Office of the Public Records Administrator
State Archives
Connecticut State Library
State of Connecticut

Public Records Standards 04-1: Electronic Records

Standards effective December 1, 2022 and supersede *Digital Imaging Standards, 2014*.

Approved scanning projects prior to this date shall meet the previous standards document except records designated as permanent, archival or life of structure on a records retention schedule.

These standards shall be read together with *Public Records Policy 04: Electronic Records Management* to ensure a full understanding of the Office of the Public Records Administrator (OPRA) and State Archives (SA) joint policy regarding (1) the preservation and authentication of electronic records, and (2) the retention and disposition of electronic records.

These standards shall be read together with *Public Records Standards 04-2: Digital Imaging* to ensure a full understanding of the Office of the Public Records Administrator (OPRA) and State Archives (SA) joint policy regarding the use of digital imaging technology for the reformatting of analog public records; and regarding the retention and disposition of original and digitized records.

These standards apply to records of public agencies (hereinafter "public records/records") that are (a) born digital or (b) digitized images of analog records (both of which are hereinafter "electronic records"). Public agencies are defined as executive branch state agencies, as defined in C.G.S. § 4-5; certain quasi-public agencies; towns, cities, boroughs, and districts; and other political subdivisions of the state (hereinafter "public agency/agencies").

Public agencies should work in conjunction with appropriate IT staff, either individually or through central IT (where applicable), to implement information systems that are compliant with the below standards.

Public agencies must establish a clear and sustainable plan for maintaining long-term electronic records and dedicate sufficient resources to this plan. Electronic records require proactive attention as they are more fragile and complex to preserve than paper and microform records. Without preservation actions, electronic records can be overwritten in databases, lost in media migrations, or become inaccessible due to incompatible legacy systems. Public agencies must be aware of the new skill sets, training, considerable significant resources, and ongoing management that will be required over many decades to ensure that the electronic records remain available to future generations.

For the purposes of this document, the term "shall" or "must" indicates a requirement and the terms "should" and "may" indicate a recommendation or best practice.

I. Legal Issues

A. Any public agency contemplating using digital imaging technology for the reproduction of public records shall be aware of all applicable statutes or regulations and any legal issues. Consultation with appropriate legal counsel regarding rules of evidence and any other legal issues is advisable.

B. References to electronic records can be found in many sections of the *Connecticut General Statutes*, including but not limited to, sections contained within Chapter 3, *Public Records: General Provisions;* Chapter 14, *Freedom of Information Act*; Chapter 15, *Connecticut Uniform Electronic Transactions Act*; Chapter 15b, *Uniform Electronic Legal Material Act;* Chapter 92, *Town Clerks;* and Chapter 899, *Evidence.*

II. Authenticity of Records

Public agencies must establish and maintain procedures to ensure authenticity of electronic records during creation and maintenance of information systems. Authenticity procedures should be designed to show that the record is unaltered from the original throughout the duration of the life of the record, including but not limited to:

1. Documentation showing that similar kinds of records generated and stored electronically are created by the same processes each time and have a standardized retrieval approach.
2. Security procedures that prevent unauthorized addition, modification, or deletion of a record and ensure system protection against such problems as power interruptions or natural disasters.
3. Identification of electronic media on which records are stored throughout their life cycle, the maximum time span that records remain on each storage medium, and the Office of the Public Records Administrator-approved disposition process for all public agency records.
4. Chain of custody detailing information on a record's lifecycle from its original creation version to its final production version to verify that the agency or vendor have not altered information either in the copying process or during analysis.
5. Checksums to detect if the contents of a file have been corrupted or changed; or audit trails to link specific records in a system and track such information as the user, date and time of event, and type of event (data added, modified, deleted, etc.).
6. Evidence that no manipulation, substitution, or falsification occurred after record creation.

III. System Trustworthiness

A. To ensure the trustworthiness of any information system, the public agency shall in consultation with IT create policies and procedures defining the normal operations and use of such systems. These written policies and procedures shall be kept up to date, be quickly accessible if needed for training and legal situations, and include the following:

1. an overview of the system that describes the purpose and uses of the system;
2. the methods used to create, modify, duplicate, transfer, and destroy records;
3. the roles and responsibilities of those individuals involved in electronic records creation, maintenance, and destruction; and
4. the systems in place to ensure consistent quality control and problem resolution.

B. The public agency shall develop and establish policies and procedures for training and support that include instructions for imaging, indexing, quality control, and retrieval, and that document user training relating to the use of the system.

C. The public agency shall implement checksums and/or audit trails to verify no unauthorized deletions, additions, or changes have entered the system and that support the public agency's ability to identify the source of any such unauthorized action.

D. Ensure the system employed includes performance assurance processes that routinely test the hardware and software and document system testing and performance issues.

E. The system shall include security protocols that limit system access and update privileges to appropriate users, prevent unauthorized modification of records, and include disaster preparedness and security backup procedures.

F. All security controls required by regulation, policy, and/or law for paper records shall be addressed for electronic records unless those regulations, policies, and/or laws state otherwise. The public agency shall ensure the protection of records that contain confidential or sensitive information.

IV. Protection of Confidential Information.

A. Public agencies and vendors have a duty to protect and shall protect against a confidential information breach all confidential information which they come to possess or control, wherever and however stored or maintained, in accordance with current industry standards.

B. Vendors shall develop, implement, and maintain a comprehensive data security program for the protection of confidential information in agency records. The safeguards contained in such program shall be consistent

with and comply with the safeguards for protection of confidential information, and information of a similar character, as set forth in all applicable federal and state law and written policy concerning the confidentiality of information. Such data-security program shall include, but not be limited to, the following:

1. a security policy for employees related to the storage, access, and transportation of data containing confidential information;
2. reasonable restrictions on access to records containing confidential information, including access to any locked storage where such records are kept;
3. a process for reviewing policies and security measures at least annually;
4. creating secure access controls to confidential information, including but not limited to passwords; and
5. encrypting confidential information that is stored on laptops, portable devices, or being transmitted electronically.

