can refer to both a system and the messages within a system. Furthermore, it can also be used to describe the action of sending or receiving a message. Here are some basic definitions to help clarify the process:

- Email *Systems/Clients*: Email systems or clients are the applications that enable users to compose, transmit, receive and manage text and/or graphic email messages and images across networks and through gateways connecting the latter with the Internet. Applications may be text or graphics-based, proprietary or open-source, public or private. A common email application is Microsoft Outlook, which many organizations sue as a front-end to Microsoft's Exchange Server. Some organizations are moving to vendor-hosted solutions such as Gmail for Business, to provide online staff access to email and as a cost-saving

[71] U.S. Courts. *Amendments to the Federal Rules of Civil Procedure*. http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/EDiscovery_w_Notes.pdf

measure.

- Email *Messages*: The communication supported by email systems sent between individuals or groups of individuals. The contents of the communication, the transactional information (metadata) associated with each message, and any attachments to the body of the message are all part of an email message.

- Email *Server*: The hardware on which the application resides and is used to route and store large volumes of email. Servers can be housed internally in your organization and managed by IT staff or housed and managed by others at a separate facility. If managed by an outside agency, proper security measures must be taken to ensure record authenticity and trustworthiness.

- *Transactional information*: Transactional information records the information, or metadata, about an email message. This metadata can include the name of the sender and all recipients, the date and time the message was created and sent. It may also include information on the systems and software applications used to create and transmit the message. Transactional information may not always be visible in every application but it is a vital part of every message and steps must be taken to preserve it. The federal courts have ruled that this information is a vital part of the message itself[72], and is an important consideration when storing email messages.

Other Electronic Messing Systems

In addition to email, there are other electronic messaging systems available to most organizations; they include voicemail, instant messaging, and text messaging. If used for official government purposes, a message created and managed in these systems may also be considered a record.

Therefore, when developing an email management program, organizations should also review all messaging systems in use and include any records covered under the existing records retention and disposition process. Work with your records management staff to develop new schedules where needed. Refer to *Preserving and Disposing of Government Records*[73] for more information on retention schedules.

Voicemail

Voicemail is a computerized system for receiving, recording, saving, and managing voice messages. Although telephone-based voicemail is well-established in many organizations and used for important public business, it has rarely been managed as a record.

Technological advances that blur the distinction between email and voicemail could make it easier to capture and manage audio records. Services utilizing Voice-over-Internet-Protocol (VoIP) are capable of delivering messages as audio files via email. Therefore, voicemail messages saved as email can be managed along with other email relating to the same topic.

[72] Armstrong v. Executive Office of the President, 1 F3d 127 (D.D.C. 1993)

However, VoIP uses the Internet, and messages sent this way are subject to the same security threats as other Internet communication methods.[74]

Instant Messaging

Instant messaging (IM) is a service that permits individuals to quickly exchange electronic messages with selected others in an informal manner that mimics conversation. Instant messaging differs from ordinary email in the immediacy of the message exchange that makes a continued exchange simpler than sending email back and forth. Most exchanges are text-only, however, some services allow voice messaging and file sharing. There are systems available to help organize, preserve, and provide access to Instant Messages over time. They often treat IMs the same as email communication.

Text Messaging

Text messaging, or texting, is another communication tool frequently being used. Text messaging uses mobile technology to send short bits of communication from one location to another. At times, these messages replace email or phone communications. If used for official government business, these types of messages should be considered when addressing your records management policy. Because space is often limited, users have developed a texting shorthand that uses abbreviations and code words for common phrases. If necessary, text messaging should be addressed in a general policy manual including when it can be used, and how to document the messages long-term if necessary.

Determining the Value of Electronic Messages

Not all email requires the same level of control. Although identification of email records relating to the activities of public organizations will always be subjective, certain categories of records will typically be important to identify and manage.

These include:

- *Policies and directives*
- *Work schedules and assignments*
- *Drafts of documents circulated for approval or comment*
- *Any document that initiates, authorizes, or completes a business transaction*
- *Final reports or recommendations*
- *Correspondence, memos, or messages about agency business*

[73] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf
[74] Federal Communications Commission. *Voice Over Internet Protocol*. February 1, 2010. http://www.fcc.gov/voip/

- *Agendas and minutes of meetings*

Transitory and personal messages that do not support government business, as well as convenience or duplicate copies of email records, should be deleted from mailboxes when no longer needed. These include copies or extracts of documents distributed or received for reference — listserv or bulletin board posts, personal messages, announcements unrelated to official business, and announcements of social events like retirement parties or holiday celebrations. These materials consume disk space, erode the efficiency of the system, and, if kept, could be subject to information requests and discoverable in legal proceedings.

Retention and Disposition of Electronic Messages

When thinking about email records and retention schedules, it is important to remember that retention periods are based not on the method by which a record is created but rather on the content, and on the legal, fiscal, administrative, or historical value of the information contained in the record.

You must prepare record retention schedules in accordance with the Records Management Act to retain or dispose of all official copies of email records relating to government business. Generally, records transmitted through email systems will have the same retention periods as similar records in other formats. Email letters and memos, for example, will be retained and disposed of according to the retention periods established for various types of correspondence. Many email messages will be part of a distinct record series. Those messages should be retained and disposed of according to the retention period established for that series.

The Minnesota State Archives has gathered together, on its website[75], several general record schedules for many types of records, for state, county, city, and township government records, including correspondence. Records having no significance beyond their initial use and when no longer needed for reference, should be destroyed according to an approved retention schedule. Keep in mind that simply deleting a message may not remove it completely from the storage media. Utility programs and other tools are available to permanently remove electronic messages and eliminate the possibility of recovery. Retention schedules must be approved by the Records Disposition Panel.

