Library of Virginia

ELECTRONIC RECORDS GUIDELINES

December 2009

Scope

This guideline, created by Records Analysis Services, is applicable to all Commonwealth of Virginia agencies and local government entities. It discusses essential characteristics of electronic records, principles of electronic records management, and long-term preservation strategies. While the recommendations set forth in this document reflect current best practices in the field of electronic records, this set of guidelines is not meant to define mandatory standards or to act as an information technology guide. Agency and locality personnel should work with in-house IT departments to address questions related to computer systems, such as network, database management, or system backups concerns.

Legal Framework

The following laws provide a legal framework for electronic records management in the Commonwealth of Virginia. Agencies and localities are mandated by these laws effectively and efficiently to preserve and maintain their public records in any format. This list is not comprehensive.

- Virginia Public Records Act (*Code of Virginia*, § 42.1-76–§ 42.1-91)
- Virginia Uniform Electronic Transactions Act (*Code of Virginia*, § 59.1-479–§ 59.1-498)
- Copies of Originals as Evidence (*Code of Virginia*, § 8.01-391)
- Virginia Freedom of Information Act (*Code of Virginia*, § 2.2-3700–§ 2.2-3714)
- Government Data Collection and Dissemination Practices Act (*Code of Virginia*, § 2.2-3800–§ 2.2-3809)
- Virginia Civil Remedies and Procedure (*Code of Virginia*, § 8.01)

Purpose

The purpose of the following guidelines is to provide best practices for public bodies that are developing an electronic records management strategy. The guidelines extend the policies and practices of paper-based recordkeeping to an electronic environment. Agencies and localities are responsible for implementing appropriate policies, procedures, and business practices in order to ensure that an electronic records management system protects the authenticity, reliability, integrity, and usability of public records. As such, the guidelines are designed to identify critical issues for public officials to

consider when designing, selecting, implementing, operating, and maintaining an electronic records system. The guidelines are divided into three sections with two appendixes:

- Section 1: Essential Characteristics of Electronic Records
- Section 2: Electronic Records Management
- Section 3: Long-term Preservation
- Appendix A: E-mail Record Decision Tree
- Appendix B: Glossary

Introduction

According to the Virginia Public Records Act (VPRA) of the *Code of Virginia* §42.1-77, a public record is defined as:

> recorded information that documents a transaction or activity by or with any public officer, agency, or employee of an agency. Regardless of physical form or characteristic, the recorded information is a public record if it is produced, collected, received, or retained in pursuance of law or in connection with the transaction of public business. The medium on which such information is recorded has no bearing on the determination of whether the recording is a public record.

Therefore, it is the content of a record, not its medium, which determines whether a record constitutes a public record. An electronic record is a record created, generated, sent, communicated, received, or stored by electronic means. Electronic record formats include, but are not limited to, word processing files, spreadsheets, e-mails, instant messages, Web sites, databases, and scanned images, as well as multimedia files that may include audio, graphics, video, and animation.

New information technologies have transformed the way information is created, used, disseminated, and stored. These new technologies enable us to collect information for and about citizens, document the business of government, and communicate, both within government agencies and between agencies and the public, in new and enhanced ways. Virginia state agencies and localities are responsible for ensuring that policies, practices, and systems for the management of electronic records are fully integrated into their records management programs. The following guidelines provide direction on the management of electronic records throughout their entire life cycle, from initial system design to final disposal or permanent preservation.

For additional Commonwealth information technology policies, standards and guidelines please see VITA's Information Technology Resource Management (ITRM) library. There are also several

organizations that develop nationally and internationally recognized standards in the field of electronic records management. The International Organization for Standardization (ISO), for example, provides a number of standards that relate to the management of records, specifically ISO 15489-1:2001 and ISO 15489-2:2001.

ELECTRONIC RECORDS GUIDELINES
SECTION 1: ESSENTIAL CHARACTERISTICS OF ELECTRONIC RECORDS

Like other government and commercial organizations, Virginia state agencies and localities face challenges in managing and preserving electronic records, as they are easily revised, deleted, changed, and manipulated. If appropriate measures are not taken, the essential characteristics of records can be altered or lost in the preservation process. Careful planning and system design are required to guarantee that the following characteristics of electronic records are both captured and maintained for the lifetime of the record.

The essential characteristics of electronic records are:

1. **Content**

   The content piece of a record is the information it conveys. Content can be composed of numbers, text, symbols, data, images or sound. The information content of a record should be an accurate reflection of a particular business transaction.

2. **Context**

   Contextual information is crucial to the evidentiary function of records. If a record lacks key information about its creator, the time of its creation, or its relationship to other records, its value as a record is severely diminished or lost entirely. In the case of paper records, much of the context can be found attached to a record's content. Electronic records pose additional challenges in this area. As a result, contextual information should always be collected, structured, and maintained with the record at the time of record creation. This involves identifying and labeling (or tagging) records and linking them to contextual information. In some cases this can be achieved by embedding key contextual information into the metadata or electronic records themselves.

   Metadata is, quite simply, data about data. More specifically, it is data describing the context, content, and structure of records and their management through time. A library catalog, for example, contains metadata in the form of entries for title, author, and subject, which are data related to books and other library resources. Effective metadata relies on a structured format and controlled vocabulary that is commonly understood among its creators and users. Because digital records can only be accessed using hardware and software, the role of metadata in electronic records is vital.

A government agency might choose to use information in a variety of ways. Whether dealing with issues of confidentiality, evidence, access, distribution, preservation, or destruction, it will be essential to understand and rely on the metadata that describes information. Government bodies use metadata to comply with records management laws, to document and design information technology systems, to document decisions and provide accountability, and to share and locate information.

Most software applications automatically create metadata and associate it with files. One example of metadata creation is the header and routing information that automatically accompany an e-mail message. Another is the set of properties created with Microsoft Word documents—certain elements such as the title, author, and file size can be automatically created, but other elements can be customized and created manually. Normally, some combination of automatically and manually created information is best for precise and practical metadata.

When creating metadata, be aware of the intended audiences as well as the information resources audiences use, the questions they ask, and their level of expertise. Furthermore, to increase the value of both metadata and the information it describes, work with other creators, custodians, and users of information. By agreeing on metadata standards, tools, and practices in collaboration with others, a more beneficial information management program is created.

The Library of Virginia encourages agencies and localities to maintain metadata relating to the

- Organization that recorded or maintained the records
- Other organizations that are, or have been, associated with the records
- Purpose of the records in fulfilling agency or locality functions
- Date of record creation
- Time period to which the records relate
- Frequency with which the records are, or will be, used
- Value or significance of the records in relation to the functions of the organization
- Record-keeping system used in relation to the records
- Relationship (if any) between the records and other records or materials
- Existence of any law, agreement, practice, procedure, arrangement, or understanding affecting the records

While such contextual information is absolutely necessary for long-term retention of electronic records, it can also improve the quality of records in active use, support information sharing, and enhance evidential quality.

3. Structure

Structure is defined as the appearance and arrangement of a record's content, including the relationships between fields, entities, language, style, fonts, page and paragraph breaks, links, and other editorial devices. Record-keeping systems must capture and preserve information about the structure of records either as part of the metadata associated with the records or in separate documentation. It is easier to preserve a record over time if it has a simple record structure. It is also advisable to base record structure on open standards to avoid dependence on a specific company or organization. Two examples of open standard include Standard Generalized Markup Language (SGML) and eXtensible Markup Language (XML).

