



Preservation Team

Format Standards: What Do I Need To Know?

Overview for Today:

1. What are Formats and Why Should You Care?
 2. Formats and Preservation Strategies for Textual Records, Structured Records, and Email.
 3. Brief Overview of Strategies and Available Resources.
-

BORN-DIGITAL

**born dig-it-al (adjective)
digital information
originally created
in electronic form**

DIGITIZED



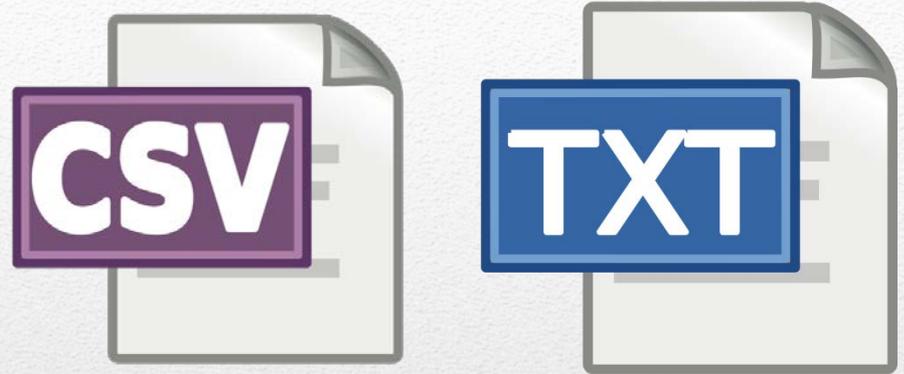
FORMAT DECISIONS AND ISSUES APPLY TO BOTH!

Understanding Formats

MEDIA



CONTENT



Adobe

Understanding Formats



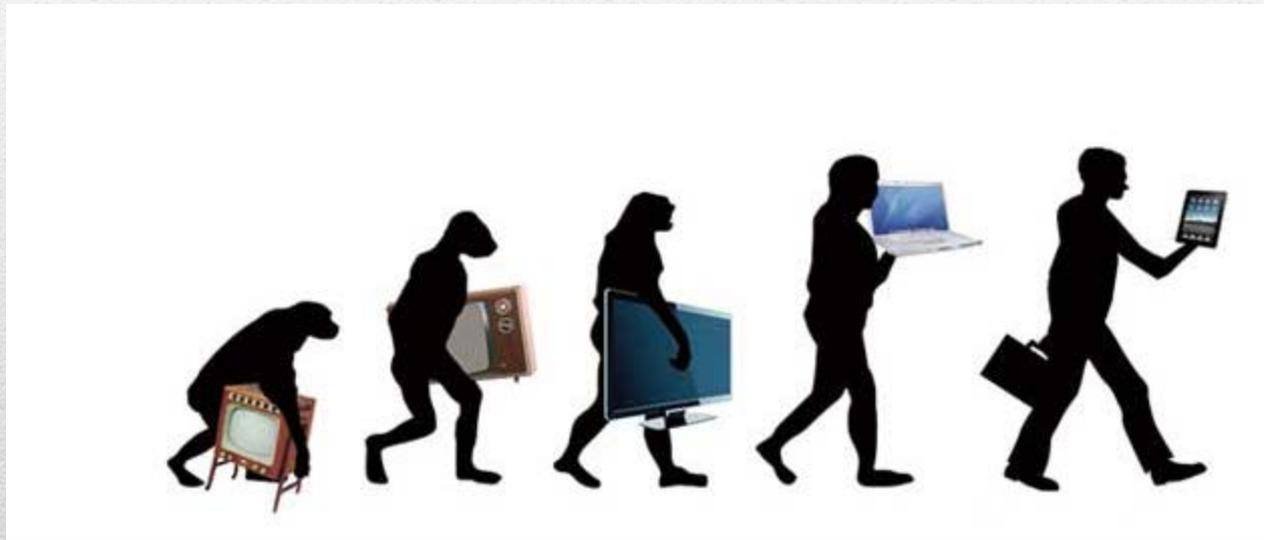
File formats are the “rules that specify how the bytes that make up [a given] file are organized, interpreted, and rendered.”

- Ciran B. Trace

Understanding Formats

The Biggest Risk with Formats?

OBSOLESCENCE!!!



Understanding Formats

Ask questions about format choices now...



...to help ensure future access.