[75] Minnesota Historical Society. *General Records Retention Schedules Available Online*. Minnesota State Archives. August 24, 2009. http://www.mnhs.org/preserve/records/retentionsched.html

Managing Documents and Metadata

Although your agency will develop unique procedures that meet your specific operational and legal requirements, bear in mind the following goals for an email record. An email record should be:

- *Complete:* Email records should completely document the transaction. For example, you cannot save the text without the sender information. Complete email records must include all of the following elements, as applicable:

    - Names of both the Sender and Recipient(s)
    - Date and Time Sent
    - Subject Line: Clearly describing the contents of the message (e.g., the subject line 'correction' is inadequate. 'Correction to Tourism Board Minutes 2005March15' provides a better description.
    - Body of email
    - Attachments should be included in full (not just indicated by file name).
    - If a distribution list was used to send a message, a list of the names of individuals who are part of the distribution list needs to be recorded. This provides a way to identify the actual recipients if the email record simply lists the group name in the recipient field. For example, the distribution list "HR" (a distribution list for all the individuals of the human resources department) should be documented so that each individual of the list is named.
    - Directory of email addresses and the corresponding staff member names (e.g., jado25@myorg.net is Jane Doe). This connects an email address listed in an email record to a person.

- *Accurate.* The contents of the email record should accurately reflect the transaction.

- *Accessible.* Unless otherwise classified under the MGDPA, email records must be accessible to the public. All email records, like other electronic records, should be reasonably accessible for the purposes of legal discovery.

- *Manageable.* Email records should be easy for staff members to manage as part of the daily workflow and records management practices. Because staff members will implement and use the email records management policy, procedures should be straightforward.

- *Secure.* The email record should reside in a secure system that controls access, storage, alteration, and deletion. This is particularly important in controlling access to non-public content. Email records present unique security concerns, because email messages are:

    - Easily manipulated or deleted in the system.
    - Easily captured and read by unintended persons.

- Easily forwarded and misdirected by mistake.

Developing an Email Policy

You should establish policies to guide users about appropriate email practices. Policies should answer questions about acceptable use, explain the management and retention of official copies, and discuss privacy and access issues of email. When developing a policy, you should make sure to receive input from all stakeholders, allow stakeholders to review and comment on the policy before it is finalized, test the procedures, and train staff. The policy itself should also be documented and all users should understand these policies and be able to apply them. More information about general policy topics is provided below.

Acceptable use
Written policies should be established for the use of email (and other electronic communication methods) in the same way they are established for the use of the telephone, fax machine, and postal mail.

Staff roles and responsibilities
Your policy should clearly define the roles and responsibilities that managers, network administrators, technical staff, records management staff, support staff, and users will have in the management of email. It should clearly communicate whether the sender or the receiver should save email records. The policy should guide staff members in determining which email messages are records and outline a procedure for grouping emails into records series with a records retention schedule for each series.

Management and retention
Because the Records Management Act requires custodians to protect their records and to work with the Records Disposition Panel to establish retention periods, your policy should describe how and where you will maintain the official copies of your email and provide for their management, protection, and retention for as long as they have administrative, legal, fiscal, or research value.

Filing and maintenance
Only the official copy of email records that relate to agency or local government business functions need to be filed and maintained in a recordkeeping system. Additional copies, transitory communications, and personal messages can simply be deleted from the email system when no longer needed. A policy should include procedures for organizing, storing, maintaining, accessing, and disposing of email records. Your policy should define how users are to manage their accounts including the regular removal of personal and transitory messages from their mail boxes.

Privacy
Your policy should make it clear that although you attempt to provide security, email messages sent or received are not private. They may be accessed and monitored by others, may be released to the public, and may be subject to discovery proceedings in legal actions. Because computers

can store messages at multiple locations within the system, even messages a user has deleted may be recoverable and used in a legal action.

Access
Because government email can be defined as a public record, email policies must comply with the state's MGDPA and Records Management Act. MGDPA gives the public the right to access records, but it also limits access to some information considered personal or private. Custodians of public records must make their records available for public inspection provided that the information is not exempt from disclosure.

Documentation
In addition, you should establish a procedure for documenting your email records policy. On an on-going basis, from initial development onward, document the development of your email records management policy, the policy itself, and changes to the policy. Include a description of the software and hardware in use, any training provided to staff, staff member responsibilities, and records retention schedules.

Training for Staff Members

All agency employees should be trained in using the records retention schedule to identify and classily the records they create. They should be aware of proper retention and disposition procedures and who to contact when records need to be transferred out of their custody. Because individuals have direct control over the creation and distribution of email messages, agencies should provide training for their employees on agency email procedures. Depending on the type of email and recordkeeping system an agency uses, the policy and procedures will vary.

Agencies will want to be sure that employees can answer legal and operational questions about email. Any training and documentation materials should set forth guidelines that will allow employees to answer questions about the legal and operational value of their records. Possible questions that employees can ask themselves include:

- Is this email an official record? Is this email message business or personal (e.g., “Thursday staff meeting to start an hour late.” or “Let's do lunch!”)?

- Does this email message have long-term significance (e.g., “New policy finalized.”)? Does this email message document a transaction or operations function (e.g., a process, a decision, or a discussion)?

- Is this email record public or not-public as set forth by the MGDPA?

- What metadata must I capture when I save this email record?

- Which records series does this email record belong in?

- Should I save the complete email record, including attachments and group list names?

- Could this email message ever be required as evidence in a legal action?

- Who is responsible for retaining the official copy of an email?

- This flowchart[76] is used by the Kentucky Department of Library and Archives to manage email messages. A similar diagram could be developed for your institution.

Processes for Preserving Email

Government agencies have responsibility for developing guidelines and procedures to incorporate email messages into their overall recordkeeping. Agency administrators should also develop policies and systems designed to ensure that email records are appropriately preserved, secured, and made accessible throughout their established retention periods. Procedures and systems configuration will vary according to the agency's needs and particular hardware and software used.

Agency records of long-term value should not be stored on individual workstations. The records should be stored on a secure drive that has the proper security features to protect the records from alteration or destruction and to provide regular back-up. Offsite employees with laptops and other personal devices should download their messages to the agency's network drives on a regular schedule. Simply backing up the email system onto tapes or other media or purging all messages after a set amount of time is not an appropriate strategy for managing email.

There are three ways to preserve email messages: online, near-line, and off-line. Each method has its advantages and disadvantages; each requires a different degree of technical support; all require supervision and management. In making your selection, be sure that:

- it meets the needs of users;

- it complies with all recordkeeping requirements;

- you have the tools, written policies and procedures in place; and

- users understand the policies and procedures, are familiar with the tools, and can apply all three consistently to all records.

Brief descriptions of each method are listed below.

Online Storage
Online storage can maintain email messages within the email application itself. This is a good method for storing temporary and short-term records (less than 5 year retention). Microsoft Outlook does have limited capability to carry out this approach which can be employed by using the "archiving" function in the application.

