# POLICY REGARDING AGENCY LONG TERM RECORDS IN ELECTRONIC FORMAT

1. **AUTHORITY**
   The Pennsylvania History Code $305 (4) states that the Pennsylvania Historical & Museum Commission (PHMC) can "Recommend such action be taken by the persons having the care and custody of public records as may be necessary to secure their safety and preservation…." .

2. **PURPOSE**. To establish policy, responsibilities, and procedures for agencies that maintain long-term textual records in electronic format. This policy does not include records that have been deemed archival by the Pennsylvania State Archives.

3. **SCOPE**
   Applies to all agencies subject to The Administrative Code of 1929, agencies under the Governor's jurisdiction and entities that store records at the State Records Center (SRC).

4. **OBJECTIVES**
   - **4.1** To ensure the efficient administration and management of electronic records and the preservation of electronic records having long-term value;
   - **4.2** To ensure the accuracy, usability, longevity and legal acceptability of electronic records maintained by agencies;
   - **4.3** To ensure the accessibility of electronic records in conformance with schedules approved by the Executive Board.

5. **DEFINITIONS**
   - **5.1** Convert – To transform a computer record from one format to another. For example, convert the records from Microsoft Word to PDF/A, or convert the records from the current version of PDF/A to a newer, future version of PDF/A.
   - **5.2** Conversion Source- the copy of the record that is stored in the old format. In a conversion *from* Microsoft Word *to* PDF/A, the Microsoft Word record is the 'conversion source.
   - **5.3** Conversion Target- the copy of the records that is stored in the new format. In a conversion *from* Microsoft Word *to* PDF/A, the PDF/A record is the 'conversion target.'
   - **5.4** Fixed Format- a format that cannot be edited.
   - **5.5** Image - the digital representation of one page.
   - **5.6** Long-Term Records-records that must be retained for more than ten years but are not permanent.
   - **5.7** Metadata - Data that describes other data. Data may be descriptive (author, title) or technical (image size, color depth, image resolution).

**5.8** Migrate - migration is the process of transferring from one operating or storage environment to another.

**5.9** Migration Source- the copy of the record that sits on the platform that is being abandoned in favor of new technology. The copy, for example, that sits on the older server.

**5.10** Migration Target- the copy of the record that sits on the newer platform after the conclusion of the migration. For example, the copy that sits on the new server.

**5.11** Page - one side of a physical sheet of paper.

**5.12** Permanent records- records that have been scheduled for permanent retention by the State Archives or by applicable statute and regulations.

**5.13** Renderable- the electronic file can be presented, using an existing and available combination of software and hardware, into a visual image. An electronic image is not considered 'renderable' if it is only theoretically so or if it requires software or hardware not in the possession of the agency that retains the file.

**5.14** Security Preservation File- a copy of an electronic record that is stored in a secure manner so that it cannot be accessed except by a limited number of authorized users and only when no other copy of the document will suffice (if, for example, a copy is suspected to have been altered and a court wants to see what the document looked like when it was captured officially).

## 6. POLICY

**6.1** Agencies are responsible for ensuring the continued accessibility of information created or maintained by their offices for the length of time prescribed by the General Records Retention and Disposition Schedule (M210.9) or their agency-specific schedule.

**6.2** Long-Term records. Long-term records must be retained either in paper format, microfilm format that has been created and maintained in conformance with applicable standards in Management Directive 210.8, or in electronic format, one copy of which (the Security Preservation File) is maintained in the manner and under the conditions set forth in this policy and its related guidelines, issued by the State Archives**.**

**6.3** Permanent records. Permanent records may be retained according to section 6.2 of this policy after consultation with the State Archives.

**6.4** Non-Permanent records. Agencies are not required to make special provisions for non-permanent records in electronic format, but all such records must be accessible for the full length of their entire retention period.

**6.5** All records, whether created or maintained on electronic systems, must be maintained in a fixed format, and must be findable, retrievable, and renderable for the entire length of the retention period designated on records retention and disposition schedules approved by

the Governor's Executive Board. It is the responsibility of agency to either (a) maintain the hardware and software required to access and display the records, or (b) establish a plan to migrate the records through each successive version of software to ensure their continued accessibility throughout their required retention period. If a record is no longer findable, retrievable, or renderable, that record will be deemed destroyed.

**6.6** Electronic record keeping systems, or procedures external to the system, must provide for the secure, confidential, irreversible destruction of all copies of electronic records (including those on backup media) at the end of the retention period specified by the General Records Retention and Disposition Schedule (M210.9) or their agency-specific schedule.

**6.7** The implementation and use of an electronic record keeping system may not limit or hinder access to records. Agencies should ensure that records maintained in such systems remain accessible and can be correlated (if applicable) with related records on paper, microfilm or other media.

**6.8** The executive director of the Pennsylvania Historical and Museum Commission (PHMC) or his designee(s), as authorized by Title 37, §305(3), shall have reasonable access to all public records maintained in conformity with this policy for the purpose of examining them and reporting on their condition. Agencies may be audited by the PHMC to ensure that long-term records maintained in electronic form are following the policy set forth.

## 7. PROCEDURES

**7.1 Administration:** While agencies must adhere to the policy, they should be aware that additional standards may be necessary to develop a robust program to preserve electronic records.