In order to maintain record integrity, agencies should follow information technology profession best practices for data preservation. Systems must ensure the following:

- A record creator's identity is verified
- Restrictions exist regarding permissions to read and write files
- Periodic system audits are conducted
- Data transmission includes data error checking and correction
- Data backup occurs regularly
- Data on off-line media are regularly refreshed to avoid loss of data due to degradation

ELECTRONIC RECORDS GUIDELINES

SECTION 2: ELECTRONIC RECORDS MANAGEMENT

Files management is the process of determining how files will be arranged, categorized, accessed, and stored. Efficient electronic filing practices ensure that the right file can be retrieved expeditiously at the right time for the lowest possible cost. Establishing a well-organized filing system involves detailed planning and designing to ensure productive workflow. This includes deciding how files will be arranged and accessed over time, creating classification systems, and selecting the proper format and media.

Electronic Files Management

File naming is an important part of any records management program. A file name is the principal identifier for a record. Especially when dealing with electronic records, it is important to have a unified naming system so that records can be placed in context with other records and series as well as proper state and local Records Retention and Disposition Schedules. Records that are named using a consistent, logical system can be more easily located and shared among users. It is important to create an agency wide file-naming procedure to accompany an electronic records management policy.

In developing a file-naming procedure, include some of the following familiar components:

- Version number (e.g., version 2 [v2, vers2])
- Date of creation (e.g., March 4, 2008 [030408, 03_04_08])
- Name of creator (e.g., Robert B. Pattinson [rbpattinson, rbp])
- 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 May 13, 2007 at 9:00 AM central time [51307_0900ct])
- Publication date (e.g., published on April 24, 2006 [pub042406])
- Project number (e.g., project number 888 [PN888])
- Department number (e.g., Department 110 [dept110])
- Records series (e.g., SeriesX)

Numbers are often used to differentiate similar documents. When numbering documents, use numbering in combination with descriptive names to make file content and purpose clear. Also be sure to use “0”s as placeholders. A sort by name of the following filing names—“committeereport1,” “committeereport2,” “committeereport10,” “committeereport27,” and “committeereport103”—would return the following order:

committeereport1
committeereport10
committeereport103
committeereport2
committeereport27

Whereas using “0”s as placeholders would result in the following order:

committeereport001
committeereport002
committeereport010
committeereport027
committeereport103

Also keep the following in mind while developing a file-naming policy:

- *Access and ease of use*. Since the purpose of file naming is to facilitate use and access, a file-naming policy should also be straightforward and simple.
- *Scalability*. Think about the number of files or file versions needed when determining a naming policy. If a particular group of records grows quickly, do not limit numerical placeholders to two spaces, or only 99 records may be created within the group.
- *Uniqueness*. Name files so that they have unique names regardless of their location. A file entitled “letter_0608” is not independent of location since this letter could pertain to a myriad of record series. Unique names will enable users to avoid the problem of files with the same name causing confusion when they end up in the same storage folder.
- *Universality*. File names should be comprehensible and should make sense to users, not just the person who created the file. Having a common classification scheme, or taxonomy, between users is essential. The benefits of standardization can be recognized when combining multiple users’ files and when integrating different types of electronic records, such as Web, database or word processing files.
- *Version control*. Determine how to manage different versions of the record. Some organizations include a version number in the file name, as shown above. Current and obsolete files with the same name become a problem when these files are transferred to a common storage folder, for example.

E-MAIL MANAGEMENT

E-mail messages—both sent and received—that provide evidence of a government transaction are considered public records. Agency and locality Records Officers must ensure that e-mail is organized for convenient retrieval, maintained, and disposed of in accordance with an approved Records Retention and Disposition Schedule, and accessible as technology is upgraded or changed. The effort to develop and implement an e-mail management policy is the responsibility of each agency or locality and involves a cooperative effort between records management staff, administration, legal counsel, and information technology departments. While IT is necessarily involved in many aspects of records management, such as server maintenance and destruction of backup tapes, creation and dissemination of e-mail management policy is the responsibility of the Records Officer.

Some examples of public record e-mails include policies and directives, correspondence or memos pertaining to the organization's business, work schedules and assignments, documents circulated for approval or comment, and any message that initiates, authorizes, or completes a business transaction, final report, or recommendation. Not all e-mail messages, however, are considered public records. Personal messages or announcements, courtesy or reference copies, routine chat on e-mail listservs, and announcements of social events are all examples of non-record e-mail correspondence. These lists are by no means inclusive.

E-mail messages are subject to the same retention requirements as all other records. This means that there are no set retention periods for e-mails as a format type. Instead, retention periods for e-mail vary according to the information contained within the message as well as the function the message performs. As mentioned above, the *Code of Virginia* §42.1-77 defines records by their content, not their format. E-mail, like paper, is a format. The life cycle of a record is determined by its Records Retention and Disposition Schedule. Often e-mail is considered correspondence, which is covered under General Records Retention and Disposition Schedule (GS-101) for state agencies and General Records Retention and Disposition Schedule (GS-19) for localities. For records that are not classified as correspondence, review the appropriate retention schedule to determine the applicable retention and disposition period.

E-mail records are also subject to the same legal requirements regarding access as other government records. In addition, e-mail records are subject to the same accessibility requirements as any other public record, unless they fall within the exemptions provided under FOIA. Requests from the public for e-mail records must be honored in the same manner as other public records. E-mail records, like all other public records, must remain accessible during their entire retention period and should be maintained in a manner that permits efficient and timely retrieval. Developing a standardized system of document naming

and filing, along with planning for indexing and retrieval points, will assist an agency or locality in maintaining the accessibility of all e-mail messages throughout the required retention period.

Some Virginia state agencies have decided to adopt a proactive and collective approach to e-mail retention by adding a common archive e-mail folder to all of their users' e-mail accounts. This folder contains subfolders, each labeled with a different record type created by that agency. Users can "drag and drop" e-mail messages that are considered public records into these folders on a daily basis. When an agency's e-mail records reach the end of their retention period, they can easily be transferred to the Archives, Library of Virginia, retained by the agency, or disposed of according to the applicable retention schedules.

More frequently, individuals are personally responsible for the records they create, including e-mails received and sent as well as any attachments. Individuals should be aware of any general or specific retention schedules that apply to their e-mail records and they must file, retain, and dispose of these messages accordingly. Many e-mail programs now include an auto-archiving function, which is not recommended for use, as e-mails are not maintained according to an approved Records Retention and Disposition Schedule. The Library encourages users to archive e-mail manually on a regular basis.

When retaining e-mail, any records management system must take into account the importance of metadata. Without this information, much of the original context in which the record was created is lost. Complete e-mail records must include all of the following elements, most of which are included in the Internet mail header:

- Names and e-mail addresses of recipients, including names and addresses of all members of distribution lists
- Name and e-mail address of sender
- Time and date that the e-mail was sent
- Subject line that describes the content of the e-mail
- Text in the body of the e-mail
- Attachments, if applicable

E-mails may be printed and maintained in paper instead of electronic format; however, all of the information above must be included in the paper copy. Often, two users may send a long string of e-mails to one another about a particular topic. In this case, only the last e-mail in a chain must be retained as long as the previous e-mail interactions appear as part of the record. This reduces the number of duplicate records and saves server space.

In addition, it is important for agencies and localities to consider e-mail security and develop procedures to provide security for e-mail so that it cannot be altered or deleted intentionally or unintentionally. E-mail records stored online should be backed up regularly to protect them from system failures, tampering, or deletion. Backup procedures should be coordinated to ensure that no copies of e-mail records are maintained after their retention period expires. Records Officers should work with their IT departments to ensure that all copies of a record are destroyed at the end of its retention period.