[76] Kentucky Department for Libraries and Archives. *How Long Should I Keep My Emails* http://kdla.ky.gov/records/recmgmtguidance/Documents/Email training/E-mail diagram-state.PDF

Another method of online storage requires the establishment of an electronic filing process using a secure shared network server. The filing process should be used to collect and store related electronic records including, but not limited to, email. Staff should be appointed to oversee the process and system including the establishment of naming protocol and file structure as well as be responsible for assigning access privileges to the system including delegation of privileges to add, delete or edit specific files and records.

Keep in mind that email systems are not recordkeeping systems, and messages should only be stored short-term within an email system. Retaining important email within the email system disconnects it from other related information and makes it susceptible to loss through regular system purges. It will be up to you to determine the amount of risk being taken for storing files within an email application.

Advantages. You retain the ability to easily search for, retrieve, or retransmit messages electronically. You may also retain important information related to the distribution of the email. Depending on the filing arrangement used, it may be an effective way to integrate similar records that are created and received in electronic form.

Disadvantages. The process requires active participation of all email users. If not consistently and accurately managed, records are difficult to locate. Unless all records are in electronic format, you will also have to coordinate filing systems for records in both paper and electronic formats. It requires the use of a separate secure shared drive controlled by a limited number of employees to protect official copies from unauthorized access and prevent storage of duplicate copies.

Near-line Storage

Near-line storage involves the transfer of the email messages and transactional information into an electronic recordkeeping system other than the email system itself. For example, an email message dealing with a particular project could be stored in a file on the agency’s network drive with other electronic files dealing with the same project. The message still retains some of its functionality, including the ability to be indexed and retrieved electronically. If the agency stores other records in electronic format, then the email messages can be integrated with other related project files.

Disadvantages to storing records near-line are the potential costs for the equipment, maintenance and service for the electronic recordkeeping system. The agency should consider the costs and benefits, and the compatibility of their email application and the electronic recordkeeping system. Storing messages external to the email application may mean converting the messages to a different format, which could result in the loss of important information. Records with retention periods of more than 5 years need to be migrated and possibly converted to new formats and systems as older ones become outdated. Finally, if the agency still maintains many of its records in paper, then the two systems (paper filing system and the electronic system) must be integrated and work together.

Off-line (Paper) Systems

In some cases, especially for permanent and long-term records, the best preservation solution may be to print the email messages, and transactional information, onto paper. This solution makes sense if the agency does not already have an electronic system in place that is designed for long-term records protection and accessibility or if a majority of its records are kept in paper form.

The biggest advantage to off-line storage is the stability of the medium. Agencies do not have to worry about hardware and software becoming outdated and the records becoming irretrievable. Email messages can be filed with other records of the same type or series directly, making the retention and disposition process easier.

The disadvantage is that the email messages lose their dynamic functionality as electronic documents. They cannot be searched and retrieved as quickly and efficiently as in a well-managed electronic system. You may also lose important information related to the distribution of the email. Furthermore, documents can be misfiled when users are responsible for printing, routing, and filing their own. Finally, with the pervasive use of email applications in the course of government business, the volume of paper records will build up quickly.

Note: No matter what storage option the agency chooses, transactional metadata *must* be properly captured and stored with the email message for the full value of the document to be preserved. This task is usually easy in email applications that readily display this information. Applications that do not display the metadata need to be configured so that the data stays with the message in whatever form the message is retained.

Keep in mind that access to email is no longer tied to an individual work computer. Email can generally be accessed from anywhere with an Internet connection. Wireless Internet connections and mobile networks make it possible to send and receive messages from almost any location. If official business is taking place in multiple systems, including through personal accounts and devices you must be sure to have methods of capturing all relevant transactions.

Key Issues to Consider

Now that you are familiar with the operational and legal importance of managing email messages as records, you can use the questions below to begin the development of your email management policy. Discussion of the questions below will help:

- Ensure that you meet your legal and operational requirements
- Gather staff member input, support, and compliance with your email management policy
- Integrate your records management policy with your overall electronic records management strategy

- Ensure that staff members manage email records at the appropriate points in the records continuum, rather than as a single records series with one retention schedule (as explained in the *Electronic Records Management Strategy* chapter of these guidelines)

Discussion Questions

- How can we ensure staff member compliance and understanding? What process is reasonable to ask staff members to comply with?
- How should we train staff members? How accountable should we make staff members for compliance?
- How should we develop our process?
- Which email messages are official records?
- What elements of an email record are required for a complete understanding of the transaction?
- What is the appropriate records series and records retention schedule for each records series? How should email records be organized for long-term storage and access (e.g., project, department, function)? How will we retrieve and dispose of email on our chosen storage media?
- How should our email retention strategy coordinate with our other records management procedures (e.g., store all project-related email with the other project documentation)? What documentation do we need for our process?
- How should we implement the procedures technically and operationally? How can we plan our implementation so the policy is widely used and accepted, but causes minimal disruption to our daily operation?
- How will this all be documented?

Annotated List of Resources

Prom, Christopher. *Preserving Email.* Digital Preservation Coalition Technology Watch Report 11-01. December 2011.
http://www.dpconline.org/newsroom/latest-news/805-email-tomorrow-and-next-year-and-forever-preserving-email-report-published

This report provides practical advice on how to ensure email remains accessible over time and address the technical, legal and cultural challenges of email preservation.

Council of State Archivists (CoSA). *Electronic Mail Policies and Management*. ARC Resource Center. January 6, 2010
http://www.statearchivists.org/arc/states/res_emai.htm

Resources specifically about electronic mail policies compiled by state.

ARMA International. *Home Page.*
http://www.arma.org/

"ARMA International is the oldest and largest association for the records and information management profession." This website provides users with access to publications and information about electronic records management, standards and best practices, professional development, and upcoming conferences and seminars. Local chapters of ARMA have been developed in many metropolitan cities, of which a list can be found on this ARMA International site.

The Sedona Conference. *Publications Home Page.*
https://thesedonaconference.org//publications

Publications address issues of eDiscovery, Email management, and other topics. (Note: minimal registration is required to download the free publications.)

