



General Commission on Archives & History
The United Methodist Church

Records Management Guidelines

Guidelines for Managing Electronic Records 2009 Edition

Guidelines for Managing Electronic Records 2009-2012

General Commission on Archives and History
The United Methodist Church
P. O. Box 127
Madison, NJ 07940
<http://www.gcah.org>
gcah@gcah.org

Table of Contents

Purpose	5
Scope	5
Audience	5
First Steps	6
The Importance of Digital Records	6
What Are Digital Records?	6
Why Manage Digital Records?	7
Challenges Associated with Digital Records	8
Digital Recordkeeping Framework	9
Integrated and Comprehensive Approach	9
Policies, Procedures and Guidelines	9
Systems Design	9
Creating Digital Records	10
Identifying and Creating Digital Records	10
Capturing Digital Records into Recordkeeping Systems	10
Creating Information about Digital Records	11
What Is Metadata?	11
Recordkeeping Metadata	11
Capturing and Maintaining Metadata	11
When Should Metadata Be Captured?	12
Determining How Long to Keep Digital Records	12
How Long Do Digital Records Need to Be Retained?	12
Disposal of Digital Records	13
Normal Administrative Practice	13
Storing Digital Records	14
How Are Digital Records Stored?	14
Selecting the Appropriate Storage Method	15
Securing Digital Records	15
Why Is Security Important for Digital Records?	15
Methods of Securing Digital Records and Systems	16
Authentication of Digital Records	16
Long-term Digital Records	17
Preserving Digital Records for the Long Term	18
Why Preserve Digital Records?	18
Planning for Technological Obsolescence	18
Techniques for Digital Records Preservation	20
Migration	20
Conversion	21
Encapsulation	21
Emulation	21

Implementing a Digital Records Preservation Strategy	21
Choosing an Approach to Digital Records Preservation	21
When Should a Digital Preservation Treatment Be Applied?	22
Planning to Implement a Preservation Strategy	22
Implementing the Preservation Strategy	23
Requirements for a Successful Preservation Strategy	24
Archival Storage of Electronic Files.	25
Quarantine.	26
Preservation/Conversion.	26
Secure Repository.	27
File Format Types -First Steps.	27
Proprietary and Non-proprietary File Formats.	27
File Format Types.	27
Executive Summary.	28

Purpose

This publication provides guidance to general agencies and annual conferences on creating, managing and preserving digital records. Digital records must be actively managed in order to ensure they are available and usable for as long as required to support accountability, good ministry and the expectations of the public. All requirements relating to permanent and temporary records described in the Guidelines for Managing Records apply to electronic records as well (see 2008 Discipline ¶ 1711.1b). This document deals issues unique to handling and preserving electronic records.

The guidelines contain advice on:

- the importance of managing digital records and how to manage them in an integrated way.
- creating and capturing digital records, and associated metadata, into recordkeeping systems;
- storing and securing digital records.
- preserving digital records for as long as they are required, including an overview of the General Commission on Archives and History approach.

Scope

The advice contained in these guidelines applies to all digital records created by agencies as evidence of business activity. Digital records include all records that are created in a digital format (born digital), or have since been converted into a digital format.

These guidelines draw upon recordkeeping requirements for the management of digital records that are set out in GCAH's Guidelines for Managing Records and best practice standards. No document like this is written in a vacuum. It relies heavily on a variety of documents and studies from around the U. S. and around the world, including reports done by the U. S. National Archives, various states and the National Archives of Australia. In particular the following studies have been useful for these guidelines:

Electronic Records Management Guidelines (version 4, March 2004), Minnesota Historical Society.
Digital Recordkeeping: Guidelines for Creating, Managing and Preserving Digital Records. May 2004. National Archives of Australia.

Audience

These guidelines should be used by all General agencies, episcopal offices and conference offices to ensure that their digital records are managed appropriately. They are relevant to all agency staff members with responsibility for managing digital records.

There are two distinct audiences for these guidelines:

- those with responsibility for managing information technology (IT) and communications, including e-business and websites; and
- those with responsibility for managing records, information and knowledge.

These guidelines contain detail that is required to meet the information needs of both audiences. Some sections may include information that is considered unnecessary by the members of one audience. For instance, information about the IT requirements of digital records has been included for the information of records, information and knowledge managers. Information about the management of digital records has been provided to assist the IT audience

First Steps

Included in these guidelines are suggestions or “first steps” which can be taken by staff almost immediately to begin managing their records. First steps should not be confused with developing a satisfactory management solution for digital records. By allowing and encouraging staff to take these first steps an agency can begin to incorporate into daily practice the necessary steps which will lead to a productive digital records management policy. The point being that such a policy should be introduced incrementally instead of all at once.

The Importance of Digital Records

What Are Digital Records?

Records are evidence of business or ministry conducted by an agency. Records can be in any form, including digital. Digital records are records created, communicated and maintained by means of computer technology. They may be ‘born digital’ (created using computer technology). Or they may have been converted into digital form from their original format (e.g., scans of paper documents).

Agencies create and store digital records in a variety of ways. Common types of digital records include word-processed documents, spreadsheets, multimedia presentations, email, websites and online transactions.

However, digital records can be found in many systems throughout an organization – including databases and business information systems, shared folders and hard drives. The following list is not exhaustive, but highlights the range of digital records covered by these guidelines.

