Guidance for

The Form of Permanent Recordation for Municipalities

The Local Government Records Committee has issued the "Policy Regarding the Form of Permanent Recordation for Municipal Offices." This guidance supplements that policy and may be edited by the State Archives with concurrence of the Local Government Records Committee. The document gives current best practices to assist an office to maintain records electronically and is numbered along with the policy. Municipalities must notify the Pennsylvania Historical and Museum Commission of their intent to maintain records permanently in PDF/A format for each records series. Refer to Appendix E of the Municipal Records Manual.

7. Security Preservation File

7.1 Administration: While municipalities covered by this policy must adhere to the policy, when developing a robust program.

ANSI/AIIM/ISO Standards. The American National standards Institute (ANSI), Association for Information and Image Management (AIIM), International Organization for Standardization (ISO), and the PDF Association are all standards organizations that provide tools for electronic records management practices. Some of the more important ones when creating permanent electronic records include:

ISO 19005 PDF/A standard. PDF/A-1a (Level A Conformance) denotes full compliance with the currently approved PDF/A Standard ISO 19005-1: Part 1. In addition to exact visual reproduction it also includes mapping text to Unicode and structuring of the document content. PDF/A-1b (Level B Conformance) is a "minimal compliance" level for PDF/A. PDF/A-1b requirements are meant to ensure that the rendered visual appearance of the file is reproducible over the long-term. It requires exact visual reproduction only. PDF/1-1a and PDF/A-1b differ primarily with respect to text extraction. The difference between PDF/A-1a and -1b has no impact for scanned documents, provided the files have not been enhanced by means of OCR for searching (PDF/A Searchable Files). PDF/A-1a requires structure in a document and is best to use for electronic documents. PDF/A-1b does not require structure in a document and is best to use for scanned documents or documents where the structure is unknown. Both PDF/A-1a and PDF/A-1b meet the minimum requirements required by this policy for permanent electronic records retention.

ISO 15489 was designed to meet the needs for recordkeeping in a business or government environment. The standard has two parts, one gives a high level framework for recordkeeping and explains the benefits of good records management, the legal considerations and the importance of making someone responsible for recordkeeping. This part also looks at what's needed for good records management, designing recordkeeping systems, records management processes, auditing and training. Another part puts the advice given in Part 1 into practice. It provides specific detail on developing records management policy and responsibility statements and suggests a process for developing recordkeeping systems. It also provides advice about developing records processes and controls such as thesauri, disposal authorities, and security and access arrangements. It discusses how you might use these tools to manage (including capturing, registering, classifying and storing) your records. It also gives specific advice about setting up monitoring auditing and training programs.

ISO 13008 specifies the planning issues, requirements and procedures for the conversion and/or migration of digital records (which includes digital objects plus metadata) in order to preserve the authenticity, reliability, integrity and usability of such records as evidence of business transactions, whether the digital records are active or residing in a repository. It provides guidance for the conversion of records from one format to another and the migration of records from one format to another. The Standard also identifies the steps, components and particular methodologies for each of these processes, covering such topics such as workflow, testing, version control and validation.

In addition, municipalities should be aware of DoD 5015.2: The Department of Defense Records Management standard. Many electronic records keeping systems have been "certified" that they follow the requirements of the standard. The standard provides guidance on these topics:

- -Access controls
- -Automatic linking
- -Data discovery
- -File plan components
- -Increased data security
- -Interoperability between solutions
- -Metadata
- -Privacy Act/Freedom of Information Act (FOIA)
- -Records retention scheduling
- -Record search and retrieval
- -Records destruction
- -Security classification
- -Vital records

7.2 Creation of electronic records that are permanent: PDF/A as described in ISO 19005 and subsequent revisions or future standards defined in this guidance is established as the required format for permanent records maintained electronically.