C. Vendors and vendor parties shall notify the public agency and appropriate legal counsel as soon as practical, but no later than twenty-four (24) hours, after they become aware of or suspect that any confidential information which vendor or vendor parties have come to possess or control has been subject to a confidential information breach.

V. Documentation for Operating Environments:

Public agencies either individually or through central IT (where applicable) should maintain documentation about operating environments used for information systems including but not limited to hardware, operating systems, configurations, associated service packs, common applications and their associated updates.

VI. Digital Preservation Systems

A. State agencies are required to contact and work with the State Archives for inclusion in the Connecticut Digital Archive (CTDA), the State Library's digital preservation system. Executive branch state agencies and quasi-public agencies shall not directly contact the CTDA.

B. Executive branch state agencies and quasi-public agencies shall consult and coordinate with the State Archives and statewide information technology prior to creating project requirements for digital preservation.

C. Local public agencies (towns, cities, boroughs, and districts) shall consult with the State Archives prior to implementing a digital preservation system and follow the requirements listed in section D below.

D. Digital preservation systems used by public agencies must meet the following requirements:

1. Follows the ISO 14721:2012 Open Archives Information System (OAIS) – Reference Model for digital asset preservation and repository construction.
2. Maintains necessary back-ups.
3. Monitors the performance of hardware and software and responds to infrastructure issues.
4. Creates a PREMIS (PREservation Metadata Implementation Strategies) metadata record to document the audit trail and chain of custody of each digital object in the digital preservation system.
5. Maintains a disaster recovery plan.
6. Supports online bulk ingest of records and supports the bulk transfer via portable media.
7. Accommodates a wide variety of file formats.
8. Automates as much electronic records processing work (i.e., virus scans, checksum generation, extraction of preservation and administrative metadata) as possible.
9. Ensures that each record has metadata sufficient to ensure its authenticity, support ongoing preservation, and facilitate access.

10. Maintains sufficient metadata required to search and understand the content, context, and structure of the records.
11. Facilitates migration or conversion to preservation-friendly formats when appropriate.
12. Safeguards against unauthorized access.
13. Ensures that records remain uncorrupted.
14. Creates and stores geographically redundant copies.
15. Supports searching and access by users.
16. Packages records in ways that facilitate their movement to next-generation preservation systems.
17. Minimizes the overall cost of preserving electronic records of enduring value.

VII. Maintenance

A. Electronic records and metadata shall be effectively and efficiently managed throughout the designated retention period.

B. The cost of ongoing maintenance and sustainability of information systems should be factored into the public agency's budget.

C. *Data Integrity:* Records shall be protected against file corruption, alteration, or deletion throughout the designated retention period. The public agency shall have policies and procedures in place to ensure the integrity of electronic records.

1. Records shall be checked regularly for integrity according to public agency policy, such as using a disk-error checking utility which is built into most operating systems such as Microsoft Windows.
2. When data is written to a storage medium, an error-checking value called a checksum is computed and written along with the data. Any time the data is read, the checksum is recalculated and compared to the stored value to verify that the data on the disk was written and read correctly.

D. *Migration*

Due to the technological advances and the potential obsolescence of technology currently in place, the public agency shall plan for future migrations to new media and systems. Storage media often become obsolete and are replaced with new technology before the end of their life expectancy. If a system stores records with retention periods exceeding the lifespan of the hardware and software in use, it becomes essential to plan for future data migration. To ensure the contents of the media remain accessible, the public agency shall migrate all electronic records and their associated metadata to a newer media platform as needed.

1. The public agency in consultation with IT should establish a migration plan *before* imaging or adoption of new media and information systems. This plan should be reviewed periodically. The reality of technological obsolescence requires that the public agency monitor technology trends and industry developments to ensure their records are accessible over the required retention periods of the systems on which they are stored.
2. The cost of migration should be factored into the public agency's budget, as migration is an ongoing expense that may grow substantially with time depending on the storage medium.
3. In some instances, it might be advisable to maintain electronic records in multiple file formats (i.e., the original Microsoft Word document and also in PDF) to avoid loss of information, understanding, or use.

VIII. Cloud Storage

A. Regardless of which cloud service providers and deployment models are adopted, public agencies are still required to manage their records in accordance with records retention schedules pursuant to C.G.S. § 11-8a. Variations among cloud service providers and deployment models, however, will affect how and by whom (agency/contractor) records management activities can be performed.

B. Cloud providers may not be able to easily meet the type of security and/or information access controls that satisfy local, state and/or federal regulation. Cloud providers may desire to charge additional costs to cover unique or specialized security requirements. Further, a public agency may have little control over who among the cloud provider's employees is authorized to access public agency data.

C. Storage of data outside of the physical and/or legal boundaries of the state may compromise the public agency's ability to manage and control its data.

D. Any data sourced to a cloud services provider shall remain the legal property of the public agency and this must be clearly articulated in any agreements with the provider.

E. The public agency must include in any agreement with a cloud provider an exit strategy to protect its data in the event the storage contract or relationship is terminated with the cloud provider including but not limited to receiving data in a usable format.

F. State agencies must obtain written approval of the state's Chief Information Officer before committing to the services of a cloud vendor.

G. For guidance and best practices refer to the National Institute of Standards and Technology (NIST) Special Publication 800-145 *The NIST Definition of Cloud Computing*.

IX. Transfer and Storage Media

A. *Transfer Media:* Transfer media is intended only for short-term storage or while moving electronic records and index data from the source (such as a vendor hired for imaging services) to the public agency's records storage or to the Connecticut State Archives.

1. For the purposes of transferring electronic records external hard drives, SFTP, VPN, or USB-drive media are preferred.
2. The use of transfer media shall not be permitted for the long-term storage of electronic records because of media instability and fragility.