Documents created using office applications:

- word-processed documents

Records in online and web-based environments:

- intranets

- spreadsheets
- presentations
- desktop-published documents
- extranets
- public websites
- records of online transactions

Records generated by business information systems:

- databases
- human resources systems
- financial systems
- workflow systems
- local church membership management systems

Electronic messages from communication systems:

- email
- SMS (short messaging services)
- MMS (multimedia messaging services)
- EDI (electronic data interchange)

These records are subject to the same legal requirements as records on paper or any other format.

To be of value as evidence, a digital record must possess content, context and structure.

This means that a digital record:

- has information content that is, and continues to be, an accurate reflection of what occurred at a particular time;
- can be placed in context so that the circumstances of its creation and use can be understood in conjunction with its information content; and
- can be reconstructed electronically when required so that each component part is brought together as a whole and presented in an intelligible way.

The best way to preserve the content, context and structure of a record, is to manage it within a recordkeeping system.

Why Manage Digital Records?

Digital records must be managed for the same reasons records in other formats need to be managed. Records allow our ministry to be conducted efficiently and effectively. There are accountability and legislative obligations that agencies must meet, and community expectations concerning the documentation and transparency of our actions. Further, special challenges, such as technological obsolescence and media degradation, make it imperative for digital records to be carefully managed.

Inadequate records and poor recordkeeping practices can contribute to accountability failures and inefficient performance. Effective recordkeeping strategies can lead to many benefits.

Records preserve an agency's history and form its corporate memory. Information about previous decisions and actions can improve service quality and effectiveness. Timely access to relevant data allows action officers to make decisions and do better business.

Challenges Associated with Digital Records

Managing digital records involves the following unique challenges:

- Digital technology evolves at a rapid rate. The software and hardware used by an agency to create digital records tends to be short-lived, quickly replaced by upgrades or improvements. Because of this hardware and software obsolescence, digital records can quickly reach a point where they cannot be read or understood. Yet, in order to meet legislative obligations, records must remain accessible for as long as they are required.
- The general manipulability of digital records means that they can quickly and easily be updated, deleted or altered. However, digital records are evidence of business activity and must be managed securely to prevent unauthorized modification.
- Metadata can be intrinsically linked to a digital record, or it may be contained within the systems used to generate or store the records. Capturing and maintaining metadata, including technical specifications, is necessary in order to ensure the preservation of digital records and their continued accessibility over time.

FIRST STEPS – FILE NAMES

While the *Guidelines* strongly recommend the use of a records management system to control digital records, there are steps that can be taken by all to begin managing digital records before a more comprehensive system is developed or installed. The first step would be the creation of a file naming convention or policy. A procedure for naming files is key in being able to select the files that are to be saved. **To avoid file names conflicting when they are moved from one location to another, each record's file name should be unique and independent from its location.** As you develop your file naming policy, you will need to be familiar with the following: developing your file naming policy, you may wish to include some of the following common elements:

- 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 initials (e.g., Department MUST [DeptMUST])
- Records series (e.g., SeriesX)

Digital Recordkeeping Framework

Integrated and Comprehensive Approach

Agencies are encouraged to pursue a holistic approach to recordkeeping that is based on legal and business requirements, rather than record format. A digital recordkeeping framework allows the management of digital records to be integrated and consistent with the management of records in other formats.

Policies, Procedures and Guidelines

Policies and procedures for managing digital records are an important element of a digital recordkeeping framework. Policies define the organization's approach to managing digital records and provide the necessary senior management authority for the implementation of the framework. Procedures outline how the policies will be implemented and provide clear instructions for their practical application. Where necessary, policies and procedures can be supplemented by guidelines to provide additional clarification and direction.

Policies, procedures and guidelines should be developed to suit the organization's size, complexity, corporate culture and structure. A small agency, for instance, may have a single policy covering the management of all digital records. Larger agencies may have multiple policies covering specific areas of digital recordkeeping, such as electronic messages, preservation of digital records, web-based digital records and digital records security.

Systems Design

The recordkeeping practices set out in the policies, procedures and guidelines need to be supported by systems that are capable of keeping records.

Innovative technical solutions can reduce the need for conscious employee involvement in the recordkeeping process, reduce reliance on staff making records selection decisions and improve overall workflow by streamlining records capture. Please note that storing digital records on structured network drives is not a substitute for a controlled recordkeeping system. However, improving the way files are stored on private drives and shared directories will assist the overall management of corporate information. For example, establishing a classification scheme and naming conventions can create a semi-structured environment for digital objects. This can be a useful interim measure, until recordkeeping systems are developed.

Creating Digital Records

Identifying and Creating Digital Records

To be of value as evidence, a digital record must possess content, structure and context, and should be managed within a recordkeeping system.

Responsibility for creating and capturing digital records into recordkeeping systems often rests with agency staff. Agency staff identify, create and capture digital records in the course of daily business. Emails sent and received are placed on file, word processed documents are drafted, altered and finalized within an agency's electronic recordkeeping system, and forms or correspondence are scanned into electronic document management systems. GCAH is ready to participate with agencies and conferences in training workshops.

Capturing Digital Records into Recordkeeping Systems

Capture is the process of lodging a document into a recordkeeping system and assigning metadata to describe the record and place it in context, so that the record can be managed over time.

The procedures and practices an agency establishes to capture its digital records will depend on the recordkeeping systems in use, the types of digital records generated and the specific recordkeeping requirements the agency must satisfy.