K&L Gates. *Electronic Discovery Law Blog.* September 10, 2009.
http://www.ediscoverylaw.com/

A blog that is frequently updated discussing court decisions, new or modified laws, and other recent events relating to electronic discovery.

AIIM Knowledge Center Blog. "*What is Email Management*?" by Atle Skjekkeland. February 13, 2009.
http://aiimknowledgecenter.typepad.com/weblog/2009/02/what-is-email-management.html

> An online presentation on email management that covers many issues surrounding email management including business drivers for using email, concerns about privacy, legal issues and security issues, policy development, and email management technologies.

Web Content Management

Summary

The impact of technology on government not only affects how government agencies complete tasks internally, it also influences the way those agencies interact with the public at large. The popularity of the Internet has resulted in government agencies growing increasingly reliant on websites to meet the information needs of citizens. As a result, agencies need to manage their web content effectively to ensure that citizens are able to find the information they want easily and are able to determine if it is accurate and current.

Web content management makes government accountable. Because websites may contain records that document government activity and the use of tax dollars, just as any paper record does, government agencies must manage web content with a carefully developed and implemented policy. Therefore, each agency should develop a plan for the management of public records maintained on its website. The plan should integrate into each agency's overall records management program.

Legal Framework

For more information on the legal framework you must consider when developing a web content management strategy refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[77]. Particularly the specifics of the:

- Official Records Act [Minnesota Statutes, Chapter 15.17[78]] which mandates that government agencies must keep records to maintain their accountability. Agencies using the web for business should have a records management plan that explicitly addresses proper handling of records posted online.

- Records Management Act [Minnesota Statutes, Chapter 138.17[79]] which indicates, like other records, your website records must be maintained according to established records retention schedules.

Additional legal considerations include:

- *Public versus not-public*. Agencies must determine which website records are public and which are not-public as described in the Minnesota Government Data Practices Act

[77] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf
[78] Minnesota Statutes, *Chapter 15.17*; http://www.revisor.leg.state.mn.us/stats/15/17.html
[79] Minnesota Statutes, *Chapter 138.17;* http://www.revisor.leg.state.mn.us/stats/138/17.html

(MGDPA)[80]. For example, you may gather and store confidential data via a web interface. This data should be protected from public access as outlined in the MGDPA.

- *Record or non-record*. The State of Minnesota, as outlined in the Official Records Act, does not differentiate among the media on which records are created or stored. The content of the web file determines whether the file is a record.

- *Web Accessibility Concerns*. Official records must be accessible to all citizens. Materials on the web are not always accessible to people with disabilities. Following web design standards will help make the content available to the most users. For more information on accessibility issues please review the *Web Content Accessibility White Paper* published by the Minnesota State Archives.[81]

Key Concepts

As you develop your web content management policy, you will need to be familiar with the following concepts:

- Website records
- Involving stakeholders
- Standard practices
- Change over time
- Long-term management

Website Records

Your website may contain records as defined by the Official Records Act; you should manage these records as part of your comprehensive ongoing electronic records management strategy.

Records may include copies of publications like annual reports, directories, fact sheets, leaflets, newsletters and other serials, research reports, technical reports, and so forth; however the official copy of these records should be managed outside of the web environment as part of your overall records management strategy.

Records available on the web are not always published documents. Some records that people access on the web are not published to the web; the record content is pulled from a back-end

[80] Minnesota Office of the Revisor of Statutes. *Minnesota Statutes, Chapter 13: Government Data Practices.* 2008. http://www.revisor.leg.state.mn.us/stats/13/

[81] Hoffman, Nancy. *Web Content Accessibility White Paper*. Minnesota Historical Society / State Archives. November 2008. http://www.mnhs.org/preserve/records/legislativerecords/docs_pdfs/WebAccessibility.pdf

database and then displayed on the web as a record. This content is generated on-demand and is controlled by user selections. The information that is used to create these records is stored elsewhere and has its own retention schedule.

Involving Stakeholders

Creating and maintaining a website involves a lot of people. If a web content management strategy will be developed, the process should include all those who are involved in website creation, administration, and use. Key groups to include are content creators and experts, website technical experts, website internal users, records management staff, and agency or department heads. Each group should be familiar with the agency's policy for web content publication, removal, storage, and disposition; how the policy affects their daily work practices, including their roles and responsibilities under the policy; and the agency's general electronic records management strategy.

Initial and on-going internal communication will be a crucial component of web management procedures and policy development because:

- Many groups are involved in the creation and administration of a website.
- Much of a website's content is interrelated.
- Website content tends to change frequently.

Consider establishing a formal mechanism to keep stakeholders informed of each other's activities related to the website. This communication allows your agency to control the content and trustworthiness your website records, since all stakeholders will know when and why content changes.

Standard Practices

Managing web content is an ongoing task. Websites are used to inform the public on recent and upcoming events as well as provide access large amounts of information from an agency. In general, a good website should follow basic web practices including but not limited to:

- *Use web design standards*. The W3C is a good source for web design standards on HTML a & CSS, scripting and Ajax, graphics, audio and video, accessibility, internationalization, privacy, the mobile web and even math on the web[82] as well as on general architecture principles, identifiers (names and addresses), protocols, and formats.[83] Using standards increases accessibility for all users; Section 508 of the Federal

[82] World Wide Web Consortium (W3C). *Web Design and Applications*. 2010. http://www.w3.org/standards/webdesign/

[83] World Wide Web Consortium (W3C). *Web Architecture*. 2010. http://www.w3.org/standards/webarch/

Rehabilitation Act[84] also provides information on accessibility concerns.

- *Version control*. Because websites are updated constantly by different individuals and groups, you should develop a method for designating and controlling versions. Make sure the content is kept up to date. Including the date last modified on the website will inform users how recent the content is.

- *File Naming*. Consider establishing a file naming protocol for web pages to help ensure ease of management, usability of the site, and internal communication about contents. For more information about website file naming, refer to the *File Naming* chapter of these guidelines.

- *Metadata*. Use a standard metadata set to define and describe web content. Dublin Core is often used. This will help search engines find and rank your web pages.

- *Organization*. Make information easy to find and read. Use clear navigation methods including breadcrumb trails and avoid using flashing graphics and colors that are hard on the eyes.

- Understand the effects of any changes you make to your webpage. For example, moving or removing a page may result in broken links or page not found error messages. Consider adding a redirect page or using persistent identifiers to help users find relocated information, as broken links may reduce site creditability.