B. *Storage Media*: Storage media is intended for long-term storage of electronic records.

1. Any storage media used shall comply with the applicable International Standards Organization (ISO) standards.
2. Storage media should be kept in a secure, dust-free area under proper environmental conditions.
3. Electronic records and their associated metadata are best stored on server-class hard drives utilizing a RAID (Redundant Array of Inexpensive Discs) configuration and/or mirroring across geographically separated data centers (geo redundancy). RAID 5 or higher is typically the preferred configuration to ensure proper protection and availability in the event of a disc failure.
4. If use of RAID 5 or other RAID level drive array is not available, storing electronic records and their associated index data on server-class hard drives which are designed for greater tolerances and durability than standard desktop PC hard drives can be used, assuming that daily backup, cloud, and off-site storage of the data is available.
5. It is recommended that public agencies implement and use an electronic content management system (ECMS) to manage electronic records during their authorized retention periods. An ECMS provides the ability to capture, store, retrieve, display, and transmit records electronically. An ECMS uses a database to manage descriptive information to aid in the retrieval of records contained in the ECMS repository.
6. Storing electronic records outside of an ECMS is not recommended due to the greater chance of accidental deletion of these records and lack of an audit trail to ensure a record's authenticity. State agencies shall consult with the Department of Administrative Services (DAS)/Bureau of Information Technology Solutions (BITS) regarding appropriate ECMS technology.
7. Individual user storage accounts (e.g., desktop folders, OneDrive, Google Drive) should not be used to indefinitely store or maintain public agency records.

X. File Formats

A. Public agencies create and receive electronic records in a variety of file formats. Over time, file formats may become obsolete or unusable, rendering these electronic records inaccessible. Monitoring and regulating the usage of file formats can help minimize the risk of records loss. Standardizing formats reduces costs and provides a platform to better manage records over time.

B. For the creation and effective management of electronic records, agencies should manage the file formats used and be prepared to migrate to more stable and widely used formats as needed. These practices are especially important for long-term and indefinite retention records as well as records that have a permanent retention managed by the agency and for those records that will be transferred to the State Archives.

C. For preservation purposes, the Connecticut State Library (based on the Library of Congress Recommended Formats Statement) strongly suggests records be saved as:

Images – tiff, jpeg, jp2, png, gif, svg

Documents – txt, rtf, doc, docx, odt, PDF, or PDF/A

Spreadsheets – csv, txt, PDF, PDF/A, ods

Databases – sql (with CREATE and INSERT statements), csv, xsl, xslx

Presentations – ppt, PDF, PDF/A, odp

Audio – Broadcast wave (bwf, wav), mp3

Video – avi (lossless), mov, mp4

Email – eml, html, mbox, msg, pst

Compression folders – zip

XI. File Naming

Creating unique, consistent, logical, and predictable file names distinguishes similar records from one another in the file hierarchy and facilitates the storage and retrieval of records. Well-named files allow users to browse file names more effectively and efficiently. In general, file names should be fewer than 20 characters, be short but descriptive, avoid special characters or spaces, include dates in the format YYYY-MM-DD, and include a version number or designate which is the draft or final version.

XII. Metadata Requirements

A. *General.* Whether embedded into image files or captured in an information system, metadata provides information explaining what each record contains, when and why it was created, what media it was recorded on, original dimensions, and whether any restrictions govern its access and use.

1. Depending on the public agency's existing record-keeping practices and level of intellectual control, the public agency may use information from the record series, file, or project as the source for administrative and descriptive metadata fields. If the components of a record have not been individually indexed with unique descriptions, the public agency may apply the series or file level descriptions to all records within that grouping. If the components of the record do not have individual titles, the public agency must apply unique identifier(s) instead.
2. Appropriate and accurate metadata (index) information is required to properly identify and later retrieve electronic records.
3. Indexing typically consists of a structured format and controlled vocabulary that allows more precise description of a record's content and often includes information such as record type, creation date, last modified date, last modified by, record creator, and disposition date, among other information.
4. The public agency shall be responsible for defining the specific metadata requirements needed to access the records efficiently.

5. Indexing shall comply with the specific requirements of the public agency and include a unique identifier for each electronic record. Unique filenames or other identifiers are preferably sequential and can be numeric, alphanumeric, or alphabetic as required by the public agency. They should be unique across all record series and storage media.
6. The index of electronic records should consist of a limited but sufficient number of field names to ensure adequate access to the records. Whenever possible, the field data should consist of objective indexing terms (such as personal names, file numbers, retention schedule numbers, and dates) from a controlled vocabulary, rather than subjective data.
7. Optical Character Recognition (OCR) can be performed to convert records into searchable text. Due to error rates, OCR should not be used as the sole tool for the retrieval of electronic records, and it is not a substitute for indexing and production metadata.
8. **Permanent records only:** If the public agency provides other metadata elements in addition to the metadata requirements in this section, the Connecticut State Library will accept that metadata as part of the transfer process.
9. ‘‘Mandatory if applicable’’ instructions in the tables in this section mean that public agencies must provide the metadata if the public agency captures the metadata as part of its business processes. Public agencies do not have to create ‘‘mandatory if applicable’’ metadata as an extra step to transfer records to the State Archives.
10. “Strongly Encouraged” instructions in the tables in this section mean that public agencies are strongly encouraged but not required to capture or create this metadata.
11. “Suggested” instructions in the tables in this section mean that it is only a recommendation that the public agency consider capturing or creating this metadata but it is not required.