**Creation of electronic records that are long-term: : PDF/A described in ISO 19005 and subsequent revisions or future standards defined in *Guidance for Policy Regarding Agency Long-Term Records in Electronic Form* is established as the required format for long-term records maintained electronically.** *Do not confuse PDF/A with standard PDF.* PDF/A is a specific variation of PDF, designed for long term preservation. The PDF/A is an ISO-standardized version of the Portable Document Format (PDF) specialized for use in the archiving and long-term preservation of electronic documents. PDF/A differs from PDF by prohibiting features ill-suited to long-term archiving, such as font linking (as opposed to font embedding) and encryption. Agencies may use standard file formats (TIFF, JPEG, PDF, etc.) for daily functions. Agencies must maintain a Security Preservation File in PDF/A format. The agency may determine at which point in their process the Security Preservation File must be created, though generally it should occur as records are closed or complete. Systems are readily available to convert most formats to PDF/A.

**7.2 Quality Control:** Agencies must ensure the quality of the Security Preservation File maintained on electronic record keeping systems. If records are scanned from paper copies (rather than born digital), it is recommended that 100% of scanned images be quality controlled.

**7.3 Data Integrity**: To enhance legal admissibility of the Security Preservation File, trustworthiness must be established by thoroughly documenting the record keeping system's operation and the controls imposed on it. Agencies shall:

**7.4.1** Review and verify the records before the Security Preservation File is created. Ensure that quality control evaluation of images and corresponding index data is performed. Before accepting records electronically, review data and verify for authenticity, integrity and freedom from computer viruses.

**7.4.2** Security procedures must prevent unauthorized addition, modification or deletion of the Security Preservation File.

**7.4.3** Provide audit trails**.** Provide for system auditing trails and system security by utilizing software capable of monitoring and recording system access and usage.

**7.4.4** Document all processes. Create the Security Preservation File through standardized, repeatable, and auditable processes that are well documented. Have written procedures for quality control, indexing, corrections, expungement, redaction, back-ups, security and migration.

**7.4.5** Maintain up-to-date and historical technical and system documentation for each information system that produces, uses, or stores the Security Preservation Files.

**7.4.6** Provide metadata with each Security Preservation File sufficient to identify the record and to prove provenance.

## 8. SECURITY

**8.1** Agencies shall implement and maintain an effective security program to protect Security Preservation Files from unauthorized access or alteration.

**8.1.1**. Network and Systems Access. Controls need to be in place to protect the Security Preservation Files against attacks or software vulnerabilities, both internally and externally.

**8.1.2** Physical Access. Controls need to be in place to prevent unauthorized physical access to resources where Security Preservation Files are stored.

**8.1.3** Confidential, Sensitive, and/or Personally Identifiable Information (PII). Agencies are responsible for identifying and classifying the Security Preservation Files, as defined by applicable laws, and must take appropriate measures to protect and maintain the confidentiality of such records.

**8.1.4** Migrate or Expunge. When the Security Preservation File is migrated or expunged, precautions must be taken so the information cannot be reconstituted.

## 9. PROTECTION

**9.1** Agencies shall implement and maintain an effective program to protect records with long-term value from loss through natural- or human-caused disaster. The program must be documented in writing.

## 10. PRESERVATION

**10.1** Agencies shall implement and maintain an effective program to:

**10.1.1** Migrate copies of the Security Preservation Files and their associated metadata from one environment to another each and every time software or hardware changes make such actions necessary to avoid technological obsolescence; and

**10.1.2** Convert copies of the Security Preservation Files and associated metadata from one format to another, as required for preservation and access or to avoid technological obsolescence.

**10.2** During each migration or conversion, records shall be accurately converted from the source file to the target file specifications.

**10.2.1** The migrated/converted records shall preserve, in the target file, any links to other files found in the source file. It is best not to use hyperlinks in a document if possible. If there are links to external files and the external files form part of the official record, the external files must be captured as PDF/A and migrated. The migrated/converted records shall preserve, in the target file, any functionality (such as searchability or sortability) present in the source file. No data or records shall be lost during the migration/conversion, either through data corruption or failure to convert/migrate (whether caused by system or human error).

**10.2.2** Each migration or conversion shall be documented in sufficient detail as to make the migration/conversion process understandable to future technologists who may be required to reconstruct the history of the record migrations/conversions.

**10.2.3** A sufficient percentage of records and associated metadata shall be sampled and the sample shall be documented as part of the record of the migration/conversion. The size of the sample shall be determined and documented by the agency which owns the records being migrated or converted.

## 11. STORAGE

**11.1** To ensure that the Security Preservation Files are accessible and useable throughout their life span, agencies must select appropriate media and systems to fulfill the retention requirements of the records. The following provisions are required for the storage and maintenance of the security preservation files.

**11.1.1** Records scheduled for long-term retention must be retained in an online environment. Store the Security Preservation Files and associated metadata on a server (or a mainframe acting as a server). The Security Preservation File must not be retained solely on any kind of removable media; copies may be retained on removable media. The Security Preservation Files may be backed up on either paper or microfilm if that proves to be a more cost effective approach to the long term maintenance of the records.

**11.1.2** Maintain copies of vital (systems critical) records to provide for disaster recovery and business continuity. The copies shall be designated:

**Security Preservation File copy:** Retain in a live (online) environment and do not use except to create the Backup and Access copies.

**Backup copy of the Security Preservation File:** A copy to be used only in the event of a disaster that disrupts or destroys the Access copy. Retain at a site that is geographically separated from the original Security Preservation Files.

**Access copy:** Use to access the records. Retain locally. May be in the native format.

**11.2** To minimize risk of loss, agencies must select appropriate locations and provide appropriate environmental conditions, as detailed in *Guidance for Policy Regarding Agency Long-Term Records in Electronic Form*, for storage of media and systems used to store electronic records. Agencies should follow OA/OIT policies.