



Vermont State Archives and Records Administration

Office of the Secretary of State

1078 US RTE 2, Middlesex • Montpelier, VT 05633-7701 • Tel: (802) 828-3700 • Fax: (802) 828-3710

FILE FORMATS GUIDELINE FOR ALL PUBLIC AGENCIES Date Effective: December 1, 2007; Last Revised: April 2020

Introduction

The purpose of this guideline is to provide guidance and advice to public agencies in the selection and use of file formats that support the interoperability and long-term preservation of public records.

Background

The accuracy and authenticity of a record is dependent on a number of factors, including: (1) the permanence of the record's content and structure; (2) reliable access for as long as the record is required to be retained; and (3) a trustworthy practice or system of recordkeeping.

Electronic records, unlike paper records, are composed largely of two parts: the format, which dictates how the data in the record is stored and displayed; and the actual content and information. Both computer hardware and software are required to process the format and the content and to present the record in a human-understandable manner.

Due to technological obsolescence, maintaining accurate and authentic electronic records is a daunting task for most, if not all, public agencies. Therefore, it is important to use file formats that better support the preservation of and access to the record's format and content. In addition, if converting paper records to electronic through imaging, using a recommended file format shortly after the time of creation or receipt helps ensure that the records are:

1. Complete and without loss of context with other records;
2. Reliable evidence of agency business activities; and
3. Well positioned for future conversions or migrations.

Furthermore, for electronic or born-digital records, using appropriate file formats at the time of creation reduces loss of context; produces more reliable records; and is less expensive than retrospective conversion processes. (Agencies using this guideline for imaging should first review the *Imaging Guideline for All Public Agencies*.)

Intended Audience

This guideline is intended for any public employee involved in the creation of electronic records which may need to be: (1) shared among other public agencies or the citizens of Vermont or (2) preserved for long term. This guideline is also intended for any public employee responsible for the interoperability or preservation of public records.

1 SELECTING A FILE FORMAT

1.1 Business and Recordkeeping Requirements

The selection of a file format should be based on business and recordkeeping requirements, which are documented in strategic plans, project charters, record schedules, and related sources. If business requirements have not been cross-referenced with recordkeeping requirements, this should be done prior to selecting a file format.

1.2 Objective

The file format selection process requires a clear objective. For example, if the objective is simply to increase access and publish State records and information electronically, then it is important to select a file format that is adaptable and designed for increasing access. On the other hand, if the objective is to preserve records electronically (regardless of the records' original format) then it is necessary to choose a file format that also supports long-term access and preservation. If both electronic access and preservation is the objective, it will be necessary to choose one format that is suitable for access and preservation or two formats: one for access and one for preservation.

1.3 Recordkeeping Requirements

It is important to review the records' appraised values and retention requirements prior to converting any files to verify that the appropriate file format is selected.¹ Also, to ensure that records remain legally viable at all times, agencies should fully articulate their plans and document their processes. Audit trails and the systematic application of established policies and procedures will be necessary to demonstrate the accuracy and authenticity of converted records. Procedures should address quality control when preparing records for conversion and verifying and validating the converted files.

1.4 Original Formats

In addition to business and recordkeeping requirements, the selection of a file format is also dependent on the format of the original records. The majority of public records are created electronically or born-digital; therefore, agencies should strive to capture their electronic records in a recommended file format rather than printing and scanning records.

¹ Appraised values appear in all record schedules issued by the Vermont State Archives and Records Administration.

2 ADDITIONAL CONSIDERATIONS

There are several other issues, such as management needs and associated costs, which need to be considered and addressed beyond selecting an appropriate file format.

2.1 Cost-Benefit Analysis

A cost-benefit analysis should be completed before choosing to implement any conversion plan. In conducting the analysis, tangible savings and benefits and intangible benefits should be equally considered.

2.2 Workflow

In most situations, the records targeted for conversion support current agency functions and activities. Therefore, the need to readily and easily convert future records after they are created and received should be introduced into the agency's workflow. Issues to consider include but are not limited to: routing procedures; interfiles and adding new records to existing files; and when records should be captured and converted.

2.2 Computer Hardware and Software

An inventory of current computer hardware and software will offer an overview of what new software or hardware may be necessary on a temporary or long-term basis to convert records or maintain the converted files. When considering alternatives, such as the use of contractors, the inventory will be an invaluable resource.

2.3 Indexing

The ability to retrieve converted records is dependent on how well the records are indexed. It is important to consider indexing by multiple access points (creator name, date of creation, agency name, subjects, record title, etc.) to increase access and retrieval by internal and external users. In addition, agencies should consider embedding metadata in each converted file; records should not be fully dependent on external indexes, such as databases, for contextual information. (See State of Vermont Information Management Standard on *Recordkeeping Metadata for All Public Agencies.*)

2.4 Staffing

It is inevitable that workflow patterns or current practices will change, therefore training needs for current staff must be considered. Agencies should also reflect on the need for additional staff or specialized expertise.

3 PREFERRED FILE FORMATS

The State of Vermont Information Management Guideline is aligned with the digital preservation guidelines set forth by the Library of Congress and the Federal Agencies Digital Guidelines Initiative or FADGI.

All Vermont public agencies should refer to the latest version of the Library of Congress' *Recommended Formats Statement*³ for preferred formats of the following: textual records, still images, audio, moving image, software and electronic learning, datasets and databases, and websites.

4 REFERENCES

CENDI Digital Preservation Task Group. (2007). *Formats for Digital Preservation: A Review of Alternatives and Issues*. Federal STI Managers Group: Oak Ridge, TN. URL: https://www.cendi.gov/publications/CENDI_PresFormats_WhitePaper_03092007.pdf (last accessed: 2020-04-16)

Federal Agencies Digital Guidelines Initiative. (2020). Federal Agencies Digital Guidelines Initiative Website. Library of Congress: Washington, D.C. URL: <http://www.digitizationguidelines.gov> (last accessed:2020-04-16)

International Organization for Standardization. (2005). *ISO 19005-1:2005: Document Management – Electronic Document File Format for Long-Term Preservation – Part 1: Use of PDF 1.4 (PDF/A-1)*. International Organization for Standardization: Geneva, Switzerland.

Library of Congress. (2007). *Sustainability of Digital Formats: Planning for Library of Congress Collections*. Library of Congress: Washington, D.C. URL: <https://www.loc.gov/preservation/digital/formats> (last accessed: 2020-04-16)

REVISION HISTORY

2020-04-15	Replaced header, removing reference to the former Department of Information and Innovation. Removed Related State Standards section and Section 4 Specifications. Revised Section 3 Preferred File Formats and added the Library of Congress' Recommended Formats Statement, which now also serves as the State of Vermont's guideline.
2010-02-04	Related Policies section was changed to "Related State Standards" and section was revised to include references to additional State standards. JPEG 2000 was added as a preferred file format for images. Links to existing References were verified and one reference that is no longer available was removed.
2008-10-17	Related standards section added. Revised to reflect changes to the Records Management Best Practice and File Formats Best Practice.
2007-12-05	"Files must pass validation" was added at the end of the first General Requirement. Microsoft Save as PDF was added as tool for Microsoft.
2007-11-28	Approved by Information Strategies: Taskforce on Archives, Records & Technology (/START) effective December 1, 2007..

³ Library of Congress. (2020). *Library of Congress Recommended Formats Statement*. Library of Congress: Washington, D.C. URL: <https://www.loc.gov/preservation/resources/rfs> (last accessed: 2020-04-16)