B. *Overall requirements.*

1. For all electronic records public agencies must:
    a. Capture the metadata specified by paragraphs C, D, and E of this section at the record series, file unit or project level.
    b. When public agencies determine that records are no longer in active use and no longer subject to changes that would alter a checksum or audit trail, public agencies must generate checksums or capture an audit trail and record them as technical metadata in an information system for each electronic record, and use them to monitor records for corruption or alteration.
    c. Create file names and unique identifier(s) for each file (although public agencies must capture other metadata at the file or item level, some might be common to multiple files or items, but not these two elements).
    d. Transfer metadata specified by paragraphs C and D of this section to the State Archives in CSV or other appropriate format as agreed upon between the public agency and the Connecticut State Library.
2. For digitized records public agencies must also:
    a. Embed the metadata specified by paragraph C of this section in each image file, capture and maintain it in an information system, associate it with the records it describes, and keep it consistent and accurate in both places.
    b. Ensure that scanning equipment or camera embeds the system-generated technical metadata specified by paragraph E of this section in each image file and that image processing does not alter or delete it.

C. *Administrative Metadata*

1. Capture in an information system the following administrative metadata:

TABLE 1 TO PARAGRAPH C.1

Metadata label	Description	Requirement level
File Name	The complete name of the computer file, including its extension.	Mandatory.
Unique Identifier(s)	The unique identifier(s) is assigned by a public agency or a information system.	Mandatory.
Records Retention Schedule Record Series Number	The records retention schedule series number assigned to the records.	Strongly Encouraged.
Relation Has Part	A related record that is either physically or logically required in order to form a complete record. Mixed-media files that contain records on multiple media types should use this element to identify all components.	Strongly Encouraged if a record includes multiple parts, such as the component parts of a case file or mixed-media file.
Relation Is Part Of	A related record or file in which the described record is physically or logically included. Records that are components of mixed media files should use this element to indicate their status.	Strongly Encouraged if file is a component of a multi-part record.

2. Capture in an information system the following access and use restrictions metadata inherited from the original source records:

TABLE 2 TO PARAGRAPH C.2

Metadata label	Required fields	Description	Requirement level
Access Restrictions	Access Restriction Status. Specific Access Restriction.	Indicate whether there are access restrictions on the record (i.e. not public). Specific access restrictions on the record, based on Freedom of Information Act (FOIA) exemptions, donor restrictions, court orders, and other federal and state statutory or regulatory provisions.	Mandatory if applicable. Mandatory if access restriction exists.
Use Restrictions	Use Restriction Status. Specific Use Restriction.	Indicate whether there are use restrictions on the record. The type of use restrictions on the record, based on copyright, trademark, service mark, donor, or statutory provisions, including Freedom of Information Act (FOIA) exemptions.	Mandatory if applicable. Mandatory if use restriction exists.
Rights Holder		A person or organization owning or managing intellectual property rights relating to the record.	Mandatory if there is a rights holder.

D. *Descriptive Metadata*
Capture in an information system the following descriptive metadata from source records at the lowest level needed to support access and preservation and to maintain contextual information.

TABLE 3 TO PARAGRAPH D

Metadata label	Description	Requirement level
Title	A name given to the original record. If a name does not exist, the mandatory metadata element File Name and/or Unique Identifier(s) serves as the title for the record.	Mandatory.
Description	A narrative description of the content of the record, including abstracts for document.	Suggested.
Creator	The agent (person, agency, other organization, etc.) primarily responsible for creating the original record.	Mandatory.
Creation Date	The date or date range of the original record.	Mandatory.
Last Modified By	The user to last modify the record.	Suggested
Last Modified	The date or date range the record was modified.	Suggested
Source Type	The medium of the original source record scanned to create a digital image.	Mandatory.
Source Dimensions	The dimensions of the original source record (including unit of measure).	Suggested.

E. *Technical Metadata*
Capture in an information system the following technical metadata describing the electronic records:

TABLE 4 TO PARAGRAPH E

Metadata label	Definition	Requirement level
File Size	The size in bytes of the image file.	Mandatory.
Format Name and Version	The format name or description of the file format.	Mandatory.
Image Width	The width of the digital image, i.e., horizontal or X dimension, in pixels.	Mandatory.
Image Height	The height of the digital image, i.e., vertical or Y dimension, in pixels.	Mandatory.
Date and Time Created	The Date or Date Time the digital image was created.	Mandatory.
Scanner Make and Model	The manufacturer and model of the scanner used to create the image.	Mandatory if using a scanner.
Scanning Software Name and Version	The name and version of the software the scanner uses to create the image.	Mandatory if using scanning software.
Digital Camera Make and Model	The manufacturer and model of the digital camera used to create the image.	Mandatory if using a digital camera.

TABLE 5 TO PARAGRAPH E

Fixity metadata label	Description	Requirement level
Message Digest Algorithm	The specific algorithm used to construct the message digest for the digital object or bitstream.	Mandatory if using checksum.
Message Digest (checksum)	The output of Message Digest Algorithm.	Mandatory if using checksum.
Audit Trail	The output of the audit trail in the information system.	Mandatory if using audit trail.

F. *Transfer metadata for permanent records to the Connecticut State Library.*

1. When a public agency transfers legal and physical custody of electronic records to the Connecticut State Library, it must also transfer the associated metadata specified by paragraphs C, D, and E of this section.
2. In addition, the public agency must follow *State Archives Policy 01: Transfer of Historical Records to the State Archives or Other Approved Archival Repository*; *Procedures for the Transfer of Historical Public Records to the State Archives*; and complete a *Memorandum of Transfer* form.

XIII. Back Up Copies

A. Public agencies either individually or through central IT (where applicable) shall perform periodic backups of all electronic records, associated indexes, and production metadata to ensure the continued accessibility of records in the event of a disaster. It is recommended that public agencies perform regular testing of the backup media to ensure electronic records have been backed up and are readable.

B. A backup copy shall be stored in a location that is geographically remote from the location where the use copies of the records are stored. An appropriate backup location is one where it is highly unlikely that the backup location will simultaneously suffer the same disaster as the public agency offices. For example, if the public agency is in or near a flood plain, the backup location should be in an area that is away from that flood plain.

C. Backup copies should be destroyed after/along with approved destruction of electronic records as part of the disposition process.