Do not confuse PDF/A with standard PDF. PDF/A is a specific variation of PDF, designed for long term preservation. 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. PDF/A documents can be created from most word processor programs. Municipalities should do an on-line search to find the latest information on how to create PDF/A from their specific software. To keep the sizes of files down and to make sure that files are text-searchable, PDF/A documents should ideally be created using a word processing program such as WordPerfect, Microsoft Word, or OpenOffice Writer, rather than being printed out in hard copy and scanned. Some scanning software is capable of directly creating PDF/A format documents. If your software does not have the capability to save in PDF/A format, the document may be "printed" to PDF/A using a PDF Creator Tool or it may need to be scanned as a regular PFDF and then converted/saved to PDF/A.

Special Considerations with Using PDF/A. PDF/A is the standard for the creation of archival documents, similar to paper documents, to ensure that they will be readable in the future, regardless of the fonts or software that are on a computer when a user subsequently views the documents. In essence, PDF/A is

similar to a frozen image of the document. This is desirable to ensure that permanent documents retain the text in its original formatting, as submitted.

With this in mind, it may be important for you to retain the original source document created by a word processor of file system, in the event that future edits are required. The PDF/A document is usually used as a final, unchangeable, static copy. Municipalities should also keep in mind that some dynamic electronic records such as databases may not be appropriate for conversion to PDF/A.

Be aware that converting any file (even a PDF) into PDF/A poses certain risks, and the converted files should be inspected for conversion errors. One way PDF/A ensures long term sustainability is by prohibiting certain features from the document. Documents cannot contain embedded movies, for example, and (more applicable to government documents) cannot be encrypted with a password, nor can they contain non-embedded fonts. So, for example, if a PDF document with non-embedded fonts is converted to PDF/A, the PDF/A will substitute a different font, which may result in an altered format, the deletion of special symbols, and similar issues (unless you are converting a file to PDF/A-1b image file). Many migration tools issue a warning whenever they encounter such a situation during conversion—they may ask the operator, for example, whether it is okay to use a substitute font—but others simply perform the conversion without such alerts. It is unlikely that this will be a problem during the migration of simple textual documents, but you should test carefully whenever you convert documents into PDF/A.

Other PDF/A Compliant Products. Many commercial products use plug-ins that are capable of creating PDF/A documents. The State Archives or the Municipal Records Committee does not endorse one over another.

Attachments and Hyperlinks in PDF/A Documents. It is best to refrain from using attachments and hyperlinks in the documents. Avoid attaching non-PDF/A documents since the former may become unreadable.

For cases where hyperlinks are necessary. When a document with active links is displayed, clicking a link will normally take the user to the referenced document or site. Two notable exceptions are described below.

Masked Hyperlinks. Hyperlinks are commonly “masked,” meaning that the fill address of the referenced file is not written out; for example, clicking the word Brief may open a brief stored in the system or elsewhere. An “unmasked” hyperlink has the full address visible to the user, such as http://www.jud.ct.gov/external/super/E-Services/efile/PDFA-Sample-Brief-Page.htm”.

Masked hyperlinks may or may not work in a PDF/A document, depending on how the PDF/A document was created. Currently, masked hyperlinks are preserved in PDF/A documents produced by the “Save As” method in Microsoft Word 2007 and 2010; the “PDFMAKER” method in Microsoft Word 2007; and OpenOffice2.3 (PDF Export”) With other production methods, such as WordPerfect, the PDF/A document includes underlined words that appear to be links, but clicking them has no effect. To avoid this potential problem, use unmasked links in permanent documents. An unmasked link would also be helpful to anyone viewing a printed copy of the document who needs to see a referenced item.

Other File Format Types. PDF/A must be used for the Security Preservation File. Use files of copies may be stored in other formats, but the municipality may save time and money if even these copies are stored in sustainable formats. Sustainable formats often include the following features:

-Published Documentation and Open Disclosure. Specifications for the format are published and accessible to the public. This means that anyone who wants to create tools to work with the format can do so with no restrictions of copyright. Formats that share these characteristics are commonly called "open-source" or "non-proprietary." Because anyone can create tools to access such formats, they have a low chance of becoming inaccessible in the future, even if the formats themselves become obsolete.