Records Analysis Services at the Library of Virginia can provide support and e-mail records training for agencies, including how to identify and manage e-mail messages that constitute public records. For more information regarding identifying e-mail records, refer to the e-mail record decision tree below.

Practical tips for managing e-mail

1. Clean up your e-mail

Agency and locality personnel are responsible for managing e-mails, including sent and received messages. The following are suggestions for managing your e-mail:

- Delete e-mails that do not need to be retained as public records, such as out-of-office responses, meeting announcements, and listserv correspondence. Start by performing a search for the following words or phrases: accepted, declined, tentative, out of office, FYI, or meeting.
- It is not always necessary to retain e-mails in which you are not the primary recipient. If an e-mail is **internally** generated and sent to a primary recipient within the agency, the e-mail should be maintained by the primary recipient. The secondary recipients, that is, those in the CC line, may filter out and delete the e-mail. If an e-mail is sent from a source **external** to the agency and an individual within the agency is copied, that e-mail may be the only copy within the agency and it should therefore be retained.
- Sort messages by sender for easy identification and purge personal correspondence.
- Retain only the final e-mail in a thread, as long as the entire thread is contained in the final e-mail. Be sure any attachments are also saved.
- Permanently delete items in your Deleted Items folder on a regular basis. You may also bypass the Deleted Items folder and delete items permanently in Outlook by highlighting the appropriate messages and holding down the Shift and Delete keys simultaneously.
- File your e-mail regularly. Once you've read and responded to an e-mail, place it in one of your e-mail folders, or delete it according to the appropriate retention and disposition schedule. Remember to file or delete sent items as well as received items.

2. Set up an archiving system

Because of mailbox size limitations in most organizations, e-mails should only be stored within electronic mail systems temporarily. In addition, if this method is in use, e-mail accounts must not be deleted until a supervisor certifies that all public records in the e-mail accounts have been transferred to another record-keeping system or that any retention periods have passed and there is no litigation, audit, investigation, or request for records pursuant to the Virginia Freedom of Information Act (§2.2- 3700 et seq.). There are three methods of archiving e-mail communications outside of the e-mail system:

- Create personal folders that are stored outside of your mailbox but can be viewed using your e-mail client. If you store e-mail on a shared drive, make sure drives are backed up properly on a regular basis.
- Store, access, and manage e-mail messages and other electronic records using an Electronic Content Management system (ECM), such as IBM FileNet.
- Print e-mails and maintain them in a manual system. Include names and e-mail addresses of recipients and sender; time and date sent; subject line that describes the content of the e-mail; text; and attachments, if applicable. If an e-mail is sent to a distribution list, include names and addresses of all members of the list.

Regardless of the chosen approach, make sure folders are named and arranged logically, just as they would be in a paper filing system. Organizing archived e-mail in folders according to records series and fiscal or calendar year facilitates the monitoring of records retention and disposition. In addition, never password-protect an e-mail archive, as doing so may result in retrieval issues in the future.

Keep in mind that there is a difference between archiving and AutoArchiving within Microsoft Outlook. Archiving involves a manual transfer by the user and AutoArchiving is an automatic process that, if turned on within Outlook, takes place at regular intervals. Rather than using the AutoArchiving function, users should set aside time to clean up mailboxes and archive appropriate folders manually.

3. Dispose of e-mail appropriately

E-mail, like all other records, may be disposed of in one of two ways: destruction or transfer to the Archives, Library of Virginia.

E-mail that does not contain confidential or privacy-protected information may be disposed of by deleting it from the e-mail system. For e-mails containing confidential or privacy-protected information, be sure that you electronically shred the e-mail or wipe clean the media on which it resides. E-mail destruction should be reported on a Certificate of Records Destruction (RM-3 Form). Work with your IT department to ensure all duplicate copies, which may reside on servers or backup tapes, are also destroyed.

E-mail may be transferred to the Archives, Library of Virginia on a case-by-case basis by completing the Archival Transfer List and Receipt and Folder List (ARC-1 & -2 Forms)

For more information about e-mail management, please see the E-Mail Management Guidelines.

Web Content Management

Government Web sites contain records that document public transactions just like paper records and, as a result, a Web site must be retained like any other record. Because of the volatile nature of Web sites, however, Web record retention has remained a challenge for archivists and records managers across the country. Static sites are uncommon, especially in government, where policies, procedures, and public notifications posted on Web pages change frequently.

In addition to traditional Web sites, social networking applications such as Facebook, MySpace, and Twitter, and other Web 2.0 technologies, including Web Portals, Really Simple Syndication (RSS), Web Logs (Blogs) and Wikis, are used increasingly for government business. New uses imply different types of content, possibly with different records management considerations. Regardless of the differences in the timeliness, presentation, context, or completeness of information yielded by the various Web applications, Web content may be a record and should be managed as such. The Library of Virginia strongly suggests agencies consider the recordkeeping implications of these tools before implementing them so that they have a plan to manage the content according to Records Retention and Disposition Schedules.

The Library of Virginia's approach to the preservation of Web content, context, and structure combines the need for record retention and disposition policies with the desire to preserve a Web site's format (look and feel) for historical and reference purposes.

An agency's Web records are composed of all publicly available information on the agency's Internet Web site and records on their extranet or intranet, as well as information they may have created on any exterior proprietary Web sites. Each of these sites contains records that need to be managed.

- *Intranet*. An *intranet* is an internal Internet site that is only accessible to persons within an organization.
- *Extranet*. An *extranet* is an intranet site that is accessible only to selected individuals outside an organization.

In order to document a Web site properly, several factors must be taken into consideration and the following elements retained:

- *Content.* The actual HTML-encoded pages themselves and additional content files referenced therein or content created by end users interacting with the Web site. Analyze the pages of the Web site to determine which elements constitute public records.

- *Context*. Administrative and technical records necessary for or produced during the management of a Web site. Maintenance of these records provides a context for Web operations, which attests to the reliability, authenticity, and integrity of an agency or locality's Web site.
- *Structure*: For those Web sites that have been appraised as records, a site map indicating the arrangement of a Web site's content pages and software configuration files of content management systems. Structure also includes technical characteristics of the record (e.g., file format, data organization, page layout, hyperlinks, headers, footnotes).

Consider creating a file-naming protocol for Web pages to help ease management of the site. Having a common taxonomy is important because it maintains consistency between Web, database, and ECM systems. In addition, because Web sites are updated frequently by various people and groups, develop a method for designating and controlling versions. This practice will help ensure that Web site content remains trustworthy.

Content management systems (CMS) can be also used to manage the content of a Web site. A CMS consists of both a content management application (CMA) and a content delivery application (CDA). The CMA can relieve the Webmaster of many of the decisions and actions required to manage the creation, modification, and removal of content from a Web site. A CDA uses and compiles the content management information to update the Web site. A CMS can also be used to create audit trails associated with content that is created directly online.

Traditional records management techniques easily apply to relatively stable contextual and structural Web site records. Managing Web page content is much more complex. Web pages are fluid in nature, and when updates or redesign of Web site maps change the organization of Web content, it may be deemed necessary to set aside a new record-keeping copy of Web site content.

Like e-mail messages, Web sites should be maintained according to Record Retention and Disposition Schedules based on the content they contain rather than their format. A Web site may contain any number of record types, including but not limited to meeting minutes, annual reports, photographs, press releases, maps, organizational charts, policies and procedures, and mission statements, for example. Different formats exist even within Web pages, such as text, image, audio, or video files.