Recordkeeping systems are specifically designed to capture evidence of business transactions. They are distinguished from business information systems by their ability to manage the content and structure of records, provide access to them over time, and maintain linkages between records and the activities they document (context).

FIRST STEPS – FILE CONTROL

General issues to consider as you develop a file control and naming policy include:

- Determining what information about the file to collect. You will need to decide what information about the file to collect and include in file names. Collection and use of information about the file in file names 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.
- Determining the official copy. Determine which file is the “official” copy.
- 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

Creating Information about Digital Records

What Is Metadata?

Metadata is data describing the context, content and structure of records and their management over time. Metadata allows users to control, manage, find, understand and preserve records over time. Some examples of metadata are:

- the title of a record
- the subject it covers
- its format
- the date the record is created
- the history of its use
- details of its disposal.

There are two main categories of metadata that are used to manage digital records – recordkeeping metadata and resource discovery metadata.

Recordkeeping Metadata

Recordkeeping metadata is structured or semi-structured information that enables the creation, registration, classification, access, preservation and disposal of records through time and across domains. Recordkeeping metadata can identify, authenticate, and contextualize records and the people, processes and systems that create, manage, maintain and use them.

Capturing and Maintaining Metadata

Capturing and maintaining good recordkeeping metadata supports digital recordkeeping by:

- protecting records as evidence and ensuring their accessibility and usability;
- ensuring the authenticity, reliability and integrity of digital records;
- enabling the efficient retrieval of digital records;
- providing logical links between digital records and the context of their creation, and maintaining the links in a structured, reliable and meaningful way;
- allowing timely destruction of temporary-value records when business use has ceased; and
- providing information about technical dependencies, to help ensure digital records' long-term preservation and usability.

When Should Metadata Be Captured?

Recordkeeping metadata will generally be identified, and/or created when digital records are captured into recordkeeping systems. This defines the point at which the information formally becomes a record, fixes it in its context and enables its appropriate management over time.

Metadata collected at the point of capture of a digital record should document its content, structure and the context in which it was created.

Some metadata may be applied at a system level. For example, all records within a finance system will share the same metadata about the organization creating the record, and the business activity being documented. This metadata can be automatically applied to all records generated within the finance system.

Other metadata will be generated as time progresses. Metadata related to business and recordkeeping processes will be added to a digital record during its lifetime. Examples include History of use (when the record was last viewed, whether it was accessed illegally), Location and Disposal status. Such metadata ensures the continued authenticity, security and integrity of the record.

Determining How Long to Keep Digital Records

How Long Do Digital Records Need to Be Retained?

As with other formats of records, digital records need to be retained until they are no longer required for any purpose. There are three general reasons digital records need to be created and kept:

- to meet the requirements of legislation and accountability
- to support the efficient conduct of business
- to meet the expectations of the community.

Generally, digital records, as with other records, will fall into one of the following categories.

- Temporary value – the records can be disposed of at an identified time (e.g., “Destroy 7 years after action completed”). Temporary-value records can range in retention length from a very short period, such as one year, up to an extended period, such as “Destroy 130 years after date of birth (of subject)”.
- Retain permanently in agency – the records have a long-term business use in the agency, but are not considered to have archival value.

- Archival value – the records cannot be disposed of but instead will be retained in the custody of the Archives indefinitely. Agencies nominate groups of records when they develop a records disposal authority. GCAH decides which records meet the criteria for archival value.

A digital record must be managed, and remain accessible, for its lifetime. How long a digital record needs to be kept will influence its management. In these guidelines, we refer to “retaining digital records for the long term”. Given the vulnerable nature of most digital media and the frequency of technology change, “long term” for digital records generally means longer than one generation of technology. Digital records that must be retained for the long term will require active management to ensure their continued accessibility.

Disposal of Digital Records

GCAH has produced guidelines for records retention which contain recommended disposal schedules for the episcopal offices and general agencies. A general agency is to produce its own retention guidelines, but they must be approved by GCAH before they can be placed into practice.

Normal Administrative Practice

Normal administrative practice (NAP) defines types of records that agencies may routinely destroy in the normal course of business. Agencies do not need to contact the Archives for permission to dispose of records within the scope of NAP. NAP usually applies to information that is duplicated, unimportant or only of short term facilitative value. For example:

- superseded system backups;
- trivial electronic messages that are not related to agency business;
- address lists and change of address notices;
- calendars, office diaries and appointment books (unless identified in a records disposal authority as having additional value);
- rough drafts of reports, correspondence, routine or rough calculations;
- routine statistical and progress reports compiled and duplicated in other reports;
- abstracts or copies of formal financial records maintained for convenient reference;
- duplicated material such as forms or templates;
- thermal paper facsimiles after making and filing a photocopy or scan.

The NAP provision must not be used to:

- destroy records of significant agency operations;
- destroy records that document the rights and obligations of the government or private

individuals;

- cull documents within files; or
- destroy business-related electronic messages before they become part of the formal record.

NAP should not be applied to records or information that can be used as evidence.

Storing Digital Records

How Are Digital Records Stored?

To ensure the ongoing protection of digital records, agencies require efficient and effective means for maintaining, handling, and storing digital records -- both active and inactive -- over time. Policies, guidelines and procedures for the storage of digital records should be an integral component of an agency's digital recordkeeping framework. There are three ways in which agencies may store digital records -- online, offline or nearline.