Change Over Time

Due to the ever changing nature of websites, continuous management of these records is required. Web pages, documents, or files may need to be removed from a website or relocated to another location for a number of reasons including:

- The information or publication no longer reflects your agency's current policy or has been superseded.

- The retention requirements of the publication have been met; the official copy has been disposed of and the online copy needs to be as well.

- The publication is perceived as no longer having value.

- Files are moved due to site reorganization.

If you do remove or move documents consider that fact that what the agency views as no longer important or relevant, others may still find useful or valuable for continuing reference or research or for historical interest. Also consider that others may have referenced or linked to individual

[84] Section 508. *508 Law*. April 30, 2008. https://www.section508.gov/index.cfm?FuseAction=Content&ID=3

resources on your website on other websites, publications, catalogs, or printed reports. If it is known that the records you want to remove are highly referenced, consider moving it to another location and using a redirect page or a persistent identifier to help users find relocated information.

Long-Term Management

As more and more records are accessible on the web and more and more records are being created directly on the web, it is important to monitor and evaluate the information that is available on the web. On one hand you must manage the day to day website and manage the content people have access to. Is the information up-to-date, accurate, and trustworthy? Do the links work? On the other hand, the web has become a place to create records and it is also important to evaluate what content is being created on the website and if that information constitutes a record as defined by the Official Record Act.

If they are records, general records management issues come into play and the records must remain accessible for as long as required. As the web is ever changing, one way to preserve websites is to capture snapshots in time of individual pages, sections, or entire websites. These snapshots, if taken at regular time intervals will document changes overtime as well as keep original web records available. This process is often called web archiving. The parameters for archiving a website through available software or contracted services can be set up to include as much as or as little of a site as desired.

As your agency expands the role of its websites to conduct agency business, it may become important for accountability purposes to document your entire website as a record rather than individual pages. Available software programs and contracted services enable you to reconstruct your entire site. In addition to capturing official website content, capturing all short-term projects with web content may also be done. For example, an agency that sets up a short-term website for a legislative initiative that includes a bulletin board for key people to discuss an initiative should, for public records purposes, take website snapshots that will allow reconstruction of the site completely as it existed at a given time. The frequency of capturing these snapshots will depend on how often web content changes.

During snapshot capture, metadata is also captured. This metadata may be basic information about the date and time of the capture, lists of the pages captures, and time the captures took or some archiving services also capture information about why, who, and what the captured content contains. Using metadata standards during web design will add to the data available during capture.

As part of ensuring that you capture enough information in a record to demonstrate the record's content, context, and structure, you will need to capture metadata. Many Minnesota government agencies have elected to use the Dublin Core Metadata Element Set as a standard (NISO Standard Z39.85; ISO Standard 15836). Some of the basic Dublin Core metadata elements include title, subject, description, creator, and date. (For more information, refer to the *Metadata* chapter of these guidelines).

Key Issues to Consider

Now that you are familiar with some of the basic concepts of web content management, you can use the questions below to discuss how those concepts relate to your agency.

You will want to discuss changes to content, organization, and administration over time; the preservation of web materials; determining who is responsible for updates and other tasks, and building staff awareness of policies and procedures. Pay special attention to the questions posed by the legal framework, including the need for public accountability, managing public and not-public records, and following records retention schedules.

You should:

- Examine your current use of the web and understand your expectations for future use. For example, you may currently publish a newsletter in paper format, but in the future, you may publish the same newsletter on the web.
- Understand the transactions that are completed online as well as the communication (e.g., bulletin boards, live chats, posted e-mails) that takes place via the website that may become records.
- Understand how to preserve website content over time and how to build this into the overall management of the website.
- Develop a plan on how to build staff awareness and compliance with web content policies, including establishing procedures to maintain and update content by authorized individuals only.
- Review pages regularly for quality, accuracy, and timeliness.

Technical considerations include:

- The site should be accessible to the most common browsers.
- Backups should be done regularly to a secondary medium. (This is not equivalent to archival preservation of the site.)
- When possible, use standard formats that are open source and non-proprietary.
- Maintain a site index or site map.

Discussion Questions

- What information will citizens seek on our website? How can we ensure that we make the information easy to find? How can we assure those seeking information of the trustworthiness of the information?

- Which elements of our website are records? Where should the official copies be stored? Are the record series included on an approved retention schedule?

- Do we need to do periodic website snapshots? How long should snapshots be kept? How can we build web content and snapshot archiving into overall website management?

- How can we build staff awareness and compliance with web content archiving procedures?

- Who will authorize website content reorganization and removal?

Annotated List of Resources

W3C. (World Wide Web Consortium). Home Page.
http://www.w3.org/

> W3C is an organization that focuses on standards for web design, applications, and architecture. Published standards and general information are provided here.

Minnesota Historical Society. *Best Practices Principles for Opening Up Government Information.* March 2011.
http://www.mnhs.org/preserve/records/legislativerecords/docs_pdfs/BestPracticePrinciplesOpenGovtMarch2011_000.pdf

> Created with assistance from the Sunlight Foundation, this paper lists best practice principles that help facilitated open government, transparency, and accessibility.

Minnesota Historical Society. *Web Archiving.* Center for Archival Resources On Legislatures (CAROL). March 2012.
http://www.mnhs.org/preserve/records/legislativerecords/carol/webarchiving.htm

> Links to an introductory paper on web archiving as well as an evaluation/comparison of Archive-It and the Web Archiving Service.

Minnesota Department of Natural Resources. *Bridges: Minnesota's Gateway to Environmental Information*.
http://www.bridges.state.mn.us

> The Bridges project was a collaboration between Minnesota's environmental agencies with the goal of providing easy access to their electronic resources such as web pages, PDF documents, databases, and geographic data. Resources were cataloged using the Dublin Core metadata scheme and made available through a simple cross-agency search engine. Although the project was completed in July 2000, the website still offers a number of resources to visitors, including best practice guidelines for web metadata, information on metadata tools, project reports, as well as links to participating agencies, other regional and federal environmental sites, and the Minnesota Governor's Council on Geographic Information.