Just as individuals are responsible for maintaining other electronic records according to Record Retention and Disposition Schedules, so are they responsible for ensuring that the information they place on their Web sites is available elsewhere in another format. If a record is only available on a Web site, the Web site is considered the record copy, which must be retained according to the appropriate retention schedules based on the content it contains. Web pages, therefore, are not considered record copies as

long as the information contained within them is retained elsewhere. However, Web sites with original content generated using Web 2.0 technologies, such as public blog comments, Facebook posts, or YouTube video responses, may need to be managed as records if they fit the definition of a public record as defined in the *Code of Virginia*.

Library of Virginia Role in Web Site Archiving

The Library of Virginia collects, preserves, and provides access to all Web sites of Virginia's state government agencies in the executive, legislative, and judicial branches of government as described in this guideline. State government Web sites are selected for inclusion based on intellectual content, research and educational use, and long-term benefit to the citizens of the Commonwealth.

A state government Web site is defined as the collection of all files identified by a state government domain for the purpose of providing publicly available information, affording access to government services, and/or conducting the state's business. Large, complex Web sites may span multiple servers and domains but are unified by Virginia government-related content.

All Web sites selected will be collected and preserved in the formats in which they were primarily distributed to the public. They will be made accessible from the Library of Virginia's catalog and/or Web site, as well as from the Archive-It Web site, a subscription service of the Internet Archive, a non-profit organization founded to build an Internet library offering permanent access for researchers, historians, and scholars to historical collections that exist in digital format.

The Library of Virginia will collect the following Web sites:

- All executive, legislative, and judicial branch state agencies, commissions, and boards as listed in *The Report of Secretary of the Commonwealth* (i.e., the Blue Book)
- All independent state agencies listed in *The Report of the Secretary of the Commonwealth* (i.e., the Blue Book)
- All statewide constitutional officers (e.g., Office of the Governor, Office of the Lt. Governor, and Office of the Attorney General), the Governor's Cabinet Secretaries, and the First Lady
- Gubernatorial initiatives and special projects (e.g., Smart Beginnings, Capitol Square renovations, the Springfield interchange, Jamestown 2007)

Web sites are collected on an established schedule, which is currently monthly for statewide constitutional officers, the Governor's Cabinet, gubernatorial initiatives, and the First Lady. All other Web sites will be crawled quarterly. In the occurrence of significant public safety and health incidents or other

noteworthy events, the Library of Virginia may, at its discretion, alter the crawling frequency of state agency Web sites.

The Library will not collect:

- Public or private college and university Web sites. Due to the size and the access restrictions of these Web resources, and because of existing state Archives practice, these sites will not be archived.
- Non-state agency Web sites. In general, captured Web sites will contain only state government information. Non-state government Web sites may be considered for capture if they contain significant state government information and assist in the formation of government policy.

Technical limitations

The Library of Virginia has partnered with the Internet Archive to collect, preserve, and provide access to the Library's Web archive collections to the best of its abilities via the Archive-It service. Web content is harvested using the Heritrix Web crawler and archived content is indexed and searchable via the Internet Archive's Wayback Machine.

As a general rule, simple, static Web pages are the easiest to archive. Limitations to capturing and playing back archival Web content are as follows:

- When a dynamic page contains forms, JavaScript, images, streaming media, or other elements that require interaction with the originating host, the archived pages might not contain the original site's functionality.
- Database-driven Web sites can be very difficult to harvest. For example, if you need to fill in a form to get access to the content, such as with a search box, the harvester typically cannot retrieve the content.
- JavaScript elements often are hard to archive and even harder to display in the Wayback Machine, especially if they generate relative links (links that do not contain the full address of the linked page).
- Web site owners can specify files or directories to be excluded from a crawl, and can even create specific rules for different automated crawlers. All of this information is contained in a file called robots.txt. The Archive-It tool respects robots.txt. exclusion headers. The Library will make every effort to contact site owners to be sure that they allow the Archive-It crawler to have appropriate access to their site.

- Password-protected sites cannot be accessed by the crawler and therefore will not be archived.
- Links to sites that are not in the same domain as a URL identified for archiving will not be captured. For example, if the Secretary of Public Safety site has a link to the Red Cross, the Red Cross's site will not be captured. However, embedded files are crawled regardless of whether or not they come from an offsite host.

The Archive-It Web crawler does not apply to locality Web sites.

Database Management

Just as e-mail and Web content must be managed according to a Records Retention and Disposition Schedule, information that is housed in databases must also be maintained. Once records in other formats are scanned and entered into a database, the originals may be legally destroyed as reference copies. The database record then becomes the official record copy, which must be retained according to the appropriate retention schedule.

Database Design

Databases must incorporate at least the following features if they are to be properly managed. First, databases must enable the user to take and store files off-line, such as in an inactive table. These inactive files must be retrievable throughout the full duration of their retention period. A technical document should accompany any inactive tables and should contain information about the database, including descriptions of each field, the relationships between data elements, and how the database is supposed to operate.

Second, databases must be able to identify records that have reached the end of their retention period. A database trigger is a stored procedure that is invoked automatically when a predefined event occurs, such as when a record arrives at the end of its retention period. This type of trigger in a date field, for example, should be included in the database to notify users when records are eligible for disposal.

Along with identification of records to be destroyed, a third essential functionality that databases should have is the ability to allow users to extract records from the database for the purpose of disposal. If records cannot be deleted from the database, they will be unnecessarily retained beyond their retention period. Deleted database records, like all records, must have their destruction documented. Keep in mind that electronic shredding for confidential records, however, is not an option in databases. Use the Certificate of Records Destruction (RM-3 Form) to report destruction of any records from a database.

Advantages and Limitations of Using Databases

Information stored in databases can be in either raw data form or uploaded as entire documents. Using databases for storage and retrieval has several advantages and limitations. Advantages include:

- Centralization of information
- Ease and speed of records dissemination and retrieval
- Multiple access points
- Increased security

The following are some of the challenges associated with database management:

- Cost of creating and implementing the database, including equipment, software, and expertise
- Possible cost of maintaining duplicate systems when the database has not fully replaced paper filing
- Preservation challenges

Database records pose many preservation challenges. First, many databases are proprietary. If the records in the database have a long-term retention period, they will eventually need to be extracted from the database structure and either reentered into a nonproprietary database, such as one created by an IT department, or transferred to an inactive table. Second, active databases are fluid by nature. Records are continually added or modified, and these changes may not be tracked. Database backups are often obsolete as soon as they are created, as information changes so quickly. For record-keeping purposes, only the most recent database copy is considered the record copy, but other copies are discoverable for legal battles and FOIA requests.

Electronic Records Management (ERM) Applications

Electronic Records Management (ERM) applications allow organizations to capture, control, store, and dispose of electronic records associated with organizational processes. Using ERM applications, agencies can implement file plans, control dispositions according to approved Records Retention and Disposition Schedules, and access agency records. Some electronic document management (EDM) companies and enterprise content management (ECM) vendors have now added records management capabilities to their products. Conversely, the capabilities of some ERM products have been extended to include many of the functions commonly associated with EDM or ECM products. An ERM product may offer version tracking or workflow tools, in addition to its standard records management functions, for

example. The NARA-endorsed Electronic Records Management Software Applications Design Criteria Standard sets forth baseline functional requirements and identifies non-mandatory features deemed desirable for ERM applications.

VITA awarded a Commonwealth of Virginia statewide contract for ECM software to IBM FileNet in 2007. Functionality available include content management, business process management, e-mail archiving, e-forms, records management, storage connectors, business activity monitors, Web content management, redaction software, and more. This business process software supports the objectives set out in Governor Kaine's "Paperless Government" Initiative through improved digital capture, document storage, and retrieval capabilities. A Shared Services platform also allows smaller to medium-sized agencies that cannot justify their own ECM solution or larger agencies wishing to pilot ECM in a limited fashion to take advantage of this offering.