XIV. Disposition of Electronic Records

A. The disposition of electronic records shall be in accordance with *Public Records Policy 04: Electronic Records Management* and *Public Records Policy 05: Disposition of Public Records*.

B. Public agencies shall have documented policies and procedures that specifically address the defensible destruction of electronic records.

C. These practices shall be consistent with the public agency's procedures for the lawful disposition of public records in other formats and should follow a regular and systematic disposition schedule.

D. *Confidential and Sensitive Information:* Electronic records shall be destroyed in a manner that ensures that any information that is confidential or sensitive, including proprietary or security information, cannot practicably be read or reconstructed. Recorded media previously used for electronic records containing information that is confidential or sensitive, including proprietary or security information, shall not be reused.

E. *Transfer:* Archival electronic records may be transferred to the State Archives or approved archival repository in accordance with *State Archives Policy 01: Transfer of Historical Records to the State Archives or Other Approved Archival Repository* and *Procedures for the Transfer of Historical Public Records to the State Archives*.

XV. Electronic Communications

A. Email, text, and chat messages (electronic communications) have changed the way public agencies communicate with their users, but management of electronic communications records are often neglected. If public business is being conducted, it is an official record. Not all communications rise to the level of official record, but generally if users are conducting official government business, any related communication is a public record. Public business on private accounts is still public. Users should avoid combining business and personal communications.

1. Public agencies need to understand how third-party tools operate prior to using them for business functions. Using email and text and chat messages for government communication complicates the process of capturing, managing, and preserving records, since these platforms are typically operated by

parties outside of government. Public agencies should also clearly understand the limits and agreements of the technologies being used and plan for records management.

B. Public agencies should have policies and procedures that clearly document how each communication technology should be used, set limits on what content may be transmitted by such technologies, and outline procedures for retention, retrieval, preservation, and disposition of communication content. Both record and non-record communication should be addressed. Agencies should consider the following questions:
   1. What type of agency business (if any) is appropriate to be conducted via electronic communications?
   2. Who in the agency can conduct agency business via text and chat messaging (e.g., elected officials, executive management, line employees, etc.)?
   3. Is conducting agency business via electronic communications allowed using personally owned devices or only using agency-owned devices?
   4. With public records created and received as email and text and chat messages, how is the agency going to:
      a. capture the communications?
      b. retain the communications for the minimum retention period in accordance with current approved records retention schedules?
      c. destroy or transfer those communications once their minimum retention period has been met?
      d. enforce these policies and procedures?

C. Public agencies must ensure that everyone who is part of the agency: (a) is aware of their agency's policies and procedures; (b) understands their responsibilities; and (c) knows how to comply with the policies and procedures.

D. Content, not format is important. Just as you would not keep a letter on yellow paper longer than one on white paper just because of its color, you would not keep or destroy communications based solely on format. Whether a message is sent via email, text, or other means, the content of the message is what determines its value and retention.

E. The responsibility for ensuring that public records of agency business conducted via electronic communications are appropriately retained lies with the agency. Third party companies are governed by their own policies, compliance with their own regulatory framework and by the agreements and contracts a public agency makes with them. Agencies need to be aware and understand what their contract with their text and chat messaging service provider covers in terms of retention of messages and the agency's ability to access those records, especially if agencies are choosing to rely on their provider to meet the agency's records retention responsibilities.

XVI. Text Messages

Text messaging is an important part of communication that is increasingly used by public agencies. Like many other forms of communication, it is important to remember that text messages that relate to public business are public records. All content relating to the conduct of the public's business, are public records, pursuant to C.G.S. § 1-200(5). As such text messages must be retained for the minimum retention period as listed on the Connecticut State Library's Records Retention/Disposition Schedules. Retention periods for text messages should be based on the record series related to the content of the text.

XVII. Options for Capturing and Retaining Text Messages

Public agencies should consider the options below regarding the capture and retention of text messages:

1. Users Save Messages – Public agencies can choose to have their users be responsible for manually saving their text messages to an agency-controlled storage device such as an Enterprise Content Management (ECM) system or a server. However, it may be difficult to demonstrate that this is done consistently.

2. Automatic Capture to Public Agency-Controlled Storage – Public agencies can choose to either configure their text messaging service or use third-party software to automatically capture each text message sent and received either into a repository or as an email sent to the agency.
3. Vendor Capture and Store Services – Public agencies can choose to use a vendor service to capture and retain their public record text messages. Again, public agencies will need to be aware and understand what their contract with their vendor service provides in terms of retention, access to the records, what happens to the text message records at the end of their minimum retention periods and what happens if the contract is terminated, or the vendor goes out of business.

XVIII. Retention of Electronic Communications

Electronic communications must be retained for the minimum retention period as listed on the Connecticut State Library's Records Retention/Disposition Schedules. Retention periods for electronic communications should be based on the record series related to the function and content of the communication, not its format or method of transmission. How long electronic communications messages need to be kept depends on the public agency's business, legal and accountability needs to retain the evidence of the transaction that is documented in the communication. The questions to ask to determine the function/content of electronic communications are:

1. What is the communication about? (content); and
2. Why was it sent and for what purpose? (function)

XIX.Social Media Sites

A. All content, including, but not limited to, comments and postings on a public agency's social media accounts, relating to the conduct of the public's business, are public records, pursuant to C.G.S. § 1-200(5). As such, comments and postings must be retained for the minimum retention period as listed on the Connecticut State Library's Records Retention/Disposition Schedules. Retention periods for social media postings should be based on the record series related to the content of the post. For example, if an agency uses Twitter™ for public relations purposes, these records should be retained in accordance with Public Relations Records series on the State General and Municipal General Schedules.