- Online – Online records can be contained on a range of storage devices (e.g. mainframe storage, network attached storage or PC hard drive) that are available for immediate retrieval and access. Generally, records stored online will be active digital records – i.e. records that are regularly required for business purposes. Electronic messaging systems and word-processed documents saved to the network server fall into this category.
- Offline – Offline digital records are contained on a system or storage device that is not directly accessible through the agency network and which requires human intervention in order to be made accessible to users. Digital records that are stored offline are usually retained on removable digital storage media (e.g. magnetic tape, CD, DVD) and are generally inactive digital records not regularly required for business purposes. Offline digital records may be stored offsite as part of an agency's business continuity plan. Digital records stored offline are not immediately available for use. Agencies must take responsibility for monitoring and guarding against environmental degradation and changes in technology that may adversely affect the storage media employed.
- Nearline – Nearline storage of digital records means the records are contained on removable digital storage media, but remain relatively accessible through automated systems connected to the network. These digital records are technically considered to be offline. The use of systems such as CD jukebox or magnetic tape silos allow them to be made available through agency networks, in relatively short periods of time and without the need for human intervention (i.e. staff are not required to physically retrieve the storage media on which the required information is retained).

Generally, digital records will begin life as online records and, as the immediate business need to refer to them diminishes over time, they will be moved to either nearline or offline storage, depending upon the technology available to the agency, the ongoing relevance and value of the records and their retention requirements.

Selecting the Appropriate Storage Method

We strongly recommend that digital records of vital significance to an agency, as well as digital records required for long-term retention within agencies, and digital records of archival value, be stored online.

Online storage devices, such as network storage devices and mainframe storage, have the following advantages.

- Digital records stored online will, in most cases, be retained on the magnetic hard drives that form an agency's core network, where they will be readily accessible to users and can be maintained and controlled as an integral part of the agency's recordkeeping system.
- Large storage capacities allow for significant quantities of digital records to be retained on a single storage device.
- Regular integrity checks of digital records can be more readily performed and, in some instances, it may be possible to automate these tasks.
- Digital records stored online have a greater likelihood of being identified and included within any changes made to agency IT systems, such as system-wide migration processes.
- Online storage devices need not be linked directly to an agency network. Where security concerns, business considerations or other factors warrant, agencies may opt to establish standalone online storage systems.
- Increasingly, online storage systems can support sophisticated automated techniques and redundant designs that aid digital record control, monitoring and backup.

GCAH does not recommend CDs, DVDs, magnetic tape or other removable digital media formats that are physically maintained but not accessible from active computer systems. Offline digital storage devices are suitable only for storing relatively low-value digital records and are not recommended for long-term digital records, vital records or records identified as being of archival value.

Securing Digital Records

Why Is Security Important for Digital Records?

Security is important for all records. The manipulable nature of digital records means that, in the absence of appropriate safeguards, it is relatively easy to alter or delete them -- whether intentionally or unintentionally. Alterations to digital records can be virtually undetectable, undermining their evidential value as records.

When implementing systems, agencies must therefore take special care to ensure they are secure, reliable and capable of producing records that are acceptable for legal, audit and other purposes.

Methods of Securing Digital Records and Systems

The following are some basic practices and protocols agencies, episcopal offices and annual conferences may adopt to ensure they maintain adequate security for their digital records and systems. This list is not exhaustive. Agencies should select a combination of methods to suit their needs.

- Limit access to digital records, and the systems on which those records are created and kept, to authorized personnel in order to protect the integrity of the records and prevent unlawful alteration or destruction of records.
- Establish network security systems, such as firewalls, to protect against unauthorized access (e.g., hackers) to systems that are accessible through external connections, such as the Internet.
- Install appropriate gateway filter software on messaging systems, and ensure that filter definitions are regularly updated, to protect against spam, denial of service attacks and malicious code, such as computer viruses.
- Implement public key infrastructure (PKI) encryption technologies to ensure, when appropriate, secure transmission of digital records to external parties.
- 'Lock' final digital records to prevent any subsequent alterations or inadvertent destruction (e.g. finalizing records as 'read-only' within an electronic recordkeeping system -- do not use passwords).
- Use digital signature technologies to authenticate digital records and provide security and confidence in authorship.
- Store vital digital records either on systems without external links or, if necessary, offline.
- Establish appropriate systems backup procedures and disaster recovery strategies to protect against loss of digital records.
- Develop and implement audit trails to detect who accesses a system, whether prescribed security procedures were followed and whether fraud or unauthorized acts have occurred, or might occur.

Authentication of Digital Records

Digital records provide evidence of agency business activity. For digital records to retain their evidential value, and be admissible as evidence in court, systems and practices must prevent the unauthorized alteration of digital records, and so ensure their continued authenticity.

To guarantee the authenticity of digital records, systems and procedures should be capable of establishing:

- if digital records have been altered

- the reliability of software applications creating digital records
- the time and date of creation and alteration of digital records
- the identity of the author of a digital record
- the safe custody and handling of records.

Version control is a useful tool for preserving the authenticity of digital records. Digital source records should be clearly distinguished from any subsequent copies. Identification may be achieved through labeling of records or by time and date stamps.

To provide evidence of business activities or action taken, agencies must be able to clearly demonstrate the provenance of digital records. This includes establishing the original conditions for the creation of the record, such as date and time of creation, software application integrity and the author or sender of a record. The ability to track when the record was last altered, by whom, and the 'chain of custody' (who was responsible for the record) will also support a record's evidential value. Clearly implemented policies and procedures demonstrate that an agency has protected the provenance of its digital records.