Kansas Information Technology Advisory Board, Electronic Records Committee and Internet Task Force. *Guidelines for Managing Records on Kansas Government Agency Web Sites*. Version 1.0, January 2004.
http://www.kshs.org/government/records/electronic/web_guidelines_approved_version1_0.pdf

The Kansas guidelines are based upon a risk analysis methodology. A summary of web-based resource types and a discussion of preservation strategies accompany a set of agency self-assessment tools.

National Archives Records Administration (NARA). *NARA Guidance on Managing Web Records*. January 2005.
http://www.archives.gov/records-mgmt/policy/managing-web-records-index.html

Guidance on the background, responsibilities, and requirements of managing web records based on statutory requirements for developing disposition schedules.

Electronic and Digital Signatures

Summary

The advent of e-government and e-services has changed the way state agencies and local government offices do business. As a result, electronic systems and processes have become as important as traditional paper and ink. In a paper environment, a hand signature, also known as a "wet signature," authorizes and authenticates the content of a document. A signature provides a level of trustworthiness and accountability that aids the conduct of business. Electronic signatures extend the function of handwritten signatures to electronic documents, providing a way for two parties to conduct business confidently in an electronic environment. Up-to-date technologies and procedures must meet the demand for trustworthiness where hand signatures are not viable.

Since signatures derive their primary importance from their legal and evidentiary value, these concerns must drive the selection of electronic signature technologies. Consequently, each agency will need to define its legal and evidentiary needs in relation to its business processes before choosing an electronic signature application.

Furthermore, the electronic signature application selected must fit the agency's technology architecture to create, preserve, and make available its records. Technical obstacles pose great challenges to the long-term preservation of electronic signatures. Policy regarding the preservation of signatures should be adopted by each agency to ensure consistent practice across the organization.

Legal Framework

Many government agencies have unique and specific legislative mandates that apply to them and their functions. Two chapters of the Minnesota statutes in particular apply to electronic signatures, Chapters 325L and 325K.

- Uniform Electronic Transactions Act (UETA) [Minnesota Statutes, Chapter 325L[85]] addresses the issues of the legal admissibility of electronic records created in a trustworthy manner and the application of the paper-oriented legal system to electronic records.
- The Minnesota Electronic Authentication Act [Minnesota Statutes, Chapter 325K[86]] defines an electronic signature uniquely in terms of digital signature using Public Key Infrastructure technology (PKI). This type of digital signature is: a transformation of a message using an asymmetric cryptosystem such that a person receiving the initial message and having the signer's public key can accurately determine: (1) whether the transformation was created

[85] Minnesota Office of the Revisor of Statutes. *2009 Minnesota Statutes. Chapter 325L: Uniform Electronic Transactions Act*. 2009. https://www.revisor.leg.state.mn.us/statutes/?id=325L
[86] Minnesota Office of the Revisor of Statutes. *2009 Minnesota Statutes. Chapter 325K: Electronic Authentication*. 2009. https://www.revisor.leg.state.mn.us/statutes/?id=325K

using the private key that corresponds to the signer's public key; and (2) whether the initial message has been altered since the transformation was made.

Each agency should their specific statutory requirements before making any choices about electronic signature technologies.

In addition to state laws, agencies must adhere to federal laws such as:

- Electronic Signatures in Global and National Commerce (E-Sign)[87], a federal law that addresses the issues of the legal admissibility of electronic records created in a trustworthy manner and the application of the paper-oriented legal system to electronic records. (Federal version of UETA.)

- Health Insurance Portability and Accountability Act of 1996, HIPAA[88]. This act is concerned with non-repudiation. Non-repudiation "provides assurance of the origin or delivery of data," so that the sender cannot deny sending a message and the receiver cannot deny receiving it. This prevents either party from modifying or breaking a legal relationship unilaterally. HIPAA holds that only a digital signature technology can currently provide that assurance.

For more information on the legal issues you must consider when considering using electronic signature technology, including what constitutes a government record, refer to the *Legal Framework* chapter of these guidelines and the Minnesota State Archives' *Preserving and Disposing of Government Records*[89].

Key Concepts

When selecting and implementing an electronic signature technology, keep in mind:

- Functions of Signatures

- Definitions of Signatures

- Electronic Signature Technologies

- Other Means of Authentication

[87] Thomas. *Electronic Signatures in Global and National Commerce Act*. S.761. Library of Congress. http://thomas.loc.gov/cgi-bin/query/z?c106:S.761**:**

[88] U.S. Department of Health and Human Services. *Understanding HIPAA Privacy*. http://www.hhs.gov/ocr/privacy/hipaa/understanding/index.html

[89] Minnesota Historical Society. *Preserving and Disposing of Government Records*. Minnesota State Archives. May 2008. http://www.mnhs.org/preserve/records/docs_pdfs/PandD_may2008.pdf

Functions of Signatures

In general, signatures serve specific functions. The American Bar Association[90] enumerates these as:

- *Evidence*: A signature authenticates a writing by identifying the signer with the signed document. When the signer makes a mark in a distinctive manner, the writing becomes attributable to the signer.
- *Ceremony*: The act of signing a document calls to the signer's attention the legal significance of the signer's act, and thereby helps prevent inconsiderate engagements.
- *Approval*: In certain contexts defined by law or custom, a signature expresses the signer's approval or authorization of the writing, or the signer's intention that it have legal effect.
- *Efficiency and logistics*: A signature on a written document often imparts a sense of clarity and finality to the transaction, and may lessen the subsequent need to inquire beyond the face of a document. Negotiable instruments, for example, rely upon formal requirements, including a signature, for their ability to change hands with ease, rapidity, and minimal interruption.

An electronic signature will have to fulfill some or all of these functions. You should determine which are pertinent to your business processes before selecting a particular electronic signature technology.

Definitions of Signatures

Using Minnesota Statutes the traditional definition of a signature is as follows:

> The signature of a person, when required by law, (a) must be in the handwriting of the person or, (b) if the person is unable to write, (i) the person's mark or name written by another at the request and in the presence of the person or, (ii) by a rubber stamp facsimile of the person's actual signature, mark, or a signature of the person's name or a mark made by another and adopted for all purposes of signature by the person with a motor disability and affixed in the person's presence.[91]

A reliance on the definition above would make it virtually impossible to use technology to deliver services and to meet all legal and evidentiary requirements. To address this problem, and

[90] American Bar Association. *Digital Signature Guidelines Tutorial*. Section of Science and Technology Information Security Committee. http://www.americanbar.org/groups/science_technology.html

[91] Minnesota Office of the Revisor of Statutes. *2009 Minnesota Statutes: 645.44 Words and Phrases Defined*. 2009. https://www.revisor.leg.state.mn.us/statutes/?id=645.44

to provide a standard approach to the use of electronic signatures, Minnesota adopted the Uniform Electronic Transactions Act (UETA)[92] in the 2000 legislative session.