Understanding Formats

QUESTION YOUR FORMATS



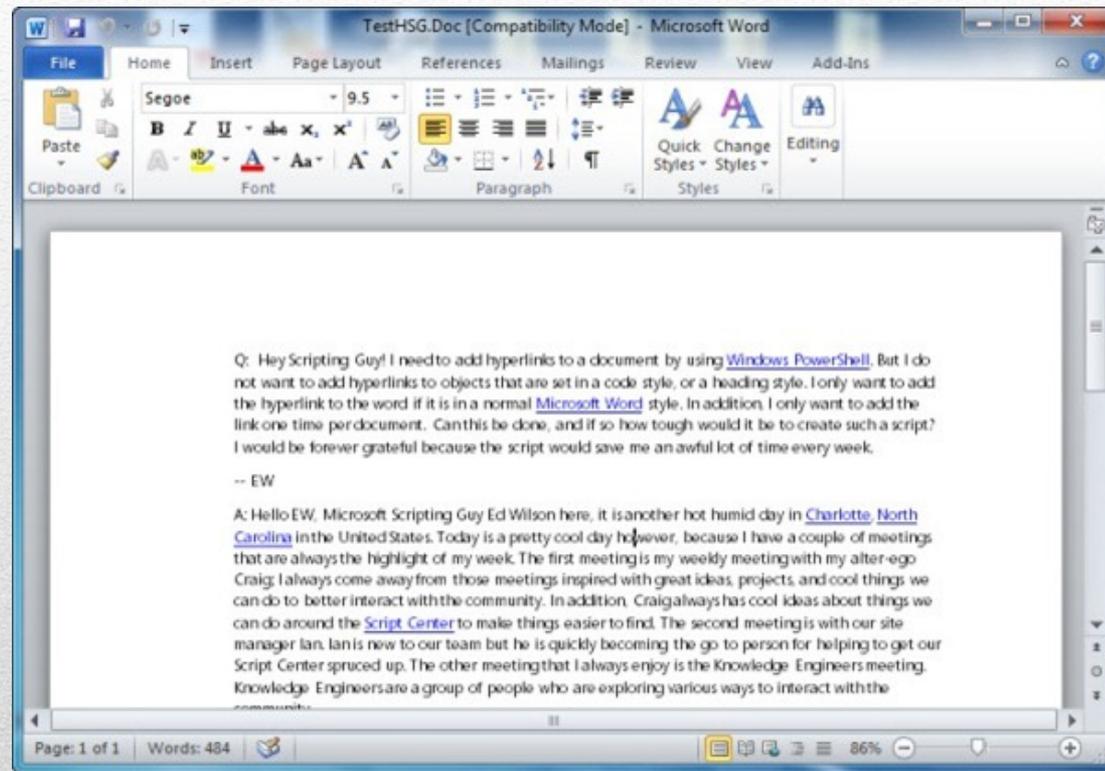
WHO OWNS?

WHO SUPPORTS?



WHO (OR WHAT) GOVERNS?