In some instances, authenticity may be demonstrated if access to digital records is restricted to authorized persons or applications. In such cases, there must be security mechanisms to prevent unauthorized persons or applications accessing the digital record. Audit trails should be able to verify that digital records have not been accessed inappropriately or illegally.

Long-term Digital Records

Agencies, episcopal offices, and annual conferences should take additional care to ensure that their security and authentication mechanisms do not inadvertently make digital records inaccessible in the long term and that the evidential value of the records does not diminish over time. This is particularly important for records of archival value.

Agencies should address the following considerations when applying security and authentication practices and protocols to digital records.

- If manipulation of data is required (e.g., encrypting a file to send via email), the process should be applied to a copy of the source record. The source record should be maintained separately in a secure recordkeeping system. This will safeguard records of long-term or archival value in the event that data loss occurs during manipulation or the file becomes inaccessible.
- Access controls should be applied and maintained within a recordkeeping system. Staff should be discouraged from using software functions to create passwords or limit access to business records. This includes specifying who can read or alter a document, preventing copying or printing, or setting an expiration date.

Documents that have been password-protected or otherwise restricted should not be captured into a recordkeeping system while the restrictions are still in place. There is a considerable risk that records will become inaccessible as staff changes occur and passwords are forgotten over time. Unrestricted records should be captured into recordkeeping systems. Authorized users can then apply access controls, in accordance with business rules.

- Metadata attesting to the validity of a digital signature, where a digital record has been authenticated using such technology, should be captured and maintained for as long as required. If such information is not kept, the evidential value of the record may be undermined.
- Digital records that have been encrypted should be decrypted prior to capture in a secure recordkeeping system. Metadata relating to the encryption and authentication process should be captured and maintained for as long as required. If encrypted records are captured and kept there is a considerable risk they will become inaccessible

Preserving Digital Records for the Long Term

Why Preserve Digital Records?

Considering the problems of technological change, and the potential instability of digital storage media, 'long term' may not be very long. When applied to the preservation of digital records, 'long term' usually means 'greater than one generation of technology'.

Many records have retention periods greater than one generation of technology. It is important that these records are preserved and accessible for use in daily business. Long-term records support strategic planning and decision-making. They act as corporate memory, reducing duplication of work and improving business efficiency.

There may also be evidentiary reasons to keep digital records for extended periods, as part of a risk minimization strategy. Inaccessible records may expose agencies to accountability failures and potentially costly consequences, such as legal action. GCAH recommends agencies develop strategies to preserve/migrate electronic records, and to ensure that all digital records are captured into a corporate recordkeeping system.

Long-term maintenance is particularly significant for digital records of archival value. Inadequate preservation strategies can render digital records inaccessible and unusable. Future approval of agency and episodic retention schedules will depend upon their having this element in their plans.

Accessibility requirements apply to all digital records, not just those of archival value. Digital records must remain accessible for as long as they are required.

Planning for Technological Obsolescence

Digital records are dependent on various combinations of hardware, software and media to retain their content, context and structure. Agencies must ensure that the technology required to render a digital record usable and accessible is available. It is not sufficient to simply retain records in digital format; the records and associated metadata must be in a format that is viewable with current technology.

Computer technology is subject to ongoing technological obsolescence, with both hardware and software quickly becoming outdated as new upgrades and versions come onto the market. This can result in digital records created using older hardware and software becoming inaccessible in their original form after a relatively short period of time.

Agencies that need to retain digital records for the long term should plan for technological obsolescence by ensuring that records can be copied, reformatted, converted or migrated across successive generations of computer technology. Such planning involves considering hardware, software, operating systems and storage devices.

Agencies need to consider a number of interrelated software and hardware issues when preserving digital records, including:

- the proprietary, platform-specific nature of many software applications and the likelihood of their continued availability;
- the cost of maintaining access to obsolete formats (including operating system software and licensing fees) for a system no longer in active use;
- the estimated physical and/or commercial life of the media on which digital records and related metadata are stored; and
- the long-term availability of the hardware and operating system platforms needed to access records stored on different types of media.

The need to plan for technological obsolescence and provide for the preservation of digital records should be incorporated, through a formal digital records preservation strategy, within the digital recordkeeping framework. The preservation strategy should outline the approach adopted by the agency for the preservation of its digital records. There are several common techniques.

In order to adequately manage the preservation of digital records over time, and ensure their continued accessibility, agencies must be proactive. They should develop and implement organization-wide strategies targeted at identifying, managing, preserving, and ensuring continued access to digital records.

An effective digital records preservation strategy should incorporate formal policies and procedures governing the agency's approach to the long-term management of its digital records and establish processes to ensure their ongoing maintenance.

An agency's digital records preservation strategy must reflect its legislative obligations, industry standards and best practice. The tools listed at the end of this chapter will be especially helpful to agencies formulating a digital records preservation strategy.

A digital records preservation strategy should be supported by a plan for its implementation that is promulgated to relevant staff. This can be achieved by:

- formulating policies, procedures and guidelines to provide a formal framework within the organization for the implementation of the strategy; and
- providing manuals, information and reference sheets, and training for staff to ensure the preservation strategy is correctly implemented. Depending on the approach adopted, this may require training for all operational areas, not simply for records and IT staff.

Agencies should assign responsibility for the management of long-term digital records to an appropriate area within the organization where staff have relevant skills and qualifications. This will generally be a specialized information or knowledge management unit headed by a senior information officer. Responsibility for policy and procedure formulation, implementation of strategies to preserve digital records, evaluation, monitoring and review of processes, and delivery of training should rest with this area.