UETA defines electronic signatures as:

> An electronic sound, symbol, or process attached to or logically associated with a record and executed or adopted by a person with the intent to sign the record.

This definition is not technology specific, and so does not mandate the adoption of any particular hardware or software application. Any technology, theoretically, that could authenticate the signer and the signed document could generate a legally admissible signature, as long as the parties could demonstrate the trustworthiness of the process that created and preserved the records in question.

In many communities there is no distinction made between the terms ‘electronic’ and ‘digital’, especially among information technology communities where “electronic” and “digital” are used synonymously and interchangeably. However, in Minnesota law there is a clear legal distinction made between electronic and digital signatures.

A digital signature is a particular type of electronic signature that relies on a Public Key Infrastructure (PKI) technology. UETA does not separately define digital signatures but permits their use under the broader definition of electronic signatures. The Minnesota Electronic Authentication Act[93] however does define a digital signature uniquely in terms of PKI. A digital signature is:

> A transformation of a message using an asymmetric cryptosystem such that a person having the initial message and the signer's public key can accurately determine: (1) whether the transformation was created using the private key that corresponds to the signer's public key; and (2) whether the initial message has been altered since the transformation was made.

A digital signature may offer the advantage of providing a unique identifier and linking the signature to the record. It can authenticate both the signer and the signed document, thus meeting legal requirements for admissibility and trustworthiness. PKI technology offers the additional advantages of adaptability to a wide range of applications and compatibility with basic office software.

[92] Minnesota Office of the Revisor of Statutes. *2009 Minnesota Statutes: Chapter 325L: Uniform Electronic Transactions Act.*   2009. https://www.revisor.leg.state.mn.us/statutes/?id=325L
[93] Minnesota Office of the Revisor of Statutes. *2009 Minnesota Statutes: Chapter 325K: Electronic Authentication.* 2009. https://www.revisor.leg.state.mn.us/statutes/?id=325K

Electronic Signature Technologies

The Uniform Electronic Transactions Act (UETA) [Minnesota Statutes, Chapter 325L[94]] purposely allows for a wide range of signature technologies. It says, "An electronic record or electronic signature is attributable to a person if it was the act of the person. The act of the person may be shown in any manner, including a showing of the efficacy of any security procedure applied to determine the person to which the electronic record or electronic signature was attributable."

An example of this is the "click through" option used on many web sites. To order a product, be it a shareware application, an airline ticket, or a book, a web user has to "click through" a page or form that indicates approval of the vendor's conditions for the sale. The system makes it impossible to transact any business without first establishing that agreement. In this instance, there is no "signature" or anything like it. Instead, the system is designed to make it necessary to move from "A" to "C" *only* through "B," with "B" serving as the equivalent of a signature. Authentication is demonstrated by the documentation of the system and its procedures, not by a signed record of a specific, individual transaction.

UETA implicitly legitimates the use of technologies such as faxes, digital imaging, the use of PIN/passwords, and digital signatures, as well as the more exotic iris scans, for electronic signatures. In all cases, the key to demonstrating the trustworthiness of a record and its signature is demonstrating the trustworthiness of the system that creates and manages the record. Having sufficient and appropriate systems documentation is the only way to establish that the signature is authentic and reliable.

Digital signatures demand the use of a specific PKI technology. PKI systems use two different keys. One key is kept secret (the private key) and the other key is made publicly available (the public key). The two keys are generated simultaneously and collectively; they are known as a "key pair." Once a message has been signed using one of the two keys, it can only be verified by the other key. The resulting digital signature is a cryptographic checksum computed as a function of the message and the signer's private key.

Because the digital signature is generated as a function of the key and a unique message, the signature serves two purposes. It authenticates the signer, since only the individual owner has (in theory, anyway) access to the private key. It also indicates the reliability and integrity of the message, since any alteration to the text would invalidate the signature.

This is not the same as encryption. PKI technology was originally developed for encryption (as in the Pretty Good Privacy applications), but the use of a digital signature does not automatically encode a message. In fact, encryption is not covered in the Minnesota Electronic Authentication Act [Minnesota Statutes, Chapter 325K[95]], which only addresses the use of PKI for digital signatures.

[94] Minnesota Office of the Revisor of Statutes. *2009 Minnesota Statutes. Chapter 325L: Uniform Electronic Transactions Act*. 2009. https://www.revisor.leg.state.mn.us/statutes/?id=325L

[95] Minnesota Office of the Revisor of Statutes. *2009 Minnesota Statutes. Chapter 325K: Electronic Authentication*. 2009. https://www.revisor.leg.state.mn.us/statutes/?id=325K

The effective use of PKI for digital signatures relies on some policy and organizational factors. There has to be some way to guarantee and to prove that a specific person actually owns a specific key. And there has to be some way to provide quick and easy access to public keys. Because it is completely impractical for each sender and each recipient of a message to work this out on a case-by-case basis, the use of PKI for digital signatures is dependent on the operation of certificate authorities.

A certificate authority is an independent, trusted third party who issues and manages key pairs. To get a key pair, individuals must prove to a certificate authority that they are who they claim to be. The certificate authority also provides secure access to public keys that allow for the validation and verification of signatures. The Minnesota Electronic Authentication Act [Minnesota Statutes, Chapter 325K[96]] creates a mechanism to license and regulate certificate authorities.

Other Means of Authentication

In addition to electronic and digital signatures, there are other methods of authenticating digital content that may be useful to your agency. These options are thoroughly discussed in two white papers: *Authentication of Primary Legal Materials and Pricing Options* and *Authentication Methods*.[97]

Key Issues to Consider

No electronic signature technology by itself is sufficient to meet all legal needs. The evidentiary value of signed records will ultimately rely on an agency's ability to produce legally admissible documentation of your recordkeeping system. In addition, the agency will, of course, have to produce the electronic records themselves. Merely preserving and providing access to electronic records present some daunting challenges. Adding electronic signatures to the equation can complicate the situation even further.