Electronic Records Destruction

Electronic records may be destroyed only in accordance with a Library of Virginia–approved Records Retention and Disposition Schedule. In addition, custodians of records must ensure that information in confidential or privacy-protected records is protected from unauthorized disclosure. "Deletion" of confidential or privacy-protected information in computer files or other electronic storage media is not acceptable. When a record is "deleted," only the index to the record is actually destroyed instead of the document itself. To be effectively disposed of, electronic records must be overwritten with meaningless data, also known as "electronic shredding," or the storage media on which the records are housed must be physically destroyed. **All** copies of a record (electronic or otherwise) that have reached the end of their retention periods must also be destroyed, including copies on backup tapes and in off-site storage. As with all aspects of electronic records, electronic records destruction necessitates good communication between records managers and information technology personnel.

Methods of Destruction

1. Digital Shredding

Digital shredding is a viable, in-house electronic records destruction method that uses software erasure programs to overwrite files with meaningless data. Limitations of this method include the following:

- Defective hard drives cannot be erased.
- There is no way to ensure visibly that erasure has been successful.
- It is a time-consuming process and subject to human error.

- Erasure tools require upgrades.

2. Degaussing

This method of destruction involves equipment that applies a strong magnetic field to magnetic media that erases all recorded data. As storage devices are typically degaussed one at a time at the rate of one minute per device, degaussing is a much more time-effective destruction method than digital shredding. Limitations of this method include the following:

- There is no way to ensure visibly that degaussing has been successful.
- Equipment creates large magnetic fields that may damage other surrounding devices.
- Different types of media require magnetic fields of varying strengths.
- Degaussing only works on magnetic media, not optical media.

3. Physical destruction of storage media

Often the most popular method of destroying electronic records is to erase or delete the data from the tape, disk, or other storage medium so that it can be reused. Software exists, however, that can extract deleted data even after it has been overwritten. For confidential or proprietary records, the safest method of destruction is to destroy the physical media itself. Depending on the volume of storage media to be destroyed, agencies may want to use outside vendors.
Limitations of this method include the following:

- Shipping records to outside and off-site vendors creates possible security problems.
- Shredded equipment may be difficult to dispose of as it may be considered environmentally controlled waste.

The best option for an agency might also be a combination of two destruction methods.

To find out more about appropriate methods of destruction for electronic records, see VITA's Removal of Commonwealth Data from Electronic Media Standard and data removal information. For more information regarding the legal disposition of electronic records, see §42.1-86.1 of the Virginia Public Records Act and Regulations Governing the Destruction of Public Records Containing Social Security Numbers, 17 VAC 15-120.

ELECTRONIC RECORDS GUIDELINES
SECTION 3: LONG-TERM PRESERVATION

Government entities generate large amounts of electronic records in various formats. Many of these records are useful only for a short period of time, but others are considered archival and thus need to be kept permanently. To ensure the reliability of these records with continuing value over time, establish and implement a preservation plan. Since electronic records create access challenges, preservation plans for electronic records must consider the limitations of storage media, the probability of hardware and software obsolescence, and the ways in which information may potentially be used.

When developing a preservation plan, first conduct a needs-analysis to help guide decision-making. How records are used will aid in determining appropriate preservation options as well as the types of metadata that will be most useful to create with the records.

Decide whether a record must be kept in electronic format or whether there is another cost-effective option for long-term storage. For example, a word processing document could be printed on paper, which might take up more physical space but would not require further migration or conversion. Printing a copy of a Web site, however, would result in the loss of the majority of its functionality. If an agency or locality chooses to retain custody of permanent electronic records, the agency or locality assumes responsibility for maintaining their reliability, authenticity, integrity, and usability.

It is also important to ascertain if access to certain data in the records is restricted by statute, such as the Freedom of Information Act (FOIA) and other state and federal laws. Long-term storage and access policies must address these obligations.

Digital Preservation Techniques

Preservation plans for electronic records must consider the probability of hardware and software obsolescence and guarantee long-term access to records. Proprietary software will eventually become obsolete as companies upgrade or stop producing the product altogether. There are several approaches, some more practical than others, to ensure that electronic records remain useful over time.

One approach is **emulation**. 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. Emulation, however, has so far proved more attractive in theory than in practice. There are few examples of success using this approach, and costs have been shown to be high.

It has a further limitation in that, at best, emulation simply reproduces earlier, less-sophisticated versions of an application.

Another approach to preservation is **encapsulation**. It involves combining the object to be preserved with all of the necessary details on how to interpret it within a wrapper or package, all 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 "media refreshing") or computer platforms in order to maintain their value. **Conversion** entails changing files from one format to another and may involve moving from a proprietary format, such as Microsoft Word, to a nonproprietary one such as a plain text file or XML. To avoid losing data in the process, testing and analysis should determine exactly what changes will occur and whether they are acceptable. With both migration and conversion, special attention must be paid to maintaining the accessibility of associated metadata. When properly planned and executed, the migration and conversion approaches represent the easiest and most cost-effective preservation methods available today. A discussion of file format and media options to consider when migrating or converting records follows this section.

One challenge of converting files is that data may be lost. Data compression is the process of encoding information using fewer bits. Compression saves storage space and enables data to be transmitted quickly and easily, but data may be lost as a result. Compression also introduces an additional layer of software dependency. Compression options vary in their degree of data loss. Some are intentionally “lossy," such as the JPEG format, which relies on the human eye to fill in the missing detail. Others are designed to be "lossless."

There are three basic types of loss that may occur during conversion or migration:

- *Data*. If data is lost, the content of the record is lost to a varying degree.
- *Appearance*. Converting records may alter the formatting of the file. For example, converting all word processing documents to RTF may cause some loss of page layout. Determine whether this loss affects the completeness of the record. If the structure is essential to understanding the record, this loss may be unacceptable.

- *Relationships*. Relationships within the data in a file, such as spreadsheet cell formulas and database file fields, may be lost.

File Format Options

File formats quickly become obsolete as technology continually advances, making long-term preservation strategies, such as reformatting, vital. Proper advanced planning can mitigate risk and ensure that legal standards are upheld and operational requirements are met.

A file format is often described as either proprietary or nonproprietary:

- *Proprietary formats.* Proprietary file formats are controlled and supported by just one software developer. Proprietary formats, such as Microsoft Word files and WordPerfect files, carry the extension of the software in which they were created.

- *Nonproprietary 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 a popular nonproprietary format.

The software in which a file is created usually has a default format, often indicated by a file name suffix, such as PDF for portable document format. Most software programs allow creators to select from a variety of formats in which to save a file, including document (DOC), Rich Text Format (RTF), and text (TXT). Some software, such as Adobe Acrobat, is designed to convert files from one format to another.

Basic file format types include the following:

1. **Text files** are most often created using word processing software. Common file formats for text files include Rich Text Format (RTF) files and Portable Document Format (PDF) files.

2. **Graphics files** store images, such as photographs and drawings, and are divided into two basic types:

    A. Vector-based files store images as geometric shapes in mathematical formulas, which allow images to be scaled without distortion.

    B. Raster-based files, also referred to as bitmapped images, store images as a collection of pixels and cannot be scaled without distortion.

3. **Data files** are created in database software programs. Data files are divided into fields and tables that contain discrete elements of information. For example, a customer service database may contain customer names, addresses, and billing history fields. These fields may be organized into separate tables. Data files can be converted to a text format, but relationships among the fields and tables may be lost. Converting data files to the Comma Separated Value (CSV) file format allows aspects of the formatting and field names to be preserved.