Agencies should ensure that their digital records preservation strategy takes into account digital records that may be created and managed by outsource providers and that these contractors are also required to actively comply with the agency's long-term digital records preservation strategy.

Techniques for Digital Records Preservation

Some early approaches to digital records preservation relied on storing records in their original format on physical media – much like boxes are used for the storage and protection of paper records. However, magnetic tapes and disks, and optical storage disks (e.g., CDs and DVDs) are manufactured for short-term storage of digital objects, not long-term archival retention. The greatest concern for this method of preservation, in addition to the relatively short life span of digital media, is the obsolescence of the hardware and software used to access the records. Rapid change in the IT industry and the move from science-based development to commercial development of software and hardware systems, has meant that media rapidly become inaccessible. Consequently, this approach to digital preservation has proven to be wholly inadequate and the GCAH strongly advises against this preservation strategy.

The most common techniques for digital preservation can be grouped into three broad categories. Any one or a combination of these may form the basis for an agency's digital records preservation strategy.

Migration

Migration relies on a program of constant transferral (migration) of digital records from older or obsolete hardware and software configurations or generations, to current configurations or generations in order to maintain accessibility. This strategy avoids the obsolescence issues of the physical media solution, preserving the functionality of the digital records and enabling users to retain access to the records -- but requires a substantial investment in resources to undertake the repetitive migration work involved. Furthermore, some characteristics of the original data format may not be retained through the migration process and, as a result, users will lose access to characteristics of the source record that may be important to its meaning.

Conversion

Conversion is the process of transferring digital records from their original data format to a standardized, long-term preservation format (also known as an archival data format). Conversion is also referred to as “normalization”, “stabilization” and “standardization”.

The conversion process is a form of migration. However, instead of migrating from an outmoded data format to a current data format, the original data format is migrated to an archival data format. Generally, archival data formats are open source, non-proprietary formats that provide greater potential longevity and are less restrictive than proprietary formats. Conversion reduces the need for repeated migrations.

Encapsulation

Encapsulation requires metadata to be bundled with, or embedded into, the digital object. The metadata allows the record to be intellectually understood and technologically accessed in the future.⁸ A viewer is then required to display the records. This packaging of contextual information ensures the integrity and authenticity of records over time. However, there is some risk that important metadata may be overlooked during encapsulation.

On its own, encapsulation cannot preserve digital records. This technique should be used in conjunction with migration or emulation to ensure the ongoing accessibility of the records.

Emulation

Emulation uses software to recreate the digital record's original operating environment to enable the original performance of the software to be recreated on current computer systems. The result is that the original data format is preserved and may be accessed in an environment that allows for the recreation of the original 'look and feel' of the record. The downside to the emulation approach is that the creation of the underlying emulator software is costly, requiring highly skilled computer programmers to write the necessary code. Furthermore, the intellectual property and copyright issues associated with the emulation of proprietary software may undermine the effectiveness and sustainability of the approach.

GCAH's approach to digital preservation uses a combination of these techniques.

Implementing a Digital Records Preservation Strategy

Whatever strategy an agency adopts to keep digital records accessible, it must address a number of common issues.

Choosing an Approach to Digital Records Preservation

Agencies, episcopal offices and annual conferences should consider the following factors when choosing an approach to digital records preservation:

- cost of implementation, including cyclical costs for ongoing preservation treatments;
- technical complexity of the selected approach and the capacity of the agency to support the approach over time (both technically and financially);
- compatibility with existing hardware and software;
- impact on business operations (e.g., whether the approach requires changing corporate work practices); and
- overall effectiveness and robustness of the approach in protecting the integrity, accessibility and functionality of the agency's digital records over time.

When Should a Digital Preservation Treatment Be Applied?

To maximize the long-term preservation prospects for digital records, preservation techniques must be applied as soon as practical, preferably while the records are still accessible. Most data formats have a limited window of opportunity during which preservation treatments can be applied before the format becomes outmoded and inaccessible. The sooner an agency addresses preservation issues and determines and implements an appropriate preservation approach, the higher the probability that the digital records will be successfully preserved.

Agencies are therefore encouraged to be proactive in pursuing their digital preservation strategies and to determine and implement appropriate digital preservation techniques before their digital records become outmoded and inaccessible.

Preservation treatments are often undertaken reactively in response to the immediate business needs of an organization, rather than as part of a considered solution to longterm digital records retention requirements. Such processes may be technology-driven exercises, initiated in response to changes in IT infrastructure or as a consequence of adopting new or upgraded software. In such cases, preservation treatments are undertaken primarily to ensure that existing digital records, particularly active core business records, are transferred from their original format into a new format capable of functioning within the upgraded IT environment.

Planning to Implement a Preservation Strategy

Periodic preservation treatments (such as migration) are often applied to digital records without necessarily considering the long-term implications for the integrity of the records. If sufficient care is not taken to protect the integrity and authenticity of the records, migrating software and hardware systems can jeopardize their evidential value.

Agencies that apply preservation treatments to data formats without properly assessing the processes, risk the loss or limitation of the functionality, format, structure and content of their digital records and the

potential loss of metadata relating to the records.

Planning for the preservation of digital records will allow agencies to retain the functionality and integrity of digital records after successive upgrades of hardware and software.