-Widespread Adoption and Use. The more widely a format is used, the more likely it is to have multiple tools used to access and manipulate it. This reduces the change of a format becoming inaccessible due to one software publisher going out of business. Widespread adoption also serves as an indicator of general format stability and as a safeguard against loss of accessibility. A wider user base means more stakeholders who have a vested interest in keeping a format going.

-Self-Describing Formats. These formats contain metadata within their structure that interprets the content, context, and structure of the file. This means that descriptive information (e.g., the file name, date of creation, identification of data within a file) can be kept within the file itself, and external documentation is not required. When discussing long-term preservation this is particularly important, since records often become disassociated from their original software environment and accompanying files. The more self-contained a format is, the better the chances of the data contained within being accessible down the road.

Some of the more sustainable formats are:

For Text files:

PDF/A (Portable Document Format/Archives). A variant of PDF that is specifically aimed at long-term preservation, its specifications are published in the standard ISO 19005-1:2005. It sacrifices certain functions, such as the ability to have external hyperlinks or embed audio or video, for the sake of greater reliability. The most notable different between PDF and PDF/A is the latter's ability to embed all necessary fonts within the file itself. This makes the file totally self-extracting, without any need to access external font information to properly present the formatting of the document. PDF/A also embeds descriptive metadata within the file itself, making it self-describing. These two factors make PDF/A the preferred format for long-term preservation of textual electronic records, both born-digital and digitized. Files can be converted to PDF/A by a number of different software tools and plug-ins to existing word-processor software.

For Still Images:

TIFF (Tagged Image File Format). TIFF was initially created in the 1980s in an effort to standardize file formats created by commercial scanners. The format has gone through a number of revisions since then, becoming an international standard for electronic images. The format is currently owned by Adobe Corporation, but the specifications are open and freely available. Unlike many image file formats, TIFF is uncompressed. This means that the files are larger than a compressed format (such as JPEG) but there is no loss of data. This ensures that the file can be reproduced over time at its full fidelity. TIFF files can

contain “tags” that store descriptive metadata about the file. TIFF files may have a file extension of .tif (Windows) or .tiff. TIFFs may be converted to PDF/A.

For Spreadsheets:

CSV (Comma Separated Values). A simple format which can be used to represent spreadsheet data. CSV files can be accessed with any spreadsheet software or text editor, but at the cost of potential loss of advanced functionality enjoyed by more proprietary spreadsheet formats. There is therefore a tradeoff with using CSV: universal interoperability is excellent for long-term preservation, but the loss of advanced formulae may compromise the core data of the record. Basic spreadsheets containing tabular data without advanced functions may be better served by CSV than others.

7.3 Quality Control: Municipalities must assure the quality of the Security Preservation File and must protect 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. Each must be indexed with accurate metadata to ensure future retrieval.

Newer scanners (generally those manufactured after 2004) often use software capable of directly creating a PDF/A. Users with older scanners can use a conversion tool such as Adobe Acrobat to convert scanned documents to PDF/A.

Scanner quality may be achieved by evaluating scanning devices for such problems as scanner noise, drop out threshold, optical flare, poor analog to digital conversion, scratches, dust, and out of focus sensors. Regularly maintaining and inspecting scanning equipment can ensure all is in good working order. The office should maintain logs of such activities. These logs must note any problems identified which each piece of equipment and the steps taken to eliminate each problem. When applicable, the local government should follow manufacturer’s guidelines for equipment maintenance unless it determines that its differing processes are more reliable. It is advisable to maintain routine service contracts for scanning equipment.