4. **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 totals two other cells. Like data files, spreadsheet files are most often saved in the proprietary format of the software program in which they were created. Spreadsheet files can be exported as text files, but the value and relationship of the numbers are lost. Converting data files to the Comma Separated Value (CSV) file format allows aspects of the formatting and field names to be preserved.

5. **Video files** contain moving images, such as digitized video and animation. These files are most often created and viewed in proprietary software programs and stored in proprietary formats. At this time, best practices in the preservation of video leans towards storage on analog or digital videotape and a consistent migration plan to refresh the media.

6. **Audio files** contain sound data. These files are frequently used in customer service environments to capture audio telephone recordings as well as to track key strokes of customer service representatives, allowing these conversations to be re-created at a later time. Audio files in mp3 or other compressed formats should be converted to a preservation format. Broadcast wave format (.WAV) is considered a type of preservation format for audio files.

7. **Markup languages**, also called markup formats, contain embedded instructions for displaying or understanding the content of a file. The World Wide Web Consortium (W3C) supports these standards. Currently, eXtensible Markup Language (XML) is the most favorable format choice for long-term preservation and use of electronic records. XML, an international standard since 1998, is a human-readable, self-describing markup language that is independent of hardware and operating systems. Because of its infrastructure-independent quality, XML is a great solution for refreshing and sharing record content. In order to use and benefit from XML, agencies must plan for certain up-front costs and time expenditures. Its structured nature, however, makes XML suitable for eventual automation and will enable the use of future open formats.

Table 1: Common File Formats

File Format Type	Common Formats	Sample Files	Description
Text	PDF, RTF, TXT, proprietary formats based on software (e.g., Microsoft Word)	Letters, reports, memos, e-mail messages saved as text	Created or saved as text (may include graphics)
Vector graphics	DXF, EPS, CGM	Architectural plans, complex illustrations	Store the image as geometric shapes in a mathematical formula for undistorted scaling
Raster graphics	TIFF, BMP, GIF, JPEG	Web page graphics, simple illustrations, photographs	Store the image as a collection of pixels that 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, MPEG	Short video to be shown on a Web site, recorded interview to be shared on CD-ROM	Contain moving images and sound
Markup languages	SGML, XML, HTML, XHTML	Text and graphics to be displayed on a Web site	Contain embedded instructions or "tags" used to transmit and display the content of a file or multiple files

File format decisions may affect electronic records management in the following ways:

- *Accessibility*. The file format must enable users to find and view the record. Records cannot be in a format that is highly compressed and easy to store if that format makes the record inaccessible.
- *Longevity*. The file format should be supported for the long-term. Proprietary software developers may not be able to ensure long-term support, thus increasing the risk of records' becoming inaccessible.
- *Accuracy*. Converted records should retain all the significant detail of the originals. The converted file should minimize data, appearance, and relationship loss.
- *Completeness*. Converted file formats should meet operational and legal objectives of existing standards for an acceptable degree of data, appearance, and relationship loss.
- *Flexibility*. The file format must meet objectives for sharing and using records. If the file format can only be read by specialized hardware and/or software, the ability to share, use, and manipulate records is limited.

Storage Options

According to the Virginia Public Records Act of the *Code of Virginia* § 42.1-85, agencies are responsible for converting and migrating electronic records "as often as necessary so that information is not lost due to hardware, software, or media obsolescence or deterioration." Both agencies and localities must ensure access to electronic records for the entire length of their retention period. This means that users must be able to find, open, and read all records throughout their lifetime. Consider digital storage options that enable accessibility through migration and/or conversion of records throughout their required retention period.

In order to determine the best long-term storage medium for records, examine the current volume of stored records, along with the size of the record files and any metadata associated with them. Next, estimate projected record volume, and take into account any data access and security requirements.

There are three basic records storage options:

- *Online storage*. Records are available for immediate access and retrieval. Online storage devices include mainframe storage and network-attached storage. Online storage provides the fastest access and regular integrity checks.
- *Nearline storage*. Records are stored on media such as network-attached storage, optical disks in jukeboxes, or tapes in automated libraries. Nearline storage provides faster data access than off-line storage at a lower cost than online storage.
- *Off-line storage*. Records are stored on removable media such as magnetic tape or optical disk. Because human intervention is necessary, this option provides the slowest access.

Vital, long-term, or archival electronic records should be stored utilizing online or nearline storage options. The advantages of online and nearline storage include large storage capacities and the opportunity for data replication. Off-line storage devices are not recommended for record copies of vital, long-term, or archival records, as they are less likely to be routinely accessed and are often overlooked when systems are upgraded and electronic records are migrated to new formats. Off-line storage is recommended for backups or security copies, however, as the records can be stored off-site.

Types of Digital Storage Media

All storage media have finite life spans that are dependent on a number of factors, including manufacturing quality, age and condition before recording, handling and maintenance, frequency of

access, and storage conditions. Under optimal conditions, the life expectancy of magnetic media ranges from 10 to 20 years, while optical media may last as long as 30 years. In less than ideal conditions, however, media life expectancies are significantly less.

The storage capacity of digital media is measured in bytes, the basic unit of measurement:

1,024 bytes	=	1 kilobyte (KB)
1,024 KBs	=	1 megabyte (MB)
1,024 MBs	=	1 gigabyte (GB)
1,024 GBs	=	1 terabyte (TB)
1,024 TBs	=	1 petabyte (PB)
1,024 PBs	=	1 exabyte (EB)

Access to digital information on digital media is divided into two types:

- *Sequential*. Sequentially ordered digital media requires the user to access preceding information in order to arrive at a specific point. For example, to view a specific portion of a videotape, a user must first fast-forward through the preceding portion of the videotape.
- *Random*. Some digital media allow users to access the stored information from any physical place on the media. For example, users can access any single file stored on a computer disk without having first to access all the files that precede it.

Digital media are divided into three types:

1. **Magnetic:** Electronic information is stored on computer drives, disks, or tapes by magnetizing particles imbedded in the material. Magnetic media include:

    A. *Magnetic disks*, such as computer hard drives that store programs and files, are randomly accessed. Fixed disks reside permanently in a drive while removable disks are encased in plug-in cartridges, allowing for storage and transfer of data.

    B. *Magnetic tape* is a sequential storage medium used for data collection, backup, and archiving. Common magnetic tape formats include Digital Audio Tape (DAT), Digital Linear Tape (DLT), and Linear-Tape Open (LTO).

2. **Optical:** Digital data is encoded by creating microscopic holes in the surface of the medium. Optical media options include:

A. *Compact Discs (CD)* can be read-only (CD-ROM), write once read many (CD-R), and rewritable (CD-RW). CDs can hold roughly 700 MB of data.

B. *Digital Versatile Discs (DVD)* are also called digital video discs. The data they store do not have to be in video form, however. DVDs can hold between 4.7 GB and 17.0 GB of data. Common types of DVDs include:

- *DVD Random Access Memory (DVD-RAM)* is a rewritable disc that provides 4.7 GB per side storage capacity.
- *DVD-R* has the same storage capacity as DVD-RAM, but can only be written to one time.
- *DVD+R* is a writable disc with 4.7 GB of storage capacity on either side.
- *DVD-RW* offer 4.7 GB per side, but can be overwritten 1,000 times. The DVD-RW technology is mainly used for video.
- *DVD+RW* is an alternative rewritable format that has a capacity of 4.7 GB per side and is used for both data and video content.