B. As a general rule, do not rely on the social media tool for recordkeeping; government bodies should keep a copy of the record within their own filing system. Methods to capture social media records include:

1. copy and paste into a word document;
2. use web crawling or other software to create local versions of sites;
3. use web capture tools to capture social media;
4. use platform-specific application programming interfaces (APIs) to pull content;
5. use RSS Feeds, aggregators, or manual methods to capture content; and
6. use tools built into some social media platforms to export content.

C. The options for successful social media capture will depend on the technical configuration of the social media platform. Agency needs may also affect which social media capture method is used. Once the agency determines the capture method, they should provide training to applicable staff on how and when to use capture tools for social media. Agencies may need to work with third-party providers to implement social media capture.

XX. Websites

A. All content on a public agency's website(s) relating to the conduct of the public's business are public records, pursuant to C.G.S. § 1-200(5). As such website management and operations records must be retained for the minimum records retention periods listed on the Connecticut State Library's Records Retention/Disposition Schedules. The records retention periods for content on a public agency's website should be based on the record series related to the content on the website. For example, if an agency uses their website for public

relations purposes, these records should be retained in accordance with Public Relations Records series on the State General and Municipal General Schedules.

B. Website management and operations are an integral part of a public agency's program. Managing web records properly is essential to effective website operations, especially mitigation of the risks associated with using the web to carry out business.

C. Public agencies should incorporate into their records management policies and procedures the department, program, users, and/or teams responsible for (a) website content and (b) website management and operation records. There might be multiple users involved in managing website management and operation records.

D. Public agencies that contract out their website development should work closely with the vendor to ensure that web management and operations records can be captured and preserved according to records management policies and procedures.

E. Public agencies should develop a governance strategy to retain web content and website management and operations records. The agency's strategy should address roles and responsibilities; capture and maintenance of web content; and determining the proper retention and disposal of web records and content.

F. ***What Constitutes a Web Record?***

1. The first step in managing web records is determining whether it meets the definition of an "official" record and shall follow Connecticut statutes concerning the creation of, retention of, and continuing access to public records. Public agency users, vendors, and partners supporting web management and operations should understand that much of the content and documentation associated with public agency websites may meet the definition of a record and must be managed as such.
2. Website-related records can be broken into two main categories:
    a. web content records representing information presented on a website, and
    b. website management and operations records, which provide evidence of the management, operations, and structure of the website.
3. *Web Content Records*
    a. Web content is comprised of information on the website itself. This can include but is not limited to content pages that make up a website (e.g., public agency information, meeting agendas/minutes, reports, policy explanations), as well as records that can be created dynamically when a user interacts with the website.
    b. For all web content, the determination must be made if official records will be managed solely on the website. Managing official records solely on the website requires the implementation of separate records management controls.
    c. An alternative option is managing web content in agency recordkeeping systems using existing records management controls and considering the website information as convenience copies of those records.
4. *Website Management and Operations Records*

    There are two categories of website management and operations records that need to be actively managed to ensure the trustworthiness of an agency website – contextual and structural.

    a. Contextual records are the administrative and technical records used to develop and maintain the website. These can include records such as policies and procedures for managing the website, site design and testing documentation, and reports that track web activity (metrics). Maintenance of these records provides a context for web operations, which attests to the reliability, authenticity, and integrity of a public agency's website.

b. Structural records provide information related to the appearance or arrangement of the information. A site map for mission-critical websites indicating the arrangement of a site's content pages is helpful in providing a framework for content records and enables the integrity and usability of both current and preserved versions of an agency website.

G. ***Retention and Disposition of Web Records***

1. *Retention and Disposition of Web Content*

a. The records series under which a web content record is classified depends solely on the information content.

i. Non-Records

(1) If web content does not meet the definition of a record, take the necessary steps to dispose of/update the content when it no longer has value and to ensure the content is not kept longer than the official record.

ii. Official Record Copy

(1) If web content meets the definition of a record and is being managed solely on a public agency's website as the official record, determine whether an existing records series applies. Web content that is the official record may not be destroyed without an approved records series and agencies must follow disposition procedures as outlined in *PRP 05: Disposition of Public Records.*

2. *Retention and Disposition of Web Management and Operations Records*

a. Web management and operations records should be retained and disposed of following disposition procedures as outlined in *PRP 05: Disposition of Public Records.*

b. Any portion of website administrative information that contains official records should be retained and disposed of per the appropriate records series. This may differ significantly from one website to another based on business function and criticality.

c. Agencies may consider capturing the following where appropriate:

i. Metadata that makes it easier to retrieve, use, or manage web records/content.

ii. How information is displayed on the website, revised and removed, in addition to having an awareness of what records are created when these actions take place.

iii. Transaction logs for transaction-based website functions.

iv. Versioning of website content and records (may want to establish the difference between a minor version and a major version and what needs to be captured).

v. Rollbacks where changes have been made affecting user views and functionality.

H. ***Capturing and Maintaining Website Records***

1. Automated or manual processes are recommended to be in place for capturing web content to document compliance with state laws and regulations. Web content that contains an official record needs to be captured and remain accessible for its entire lifecycle, which can be accomplished via electronic content management systems (ECMS) or similar tools.

2. Where possible, web content should be a copy of the record and the official record copy should be maintained within the agency's recordkeeping system.

I. ***Decommissioning Websites No Longer in Use***

1. Public agencies should develop a content lifecycle strategy that includes what to do in cases where websites or their sub-sites have become stale or obsolete. When websites or sub-sites are ready to be decommissioned, consider the following before deleting the site and associated content of the site from the web server:

a. There may be official records retained on the site that need to be managed per a record series and where the content must be moved and appropriately retained through alternative storage for its entire lifecycle.
b. Public agency users can work with IT to collect and archive the pages just prior to decommissioning of the website.

XXI. Google Workspace

1. Google documents, slides and sheets require different methods of handling, as they exist as data that is rendered within the browser, rather than as distinct files. Depending on the type of documents, slides and sheets you may be able to download into Microsoft Office or PDF formats. In some instances, the original format cannot be downloaded and rendered as is possible with a Word Document or PDF file.
2. For office style documents and spreadsheets, the Microsoft or Open Office formats offer the most similar functionality.