Development of preservation strategies, and the selection of an appropriate approach, should be the result of a collaborative effort between the records and IT sections within an agency. Best practice recordkeeping issues need to be carefully considered, and the input of agency records and information personnel taken into account, before any preservation processes are applied to an agency's digital records.

Implementing the Preservation Strategy

Although the three main preservation techniques – migration, encapsulation and emulation -- differ substantially in their method of preserving digital records, they share common ground in the process of implementation. The following steps outline the implementation process.

1. Identify records requiring preservation – Identify and select digital records that require the application of preservation treatments in order to ensure their continued accessibility.
2. Research technical solutions – Investigate the hardware and software technologies required to successfully implement the agency's preferred preservation approach. In the case of emulation, this may involve the development of specialized software capable of re-creating the source records within a new computer environment. In the case of migration, this may involve identifying suitable migration paths (i.e., software applications with sufficient backward compatibility to transfer source records from an outmoded data format to a current data format). In the case of encapsulation, this may involve software with the ability to embed metadata or 'package' it with the record.
3. Test proposed solution – Before a preservation approach is fully implemented, agency staff must conduct comprehensive testing of the technical processes. Testing should be performed on duplicates of source records.
4. Back up records identified for preservation – Prior to implementation, all digital records identified for preservation treatment should be backed up. The integrity of the duplicates should be verified before they are removed to a secure storage area. These duplicate source records should not be subjected to a preservation process and will serve as master copies should the selected preservation treatment be unsuccessful.
5. Apply the preservation treatment – After successful testing, the treatment should be applied to all digital records identified for preservation treatment. For migration and encapsulation techniques, this would entail applying preservation treatments to the source records, thereby altering their format. For an emulation-based technique, the records

identified for preservation would be transferred to the new environment – without altering the records themselves.

6. Audit the integrity of preserved records – Following implementation of the preservation process, the preserved records should be subjected to rigorous testing to ensure that any reduction in functionality, or loss of content, structure or format, is within previously set limits of acceptability. The integrity of all relevant metadata associated with the preserved records should be verified. Metadata should also be updated to record the preservation treatment.

If the records cannot be verified, the preservation process will need to be repeated on new duplicates of the source records (steps 4 to 6). In some instances, the preservation strategy itself may require re-evaluation.

7. Destroy source records where appropriate – Once the preservation process has been completed and the integrity of the preserved records has been verified, agencies may destroy the duplicate source records..

8. Establish monitoring regimes – The integrity of the preserved records, their functionality, structure, content and context, and associated metadata, should be monitored periodically following preservation to ensure the stability of the preserved records and to identify when subsequent preservation treatments are required.

Please note that, if it appears likely at any stage during the application of a preservation treatment that digital records of archival value may be lost or significantly altered as a result of the preservation process, the Archives should be consulted immediately so that alternative arrangements may be considered.

Agencies experiencing significant difficulty in ensuring the continued accessibility of their digital records should contact the Archives for advice.

Requirements for a Successful Preservation Strategy

A successful preservation strategy ensures the continued integrity of the digital records, as well as their continued accessibility and functionality. The preservation of integrity requires that the records, and their associated metadata, remain reliable, complete and authentic.

The following steps will ensure successful preservation of digital records.

- Care is taken in selecting and testing software applications and hardware required for preservation processes.
- Where possible, non-proprietary, fully documented, open source data formats are used – particularly when implementing migration-based preservation techniques. Proprietary data formats are not recommended for long-term storage of records.

- Preservation processes are applied systematically to all digital records, both current and non-current, retained by an agency. Failure to include non-current digital records can result in their inaccessibility.
- All relevant metadata (for the records and the preservation process) is captured at the time of preservation.
- Preservation processes are fully documented and the documentation retained to help inform future preservation efforts. Any copying or reformatting of data for migration or conversion should be documented in the recordkeeping metadata.
- Preservation processes are carried out in accordance with relevant recognized recordkeeping, information and data management standards.
- Guidelines and procedures are issued and staff are encouraged to adopt common usage rules to help standardize the application of the selected techniques across all agency systems.
- Where records are migrated, converted, copied or reformatted, the success of the process must be verified and data integrity confirmed before the duplicate source records are destroyed.
- Any alteration or loss of functionality, structure, content or appearance that occurs as a result of preservation is fully documented in the recordkeeping metadata.
- Thorough checking regimes are put in place following preservation to monitor record integrity and identify when further preservation treatments are required.

Archival Storage of Electronic Files

While the above guidelines have been focused on office staff and how to treat the records in an office environment, they are also important to archivists. In order for the record to be trustworthy it needs to be reliable and authentic, complete, accessible and durable. These are all core values for archivists as they are for office workers. This has been the purpose of recordkeeping since the inception of records. Our tasks as archivists is to ensure the trustworthy character of the record over time. We need to set up a storage and preservation system that will convince the researchers of the future that the records we have bequeathed them are trustworthy.

We suggest a three step process.

! Quarantine

! Preservation/Conversion

! Secure Repository

Quarantine

Records will be brought into the archives either by the Internet or on some type of media, e.g. CD, DVD or a USB drive. Hopefully a list, or manifest, of the material will have been created as part of the transfer. If not then one needs to be made. On a computer isolated from the rest of the network the contents of the transfer media are checked to confirm that the files received are those that the originator intended to send and all media are checked for the presence of computer borne viruses.

Once a virus check is passed the records are copied to a carrying device, disconnected from the Quarantine network and stored for a period of 28 days. During the 28 day quarantine period, the virus definitions on the Quarantine network are updated daily.