3. **Solid state:** With no moving parts, a solid state device uses electronics instead of mechanics. These devices are much faster and more reliable than magnetic and optical media. Solid state devices include:

A. A computer's BIOS (Basic Input/Output System) chip
B. PCMCIA Type I and Type II memory cards, which are used as solid-state disks in laptops
C. Flash and Dynamic Random Access Memory (DRAM-based) solid state drives
D. CompactFlash, SmartMedia, or Memory Stick, which are most often found in digital cameras

When choosing digital storage media, each option's performance characteristics must be evaluated in relation to the users' records management needs. Consider:

- How quickly users need to access the records. Some types of records require quick retrieval, while others do not.
- The volume of records that can be stored on the medium. Examine the current volume of the records and try to determine future needs.
- How long the industry will support various media options and compare those figures with the time period that records must be kept according to the approved records retention schedule. A

medium that meets many needs but is not widely used or has a high risk of becoming obsolete has limited long-term value.

- How easily a given medium can be damaged or will deteriorate. A medium that deteriorates after three years might be a suitable option for records that need to be retained for only one year.
- The types of file formats a medium can store. For example, a floppy disk cannot store large graphics files, but a CD or a DVD can store graphics, text, audio files, and video files.
- The backward and forward compatibility of the digital media. A backward compatible component retains the functionality of an older component. Forward compatibility refers to the ability of the media to read information created for later versions. DVD-ROM drives are backward-compatible to CD-ROMs, but a CD-ROM drive is not forward-compatible to DVD-ROMs. This assessment will help determine how often to upgrade supporting systems, migrate and/or convert records.
- Costs and benefits of each medium, including the cost of converting and/or migrating records.

Permanent Electronic Records in the Archives

The Archives and Records Management Services Division (ARMS) of the Library of Virginia will accept the transfer of permanent electronic records from state agencies and localities on a case-by-case basis. Agencies and localities are responsible for all permanent public records created and maintained during the course of business, regardless of the record's format. With assistance from ARMS staff, each organization will establish appropriate methods to maintain the authenticity, integrity, and accessibility of these records until their official transfer to the permanent Library collection. ARMS staff highly recommends that organizations create and implement a records policy, as well as train staff on the proper methods for maintaining public records.

The Archives at the Library of Virginia will make every reasonable effort to take in all permanent records that have reached their stated disposition. However, as a result of limited capacity to administer electronic records at this time, the Library initially may be unable to accept all electronic records designated as permanent. In those instances, agencies and localities will remain responsible for ensuring that the essential evidence contained in the records is preserved until the Library is able to accommodate them. Records should not be transferred until official notification is received from ARMS staff.

If an agency or locality chooses to retain custody of any record, including permanent electronic public records eligible for transfer to the Library, it will assume the responsibility for maintaining that record in a manner that preserves its authenticity and integrity, and ensures its continued accessibility. Records stored electronically require a migration plan to ensure that access is not interrupted by degradation or by hardware, software, or media obsolescence. Please contact your designated Records Management Analyst at the Library of Virginia for additional guidance.

Transferring Permanent Electronic Records to the Library of Virginia

If records have been designated as archival by Archives and Records Management Services (ARMS) staff and will be added into the Library of Virginia's permanent collection, follow these procedures for transfer. These procedures do not include electronic back-up tapes or vital information created and maintained for internal use only. For more information and step-by-step instructions, refer to the chapter entitled "Archival Transfers and Procedures" in the Records Management Manual.

1. Review appropriate General and Specific Records Retention & Disposition Schedules and locate the series for the records. The schedule will indicate which records series have a permanent retention and may be or are required to be transferred to the Archives.

   Please note: Due to a limited capacity to administer permanent electronic records at this time, the Archives and Records Management Services Division will evaluate each permanent records transfer prior to approval. If approved, the Records Analyst will make arrangements for transfer of the records to the Library of Virginia.

2. Upon receiving transfer approval, send a completed Archival Transfer List and Receipt and Folder List (ARC-1 & -2 Forms) to your assigned LVA Records Analyst. Although most electronic transfers will be on physical media (e.g. tape and/or optical disc), under certain circumstances other transfer options may be available.

   Physical media can be transferred to the Archives similarly to traditionally formatted records. Arrangements for delivery or pickup will be made by the LVA Records Analyst. Be sure to remove all password protection applied to any files prior to copying to the transfer media.

   The LVA recommends the following media be used for transfer of electronic records:

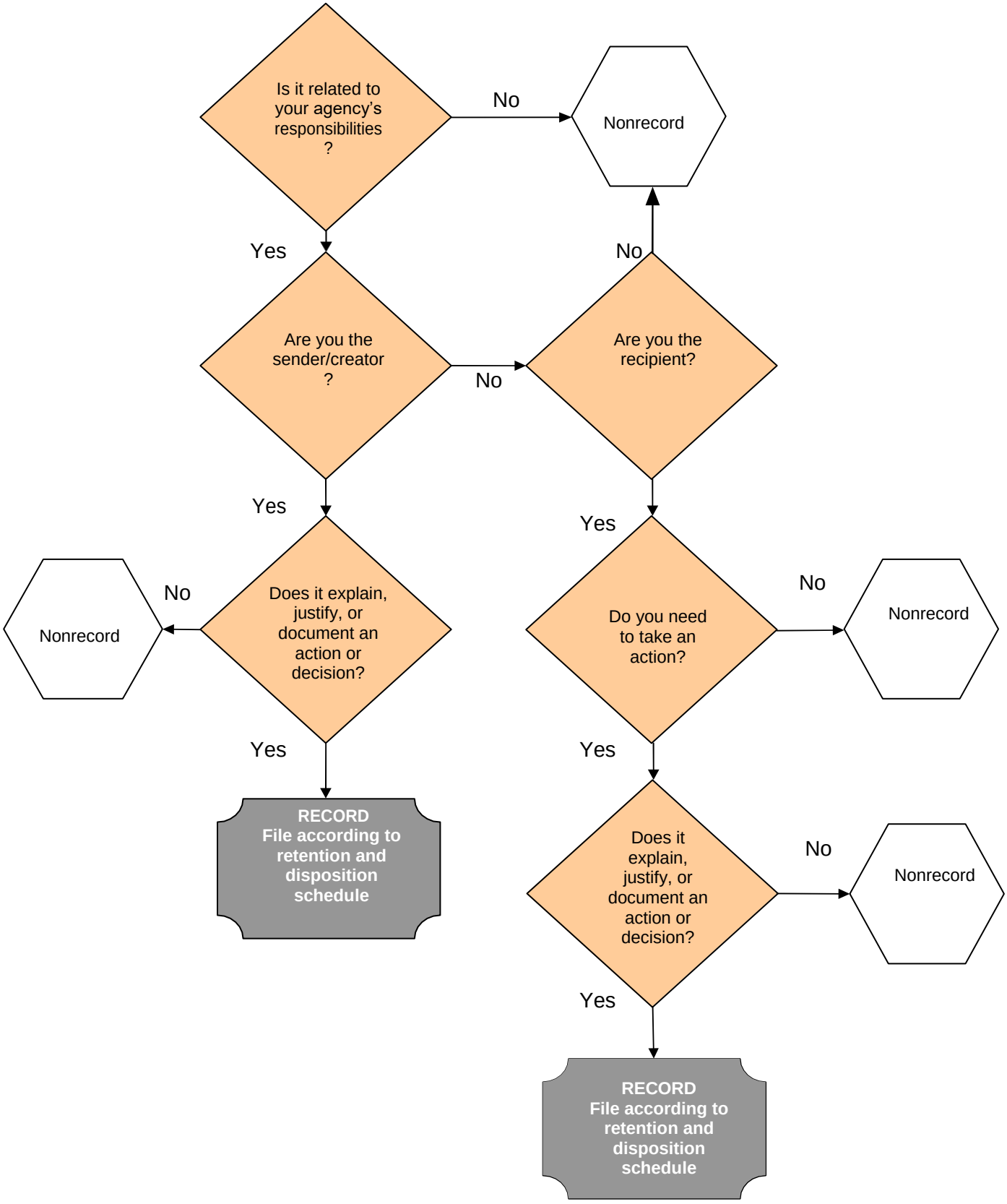
   - **Tape** – Linear Tape Open (LTO) and Digital Linear Tape (DLT) are preferred. Digital Audio Tape (DAT) should not be used.
   - **USB Flash Drive** – Flash memory may be used as temporary storage to transfer files to the LVA.
   - **Removable Hard Disk Drive (HDD)** – Removable HDD is an option when transfers are significant in size.
   - **Optical Media (DVD-R/CD-R)** – Use high quality optical disks (e.g. Taiyo Yuden, MAM-A or equivalent) to store and transfer electronic records. Records storage on lesser quality discs may result in a corrupted or incomplete records transfer. Please note that the Archives cannot accept HD-DVD or Blue-Ray DVD at this time.