Definitions

The definitions below are from the National Archives and Records Administration (NARA) and the Society of American Archivists (SAA) *Dictionary of Archives Terminology*, except where noted. See also *OPRA Records Management Terms;* policy and standards resources, and additional information are available on the Office of the Public Records Administrator website (https://ctstatelibrary.org/publicrecords/).

Accessible is information arranged, identified, indexed, or maintained in a manner that permits the custodian of the public record to locate and retrieve the information in a readable format within a reasonable time. (Wisconsin Public Records Board (PRB), *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*)

Authentic/authenticity means that the record is unaltered from the original; that it is what it purports to be, and/or that its representation is transparent.

Approved archival repository is a repository that meets professionally accepted archival facility and infrastructure requirements including but not limited to the care, management, security, preservation, and accessibility of public records. Contact the Office of the Public Records Administrator and the State Archives prior to any archival records transfer to an archival repository.

Audit Trails link to specific records in an information system and track such information as the user, date and time of event, and type of event (data added, modified, deleted, etc.). Since audit trails may play an integral part in prosecution, disciplinary actions, or audits or other reviews, public agencies are responsible for ensuring that internal management policies are in place for retaining audit trails as long as necessary for these purposes following the minimum retention period as listed on the Connecticut State Library's Records Retention/Disposition Schedules. Audit trails help prove a record's authenticity.

Batch is a group of files that are created under the same conditions or are related intellectually or physically. During digitization, batches represent groups of records that are digitized and undergo Quality Control inspection processes together.

Chain of Custody is the complete, documented, chronological history of the possession and handling of a piece of information or a record from the time of its creation through its authorized destruction. The ability to demonstrate an unbroken chain of custody is an important test of the authenticity of records. This includes all information on a file's travels from its original creation version to its final complete version or a detailed account of the location of each document/file from the beginning of a project until the end. A sound chain of custody verifies that the agency or vendor has not altered information either in the copying process or during analysis.

Checksum is a function that takes an input string, which can be of any length, and generates an output of fixed length. The output, or hash, is used to authenticate information, such as whether a file has been corrupted or modified. The values returned by a hash function are called hash values, hash codes, digests, or simply hashes. A digital signature is a special form of checksum, whose hash value is generated by a private key and verified with a public key.

Cloud or Cloud Computing consists of three parts: 1) delivery of hosted services over the internet, or an organization's intranet, instead of on a user's local computer; 2) storing, accessing, sharing, and using data with those hosted services; and 3) the hardware, networks, and staffing required to maintain the data and services. When the Cloud resources are owned and operated by an organization itself, it is known as a "private cloud." Most commonly, cloud resources are offered as a service from a third-party provider and are known as a "public cloud."

Color management is using software, hardware, and procedures to measure and control color in an imaging system, including capture and display devices.

Content means the basic data or information carried in a record. (Wisconsin PRB, *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*)

Context is the relationship of the information to the business and technical environment in which it arises. "Context" can include, but is not limited to, such elements as: the origin of the record; date and time the record was created; identification of the record series to which the information belongs. (Wisconsin PRB, *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*).

Defensible disposition is a process by which content is systematically deleted with an audit trail that is legally admissible in court.

Digitization project is any action an agency (including an agent acting on the public agency's behalf, such as a contractor) takes to digitize records. For example, a digitization project can range from a one-time digitization effort to a multi-year digitization process, can involve digitizing a single document into an electronic records management system or digitizing boxes of records from storage facilities; or can include digitizing active records as part of an ongoing business process or digitizing inactive records for better access.

Digitized record is an electronic record created by converting paper or other media formats to a digital form that is of sufficient authenticity, reliability, usability, and integrity to serve in place of the original source record.

Disposition is "a final administrative action taken with regard to records, including destruction, transfer to another entity, or permanent preservation." (ARMA)

Electronic Content Management System (ECMS) is a software system that provides the strategies, methods and tools used to capture, manage, store, preserve, and deliver content and documents related to organizational processes. An ECMS can include features such as document management, content taxonomies, auditing capabilities, check-in/check-out and other workflow controls and security mechanisms.

Enterprise content management (ECM) is used to create, store, distribute, discover, and manage unstructured content (such as scanned documents, email, reports, medical images and office documents) and ultimately analyze usage to enable organizations to deliver relevant content to users where and when they need it. (Gartner Information Technology Glossary)

Extranet is a computer network that allows controlled access from the outside, for specific business or educational purposes. In a business-to-business context, an extranet can be viewed as an extension of an organization's intranet that is extended to users outside the organization, usually partners, vendors, and suppliers, in isolation from all other Internet users. (Wisconsin PRB, *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*)

File (noun) is a document or group of documents related by use or topic, typically housed in a folder (or a group of folders for a large file).

File (verb) is the action of placing a record or document in a folder.

Folder is a container used to group records.

Geographically remote means storing backups or duplicate copies outside of the building in which the server resides.

Image quality is a measurement of a digital image's overall accuracy in faithfully reproducing an original. A digital image created to a high degree of accuracy meets or exceeds objective performance attributes (such as level of detail, tonal and color fidelity, and correct exposure), and has minimal defects (such as noise, compression artifacts, or distortion).

Information system is an organized set of procedures, tools, and techniques designed to store, retrieve, manipulate, analyze, and display information. Note: "Information system" usually connotes the use of computers. If automated, 'information system' also refers to the hardware and software. Automated information systems are generally distinguished from real-time control systems, message-switching systems, and software engineering environments.

Intellectual control is having the information necessary to identify and understand the content and context of the records. This includes knowing the disposition schedule under which the records fall, the date range when the records were created, and any access or use restrictions that apply to the records.

Integrity means that the image is an exact copy of the original and that the data has not been altered through loss, tampering, or corruption. This is verified using an audit trait or checksum.