After 28 days have passed, the carrying device is again connected to the Quarantine network and the data is again scanned for the presence of viruses.

The first thing an archives has to do is to make sure that its records are "clean." Just as we investigate for mold and insects we need to be sure that no damaging viruses are brought into the archives. The 28 day time lag ensures that virus protection will catch up to any new viruses.

Preservation/Conversion

If necessary the records need to be converted to a standard format, either proprietary or an open standard such as openDocument. We recommend conversion for the following reasons.

- ! Keeping working copies of older generations of PCs , operating systems and software is just not a viable option. The expertise to manage these older systems is huge. Merely finding parts for many of these machines would prohibit this as a workable solution.
- ! Building software that emulates the older operating systems and software. Again, the expertise to do this would be huge, and costly. Almost all of the older systems are of a proprietary nature and there would be charges and challenges in creating something that worked like an older system.
- ! Converting the files to a standard format reduces the complexity of the number of file types to deal with. If, and when, that "standard" or "open" file type becomes obsolete, then the associated costs of the next transfer is also less since the conversion is from one file type to and other. If the conversion is well-designed it will have minimal impact on the fixed nature of the record.

So, as the next step once the material has been determined to be virus-free a carrying device is connected to the Quarantine network and taken over and connected to the Preservation network and the individual data files are converted to your selected standard preservation format. The files created by this process should be recorded on a second carrying device for transfer to the Digital Repository.

Secure Repository

The carrying device is connected to the Digital Repository network and the files of the original data are copied to long-term storage arrays. These storage files should be in a secured space, with limited access and a very secure password. Only a few accounts should be on the system. Logs should be kept of who access this system. For general public access to these records, copies should be made and placed on the public network. But even these files should be placed on a read-only drive.

File Format Types -First Steps

There are a variety of different file formats. Offices are advised to limit the number of types and the corresponding software that supports them. The more file formats and software in use, the greater chance for loss of information because the files become inaccessible. You consider the file format options available to you, you will need to be familiar with the following concepts:

- Proprietary and non-proprietary file formats
- File format types

Proprietary and Non-proprietary File Formats

A file format is usually described as either proprietary or non-proprietary:

- Proprietary formats. Proprietary file formats are controlled and supported by just one software developer.
- Non-proprietary formats. These formats are supported by more than one developer and can be accessed with different software systems. For example, eXtensible Markup Language (XML) is becoming an increasingly popular non-proprietary format.

File Format Types

Files fall into the following large categories

- Text files. These are files associated with MS Word, Wordperfect and other word processing files.
- Some use proprietary files, such as Word, while other files, ASCII files are used by simple word processors like NoteTab.
- Portable Document File (PDF) is a popular proprietary file type used by Adobe Acrobat
- Graphic files,
 - Vector based files which store images as mathematical formulas. Most frequently used in architectural files and PostScript files used in publishing. These images can be scaled without distortion;
 - Raster-based files that store images as a collection of pixels. These are also called bit-

mapped images. They cannot be scaled without some distortion;

Bitmap (BMP) one of the earliest. Low quality files often used in word processing.

Tagged Image Format file (TIFF), widely used in many programs. No compression is used in storing the data;

Graphics Intechange Format (GIF) file, widely used on the Internet;

Joint Photographic Experts Group (JPEG) is mostly commonly found in digital cameras today. This popular file uses compression when storing an image.

- Data files. Files used by database software. Most often today these are relational files, which means the file structure is placed in a type of tabular structure. However, internal structure can vary and the indices which accompany the database are often have a unique structure. Also, the growth of popularity of the XML database adds to the complexity of this general file type.
- Spreadsheet files. Spreadsheet files hold information in a tabular format as well as relationships between the various cells of the tables. These are often proprietary.
- Video and audio files. These files hold moving images and sound. Almost all of the popular formats are proprietary.
 - WAV files. An audio file which captures sound with little or no compression.
 - MP3 files. An audio file used most often in portable players and on computers.
 - MPEG files. A compressed movie file.

When creating or saving files it is always better to use a standards-based file and one which uses little to no compression.

Executive Summary

Agencies at all levels of the denomination are creating more and more electronic records. the rapid obsolescence of digital technology, agencies should plan for the long-term preservation of digital records. Digital records that are to be retained indefinitely by the agency require preservation to ensure their ongoing accessibility. Because digital records can be easily modified, their security is very important. Agencies should plan for disasters -- loss of digital records can be crippling. Agencies should develop an integrated and comprehensive framework for digital recordkeeping.

First must be sure we understand that we are talking about the preservation of electronic records, not electronic publishing. Electronic publishing is the conversion of an existing document, book or image into a digital format and making it available over the Internet. What is currently happening is analogous to the spat of publishing that took place at the end of the 19th and into the 20th centuries. The papers of individuals were collected and published. Electronic publishing allows the originals to be available for the public and scholars. While this is not necessarily a bad thing, it ties up a significant amount of resources and time in recreating what already exists. In many cases libraries are taking existing documents and just digitizing them. Scholars and Librarians tend to think in subject areas first and so there is a propensity

when they do turn to archival material to select material by subject matter and then to publish it on the Internet. They become publishers and researchers instead of dispensers of information.

What we are concerned about is the growing body of material which is born digital and which must be preserved for the future. This calls on us to sharpen our appraisal skills and to develop new ways of managing a large body of documentary material.