Office documents that contain no type fonts smaller than 6 points must be scanned at a minimum density of 200 dpi. Engineering drawings, maps and documents with type fonts smaller than 6 points must be scanned at a minimum density of 300 or 600 dpi. A grey scale level of at least 8 bits or a color level of at least 16 bits, and a minimum of 300 dpi, are recommended for the use of optical character recognition.

Digitization must capture each image with the same level of clarity as the original document page so that every legible line and character on the original document appears and is legible in the image. If the original document is illegible, the scanning process must include steps to identify the illegible original so that future viewers do not assume inadequate quality control of the scanned images.

Measures must be taken to ensure the quality of the image. Quality assurance procedures must be in place to ensure the creation of accurate and authentic images and accurate metadata. If records will be maintained permanently by scanning paper documents, quality assurance must be conducted before the destruction of any original documents. Each image of every page of all digitized documents must be inspected to ensure clarity, readability, and accurate representation of the original record. Similarly,

each metadata field must be checked against the original or imaged record. A person or persons other than those digitizing or indexing a particular record should perform the final quality control procedures outlined within this document. In most situations, quality assurance is performed in a two-step process: the scanner or digital camera operator will perform an initial quality check and supply the required metadata during the imaging process; then a different individual will perform a second review in a separate process. If a vendor is conducting the digitization, either the government agency utilizing the services of a vendor or a trusted third party must conduct the quality assurance process.

It is recommended that 100% of scanned images be quality controlled.

The quality control process must be documented and maintained throughout the digitization process. Information to document includes problem resolution procedures and reporting requirements for each step of the digitization project.

Quality control steps for digitized images must verify the following items:

- -Correct file naming convention, as per your scanning process documentation.
- -Correct file format.
- -Quality of image is the same as the original and not manipulated or changed.
- -Correct image size.
- -Correct resolution; DPI for scanners, PPI for cameras.
- -Proper viewing orientation (landscape or portrait).
- -Pages are in the appropriate order.
- -Image is not skewed or cut off.
- -Image has appropriate contrast and is neither too light nor too dark.
- -Nothing distorts the text including curvature of the page.
- -Nothing obscures the image including extraneous materials such as sticky notes, fasteners, etc.
- -No additional information was added to the image that is not part of the official record.
- -Appropriate metadata is associated with the image.

Upon inspection, any image deemed of unacceptable quality must be re-digitized and the new image re-inspected until an adequate image is achieved. PHMC reserves the right to examine and report on the condition of a municipalities' permanent electronic records.

If producing access copies and/or thumbnails of the images, they should be derived from the source image that was previously 100% quality verified.

8. Security

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

-**Unencrypted Files:** Security Preservation Files should not be encrypted in any way, as this can severely compromise the future accessibility of the records. Network, system and application security should be used to protect the records, restricting access to records as needed, while leaving the security preservation records themselves unchanged.

Should there be a need to transmit a copy of the Security Preservation File that may contain confidential, sensitive, or personally identifiable information, then encrypting the records while in transit should be enforced.

-**Network and Systems Access.** Controls, such as firewall policy and rules and intrusion detection, need to be in place to protect the Security Preservation Files against attacks or software vulnerabilities both internally and externally. Access to systems and applications should require login ID's and password protection or multi-factor or biometric authentication to ensure appropriate access and that files are not compromised. This includes mobile devices such as a mobile phone, smartphone, laptop or tablet that are authorized to connect to the municipal network.

-**Physical Access**. Controls need to be in place to ensure Security Preservation Files are not physically compromised. Ensure physical security measures are implemented for such areas as municipal information technology facilities and/or resources, off-site storage locations, and storage devices and servers that house Security Preservation Files such that only authorized individuals may access the areas to further limit and prevent physical tampering, damage, theft or unauthorized physical access.

-**Confidential, Sensitive, and/or Personally Identifiable Information (PII).** Offices are responsible for identifying and classifying the Security Preservation Files they create, collect, store, use, and/or disclose as defined by applicable laws, and must take appropriate measures to protect and maintain the confidentiality of such records.