Intranet is a private network inside a company or organization, which is for internal use only and not accessible to the public or outside the organization's network. (Wisconsin PRB, *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*)

Lifecycle means all phases of a record's existence: creation, active use, preservation and management through to disposition. "Disposition" includes permanent preservation as well as designation for destruction. (Wisconsin PRB, *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*)

Mass digitization is the large-scale scanning of source records using scanners capable of high-volume throughput. Mass digitization approaches are appropriate for paper records of uniform size and type that can be digitized without being damaged by the equipment, and in which there is no information requiring higher specifications to ensure accurate capture (such as fine detail or precise color accuracy).

Media are the physical forms on which records are stored, such as paper, photographs, compact discs (CDs), digital video discs (DVDs), analog tapes, flash drives, local hard drives, or servers.

Metadata is the characterization or description documenting the identification, management, nature, use, or location of information resources (data). Note: Metadata is commonly defined as "data about data." Metadata is frequently used to locate or manage information resources by abstracting or classifying those resources or by capturing information not inherent in the resource. Typically, metadata is organized into distinct categories and relies on conventions to establish the values for each category.

> **Administrative metadata** is necessary to manage and use information resources and that is typically external to informational content of resources. Note: Administrative metadata often captures the context necessary to understand information resources, such as creation or acquisition of the data, rights management, and disposition.
>
> **Descriptive metadata** is information that refers to the intellectual content of material and aids discovery of such materials. Note: Descriptive metadata allows users to locate, distinguish, and select materials on the basis of the material's subjects or 'aboutness.' It is distinguished from information about the form of the material, or its administration.

Embedded metadata are textual components that exist alongside the content (usually binary data) within the file. Embedded metadata may be used to make self-describing digital files that contain specified administrative, rights, and technical metadata and can be appropriately managed outside of a recordkeeping system.

Preservation metadata is technical information that can help support the longer-term sustainability of digitized content. information about an object used to protect the object from harm, injury, deterioration, or destruction.

Structural metadata is information about the relationship between the parts that make up a compound object.

Technical metadata are elements of information that describe processes used to create electronic files, and parameters that aid a system in rendering the files properly. Technical metadata may include elements such as a file's byte size, file format and version, color encoding, and the type of equipment used to make the file (camera name, scanner manufacturer, etc.).

Mixed-media files are records in different forms of media. A file, when used in the phrase ''mixed-media file,'' is a group of records—regardless of location and type of media—that belong together or relate to a topic, such as a case file. For example, a mixed-media case file could be a box with paper notes, audio recordings of interviews, and a CD of photographs, along with physical evidence stored separately in an evidence locker. Records in a file may be in more than one media type due to changes in how agencies create, maintain, and use records, shifts in technology, and the topic or activity involved.

Official Record Copy is the specific copy of a public record, as provided in C.G.S. § 1-200(5), designated by the public agency as the legally recognized copy that must be maintained for records retention, preservation, and authentication. For example, if records are kept in both electronic and hard copy format, the agency must identify the official record copy.

Physical control is having the information necessary to physically manage the records. This includes knowing where the records are housed, whether any records that fall within the project's scope are missing or stored separately, and the records' physical form (such as media types, the records' dimensions, and the smallest level of detail used to convey information).

Project plan establishes the vision and goals for the project, summarizes key points of historical or referential context, identifies stakeholders, addresses any areas of concern or risk for the long-term preservation of and access to digitized materials, and communicates in broad strokes the overall plan for the project.

Public Record as defined by to C.G.S. § 1-200(5), is "any recorded data or information relating to the conduct of the public's business prepared, owned, used, received or retained by a public agency, or to which a public agency is entitled to receive a copy by law or contract under section 1-218, whether such data or information be handwritten, typed, tape-recorded, printed, photostated, photographed or recorded by any other method."

Quality control (QC) is the process by which a public agency reviews the quality of all steps in the creation and maintenance of electronic records through inspection or testing to determine if they meet their specifications. The purpose is to detect defects (deviations from predetermined requirements) in records or processes.

Reference target is a chart of test patterns with known values used to evaluate the performance of an imaging system.

Reflective digitization is a process in which an imaging system captures reflected light off of scanned objects such as bound volumes, loose pages, cartographic materials, illustrations, posters, photographic prints, or newsprint.

Reliable means the electronic record produced accurately reflects the initial record each time the system is requested to produce that record. (Wisconsin PRB, *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*)

Reproduction scale accuracy measures the relationship between the physical size of the original object and the size in pixels per inch (ppi) of that object in the digital image.

Resolution is the level of spatial detail an imaging system can resolve in an image.

Rollback is the operation of restoring information to a previous state by canceling a specific transaction or transaction set. Rollbacks are either performed automatically by database systems or manually by users. (Wisconsin PRB, *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*)

Sharpening is used to artificially enhance details to create the illusion of greater definition.

Source record/original source record is the record from which a digitized version or digitized record is created.

Structure is the appearance or arrangement of the information in the record. "Structure" can include, but is not limited to, such elements as heading, body and form. (Wisconsin PRB, *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*)

System trustworthiness means a system that is believed to be capable of operating within defined levels of risk despite the environmental disruptions, human errors, structural failures, and purposeful attacks that are expected to occur in its environment of operation.

Transaction logs is a system generated history of actions for a specific business purpose. (Wisconsin PRB, *Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A*)

Transmissive digitization is a process in which the system transmits light through a photographic slide or negative.

User is any person who creates, modifies, deletes, or accesses electronic records. In the present context, users include, but are not limited to public agency employees, contractors, individuals on a PSA, interns, volunteers, or the public.

Versioning is creating updated versions of content records. (Wisconsin PRB, Guidance for Managing Web Records for State Agencies and Local Units of Government Appendix A)

Web Archiving is the process of collecting, preserving, and providing enduring access to web content.