- Hardware and software obsolescence make it difficult, if not impossible, to preserve and provide long-term or permanent access to both the electronic signature and the associated electronic record. For example, if an agency is using different technologies to create and to sign a record, those technologies might "age" at different rates. In a digital signature (PKI) system, the signature is a function of the content of the document. Due to this relationship, any migration or conversion of the document's content for preservation will nullify the original digital signature and prevent its use as a means to ensure the authenticity and reliability of that document. Therefore, agencies will need to plan for technology

[96] Minnesota Office of the Revisor of Statutes. *2009 Minnesota Statutes. Chapter 325K: Electronic Authentication*. 2009. https://www.revisor.leg.state.mn.us/statutes/?id=325K

[97] Minnesota Historical Society. "Authentication of Primary Legal Materials and Pricing Options" and "Authentication Methods". *Center for Archival Resources On Legislatures (CAROL).* 2011. http://www.mnhs.org/preserve/records/legislativerecords/carol/authentication.htm

obsolescence of both the record and the signature if long-term preservation of electronic signatures is desirable.

- Agencies should plan to document their decisions and transactions. Understanding legal needs and addressing them at the design phase of an application are important factors to making this work. Keeping documentation up-to-date is an on-going responsibility, which could be complicated if relying on a third party. For example, when using digital signatures agencies should make sure that the certificate authority is managing its records and documentation adequately.

- Agencies should make sure that the electronic signature technology is interoperable with their and their constituencies' other software applications. Requiring complex or expensive solutions is probably not practical. It would be especially difficult to ask citizens to buy and maintain multiple signature technologies.

- Agencies should assess risks associated with the use of electronic signature technology and develop a well-documented risk management plan based upon the risks identified.

- The human side of the equation is critical: no technology will completely address your legal requirements. For example, a digital signature is only as reliable as the certificate authority standing behind it as well as the ability of the users to protect personal certificate information from loss or inappropriate use.

Selecting the appropriate electronic signature technology means defining the most important criteria and then determining if the system and proposed application meet those criteria. The criteria should give priority to legal concerns, since signatures are primarily valuable for evidentiary purposes. A selection decision should also reflect consideration of other factors, such as technology architectures, costs/benefits, agency business practices, and all pertinent policies, hardware, software, controls, and audit procedures. A specific example of the criteria pertinent to a digital signature application can be found in the American Bar Association's *PKI Assessment Guidelines.*[98]

Discussion Questions

Use the following questions to help determine why you need to use electronic signatures, who will use them, what technologies are appropriate, and how other records management issues relate to electronic signatures.

- Why do you want to use electronic signatures? What business functions will the technology support?

[98] American Bar Association. *PKI Assessment Guidelines: Guidelines to Help Assess and Facilitate Interoperable Trustworthy Public Key Infrastructures*. Chicago, Ill; American Bar Association. 2003. http://openlibrary.org/b/OL12199471M/Pki_Assessment_Guidelines__Guidelines_to_Help_Assess_and_Facilitiate_Interoperable_Trustworthy_Public_Key_Infrastructures

- Who will have to use and rely on the electronic signature?

- How long will the signatures and the records to which the electronic signatures are affixed have to be preserved?

- Which state and federal statutes pertain to the functions and transactions that generate your signed records? What case law is there?

- How does the electronic signature technology fit into your overall technology architecture? What's the total cost of the technology? What's the cost per transaction?

- What sort of electronic signature technologies do your customers use? Will you have to share these records with any other organizations or agencies? What technologies do they use?

- What methodology will you use for documenting your information systems, policies, and practices?

Annotated List of Resources

Government Printing Office (GPO). Authenticity of Electronic Federal Government Publications. June 13, 2011.
http://www.gpo.gov/pdfs/authentication/authenticationwhitepaper2011.pdf

> This paper describes the tools and evidence that the GPO provides to users to help them verify that they can trust the source of the content, and that unauthorized alterations to content have not occurred. In doing such, the GPO provides evidence that the electronic information it maintains is from a trustworthy repository and the history of each item in the repository can be documented. The GPO also provides content integrity tools such as digitally signed PDF files and cryptographic hash values.

Minnesota Historical Society. *Authentication.* Center for Archival Resources on Legislatures (CAROL). March 2012.
http://www.mnhs.org/preserve/records/legislativerecords/carol/authentication.htm

> This resource includes information on the Uniform Electronic Legal Materials Act (UELMA), as well as white papers and resources that introduce authentication, explore options, and associated cost models.

The National Archives Records Administration (NARA). *Records Management Guidance for PKI Digital Signature Authenticated and Secured Transaction Records*. March 11, 2005.
http://www.archives.gov/records-mgmt/policy/pki.html

> Guidelines for the use of Public Key Infrastructure (PKI) digital signatures as authenticated and secure electronic transmissions.

The National Archives Records Administration (NARA). *Records Management Guidance for Agencies Implementing Electronic Signature Technologies*. October 18, 2000.
http://www.archives.gov/records-mgmt/faqs/pdf/electronic-signiture-technology.pdf

> Records management information for agencies concerned about ensuring the trustworthiness of their records. These guidelines address record management issues including trustworthiness, define key terms, and provide resources for further information and assistance.

National Institute of Standards and Technology (NIST), U.S. Department of Commerce. *Cryptographic Toolkit: Digital Signatures*. Washington, D.C.: NIST, 2001.
http://csrc.nist.gov/groups/ST/toolkit/index.html

NIST's web site provides access to three Federal Information Processing Standards (FIPS) standards for digital signature algorithms, along with a variety of other resources on cryptography.

Artic Soft Technologies Limited. *An Introduction to PKI (Public Key Infrastructure).* 2010. http://www.articsoft.com/public_key_infrastructure.htm

Introduces PKI, explains public and private keys used for digital signatures, certificates, storage methods for keys, certificate authorities, registration authorities, and certificate management techniques.

MBA Knowledge Base. *How Public Key Infrastructure (PKI) Works?* 2010. http://www.mbaknol.com/business-finance/how-public-key-infrastructure-pki-works/

Explains encryption, digital certificates, digital signatures, PKI, certificate authorities, and registration authorities.