-**Migrate or Expunge.** When the Security Preservation File is migrated or expunged, the electronic storage containing such information must be electronically wiped clean or physically destroyed in such a manner that the information cannot be reconstituted. When records are expunged, municipalities should ensure that all records that have met their retention or are obsolete are destroyed.

9. Protection

9.1 Municipalities shall implement and maintain an effective program to protect records with permanent value from loss through natural- or human-caused disaster. The policy and procedure must be documented in writing.

Backups. Backing up records and information is an essential element of data management. Regular backups protect against accidental or malicious data loss. Regardless of whether the backups are on physical media and stored off-site or are backed up to storage devices in another location, the same security measures apply for physical access, network and systems access, and protecting confidential, sensitive and PII data, just as with the Security Preservation Files.

Back-up Policy. In addition to defining the particulars for the backups, such as backup time, incremental or full backup and rotation of backup media, the policy should also include regular verification and validation processes of the backups including periodic restore drills.

10. Preservation

10.1 Municipalities shall implement and maintain an effective program to:

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

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 responsible for the migration or conversion.

When migration of Security Preservations Files is necessary, 100% of those records must be migrated.

It is best not to use hyperlinks in a document, however, if they are and it is part of the official records, the linked document must be migrated also. All external documents should replace the name of the referenced document in addition to the URL address.

11. Storage

11.1.1 Records scheduled for permanent 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 permanently valuable or 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.

For the Security Preservation File stored on electronic media, offices should consider the following:

Access: maintain the records in a usable format that keep up-to-date materials needed to access them, including indexes and other documentation.

Backups: Maintain backup copies of records and all materials required to access them in an off-site, preferably geographically different, location that does not share the same disaster threat. Cloud storage **of backup copies is acceptable so long as the storage meets the requirements of this policy. Create policies and procedures for backing up records.**

Cost: There is a significant cost associated with the conversion and ongoing maintenance of permanent records in PDF/A in a live server environment. Municipalities should consider the resources involved before determining which permanent records to keep in electronically.

11.2 To minimize risk of loss, municipalities must select appropriate locations and provide appropriate environmental conditions, as detailed in "Guidance for the Form of Permanent Recordation for Municipal Offices," for storage of media and systems used to store electronic records. Municipalities should develop their own policies.

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

Location. Servers must be located in areas with as low an environmental disaster risk as possible. Do not locate in buildings that are in known flood plains. Locate in inner rooms without windows in areas with tornados. Storage room should have fire rated walls, smoke alarm(s), and fire suppression rated for electrical fires.

Temperature. To keep equipment from overheating and being damaged, the server room should be temperature controlled, and set at about 65-75 degrees F, relative humidity between 35-50%. It is a good idea to have separate temperate controls for the server room; relying on central air conditioning that cools an entire building is not a good idea, as temperatures can fluctuate drastically throughout the building.

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

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

Security. Allow only approved people access to the storage facility. You will want to consider, among other things:

-A controlled auditable entrance (e.g. security code keypad, smart-card swipe).

-An alarm system that sounds if an unauthorized person attempts to enter the storage facility.

Ventilation. Good ventilation will help prevent dampness, mold, and pest infiltration. The air in the storage facility should be free from pollutants (e.g., strong cleaning solution fumes). Dust can also be particularly damaging to digital media.

Disaster Recovery (Continuity of Operation) Plan. As part of your records management policy, include a disaster recovery plan that provides a series of detailed actions (including who is responsible for executing each step of the disaster plan) if a disaster should occur at the storage facility. Include the response procedures for multiple types of disasters (e.g., flood, fire, smoke and explosion). The goal of the plan should be to have the facility operational and the greatest number of records recovered in the least amount of time. Train staff members and practice the disaster recovery plan. The Pennsylvania State Archives provides templates and further guidance for disaster planning on its website.

Approved March 28, 2019

Municipal Records Committee