Records custodians will remain responsible for the security of the media and the information contained on them as defined by Virginia Information Technology Agency (VITA) policy and standards until an official transfer to the Library is complete. See the IT Information Security Standard (SEC501-01) for additional information.

3. Following the initial transfer of records to the LVA, all electronic files will be quarantined and inspected. Do not delete or destroy any electronic files identified as permanent prior to official notification from the LVA indicating that the inspection process has been concluded and a re-transfer of the files is not required.

ELECTRONIC RECORDS GUIDELINES

APPENDIX A: E-MAIL RECORD DECISION TREE



URLs updated 2018-08

ELECTRONIC RECORDS GUIDELINES

APPENDIX B: GLOSSARY

Archival quality: A quality of reproduction consistent with established standards specified by state and national agencies and organizations responsible for establishing such standards, such as the Association for Information and Image Management, the American National Standards Institute, or the National Institute of Standards and Technology.

Archival record: Material created or received in the conduct of affairs that is preserved because of the enduring historical value or as evidence of the roles and responsibilities of the creator.

Backup: A copy of an electronic record maintained to protect information loss.

Bits: Short for binary digits, the smallest unit of information in a binary system. Each bit is assigned a 1 (high current) or a 0 (low current), where eight bits make up a byte.

Conversion: The act of moving electronic records to a different format, especially data from an obsolete format to a current format.

Compression: A computer process using algorithms that reduces the size of electronic files so that they occupy less digital storage space.

Data compression: Any of several techniques that reduce the number of bits required to represent information in data transmission or storage.

Digital image: See Image

Digital imaging: See Imaging

Disposition: Action to be taken on a records series at a specific time. May entail destruction, reformatting, transfer, or permanent retention.

Enterprise content management (ECM) system: Technologies used to capture, manage, store, preserve, and deliver content and documents related to organizational processes. Electronic Document Management (EDM) and Electronic Records Management (ERM) are components of ECM.

Electronic document management (EDM) system: Software that controls the capture, indexing, processing, storing, transferring, and use of electronic documents to facilitate workflow. Manages documents as individual units, as opposed to preserving relationships to larger groups of documents that provide evidence of the same particular organizational function.

Electronic records management (ERM) system: Software that enables the capture and management of electronic documents as records. Typical ERM functions include declaration, capture, organization, security, retrieval, preservation, audit/oversight, and disposition.

Electronic shredding: The process of overwriting data instead of merely deleting it. Involves overwriting the file's data clusters, renaming the file with a randomly generated name, truncating the file to 0 bytes in length and deleting the renamed and truncated file.

File format: A specification for organizing data. Digital images and their associated metadata may be presented in a number of formats depending on compression schemes, intended use, or interoperability requirements. Some image formats are broadly decipherable, while others may only be accessible to certain application programs.

Identifying (or privacy-protected) information: According to §18.2-186.3 (C) of the *Code of Virginia*, identifying information includes social security numbers, driver's license numbers, bank account numbers; credit or debit card numbers, personal identification numbers (PIN), electronic identification codes, automated or electronic signatures and passwords.

Image: A graphic representation of an object. More specifically, a raster-based, two-dimensional, rectangular array of static data elements called pixels, intended for display on a computer monitor or for transformation into another format, such as a printed page.

Image compression: The application of data compression on digital images.

Imaging: The process of electronically capturing the visual appearance of (usually) paper documents, also called “scanning” or “digitizing.”

Index: Descriptive data associated with an image for retrieving that specific image from storage.

Legacy system: An existing computer system that must be accommodated when building new systems.

Life cycle: The creation, use, maintenance, and disposition of a public record.

Lossless compression: Reduction in file size without loss of information, achieved by storing data more efficiently.

Lossy compression: Reduction in file size that involves permanent loss of information. Algorithms selectively discard data in order to attain a greater size diminishment than is possible with lossless compression.

Magnetic media: Tape or disk coated with a magnetic surface used for storing electronic data.

Master image: A faithful digital reproduction of a document optimized for longevity and for production of a range of delivery versions.

Metadata: Data describing the context, content, and structure of records and their management through time.

Migration: The process of moving records from one hardware and/or software platform to another.

Nonproprietary: A format that is not owned by a private individual or corporation under a trademark or patent. It is in the public domain and is easily portable between various hardware and software systems.

Optical media: A data storage medium that utilizes laser technology to read information.

Pixel: Short for picture elements, which make up an image, similar to grains in a photograph or dots in a halftone. Each pixel can represent a number of different shades or colors, depending on how much storage space is allocated for it. Pixel size, frequency, and color determine the accuracy with which photographic images can be represented.

Proprietary: A format that is owned by a company or a private owner. Some proprietary formats are published and protected by intellectual property rights or copyright. Other proprietary formats are not published.

Public record: Recorded information that documents a transaction or activity by or with any public officer, agency, or employee of an agency. Regardless of physical form or characteristic, the recorded information is a public record if it is produced, collected, received, or retained in pursuance of law or in connection with the transaction of public business. The medium on which such information is recorded has no bearing on the determination of whether the recording is a public record.

Quality assurance: The process by which the total product is examined to ensure that the quality criteria initially established in the preproduction test has been met.

Quality control: Techniques to ensure accuracy and high quality through various stages of a process.

Record copy: An original, official, or master record that is distinct from a "working" or "convenience" copy, which is a duplicate used for reference purposes.

Records analysis: The examination and evaluation of systems and procedures related to the creation, processing, storage, and disposition of records.

Records Retention and Disposition Schedule: A Library of Virginia–approved timetable stating the required retention period and disposition action of a records series. The administrative, fiscal, historical, and legal value of a public record shall be considered in appraising its appropriate retention schedule.

Refresh: The process of periodically moving records from one storage medium to another.

Resolution: The measure of the quality of a digital image, usually expressed in dots per inch (DPI).

Retention period: The length of time a record is kept.

Scanning: See Imaging

Solid state media: A data storage medium that uses solid-state memory with no moving parts.

Standards: Rules typically developed, adopted, and promoted by large organizations that can advocate for their broad usage. Data standards enable the exchange of data while technology standards enable the delivery of data between systems.

Taxonomy: System or technique of classification.

Vital record: A record essential to the operation of an organization and/or resumption of operations following a disaster.

Workflow analysis: The examination and evaluation of the tasks, procedural steps, staff involved, required input and output information, and tools needed for each step in a business process.