Understanding Formats

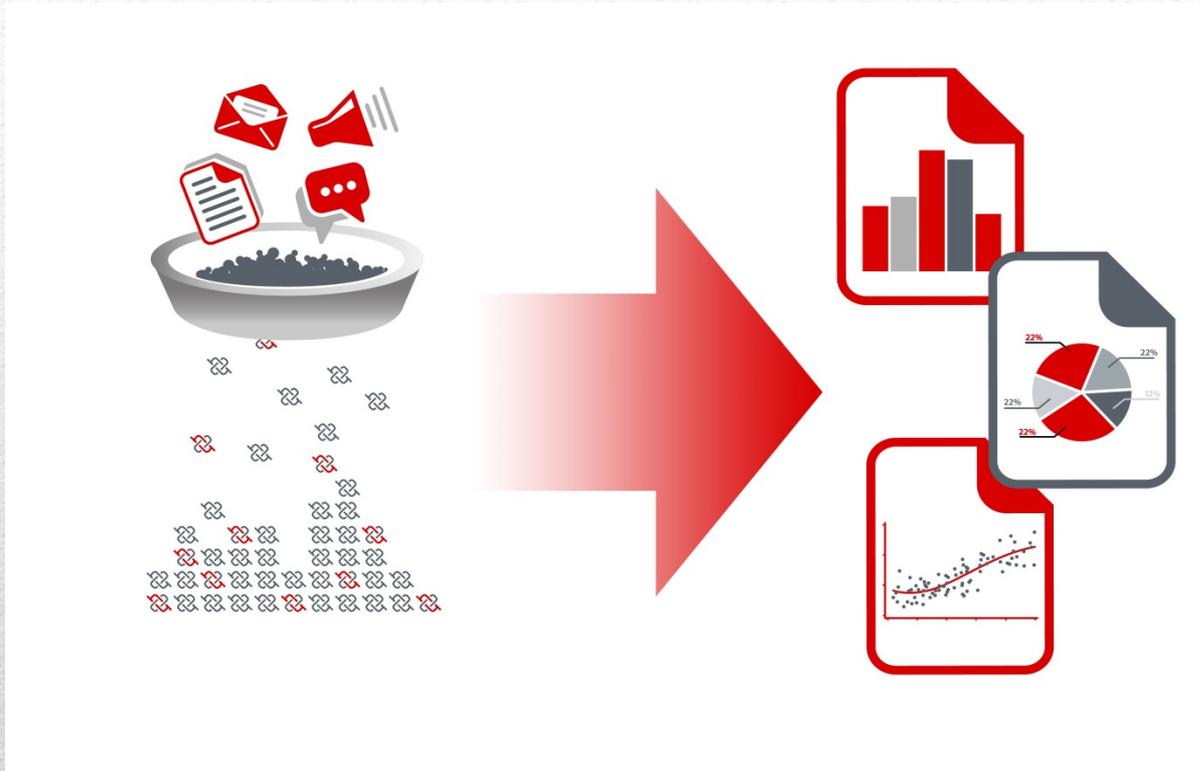


Textual Records

COMMON FORMATS



Textual Records



Structured Data Records

COMMON FORMATS



Structured Records



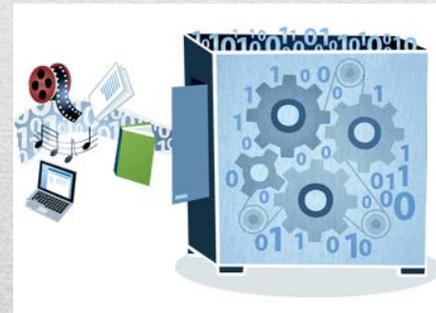
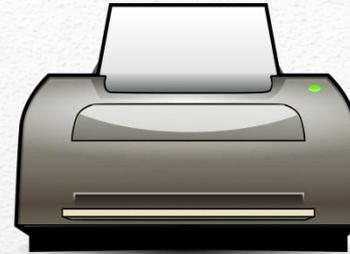
Email

COMMON FORMATS



Email

Policies & Procedures



Strategies and Resources

Professional Guidance



NATIONAL ARCHIVES

Blogs | Bookmark/Share | Contact Us

Search Archives.gov

GO

Research Our Records

Veterans Service Records

Teachers' Resources

Our Locations

Shop Online

Records Managers

Home > Records Management > Policy > Appendix A: Tables of File Formats

Records Management Resources

E-mail Management

Records Management FAQs

Memorandums to Agency Records Officers

Federal Records Centers (FRC)

Guidance and Policy for Accessioning

Electronic Records Management (ERM)

Toolkit for Managing Electronic Records

Records Management Handbook

Records Management Policy and Guidance

Records Management Initiatives (RMI)

Records Management Self Assessment (RMSA)

Records Management Training

Electronic Records Archives (ERA)

NARA 2014-04: Appendix A, Revised Format Guidance for the Transfer of Permanent Electronic Records – Tables of File Formats

Quick Links

1. Computer Aided Design	2. Digital Audio	3. Digital Moving Images
3.1 Digital Cinema	3.2 Digital Video	4. Digital Still Images
4.1 Digital Photographs	4.2 Scanned Text	4.3 Digital Posters
5. Geospatial Formats	6. Presentation Formats	7. Textual Data
8. Structured Data Formats	9. Email	10. Web Records

Symbol Key

Preferred Formats



Acceptable Formats



Acceptable for Imminent Transfer Formats



NARA 2014-04: Appendix A

- Provides format standards for a variety of digital objects encountered in an archival digital preservation environment.
- Tiered categories for formats that provides greater flexibility.
- <https://www.archives.gov/records-mgmt/policy/transfer-guidance-tables.html#presentationformats>

Strategies and Resources

Professional Guidance

LoC Recommended Formats Statement:

- "...purpose is to inform the creative and library communities on best practices for ensuring the preservation of, and long-term access to, the creative output of the nation and the world."
- Provides format recommendations and standards for both analog and digital materials.
- <https://www.loc.gov/preservation/resources/rfs/>

The screenshot shows the Library of Congress website's "Recommended Formats Statement" page. The page is titled "Recommended Formats Statement" and includes a search bar, navigation links, and a table of contents. The table of contents lists six main categories of creative content: I. Textual Works and Musical Compositions, II. Still Image Works, III. Audio Works, IV. Moving Image Works, V. Software and Electronic Gaming and Learning, and VI. Datasets/Databases. Each category has sub-links for different formats and media types.

LIBRARY OF CONGRESS ASK A LIBRARIAN DIGITAL COLLECTIONS LIBRARY CATALOGS Search GO

The Library of Congress > Preservation > Resources > Recommended Formats Statement

PRESERVATION

Search this site GO

- [Preservation Home](#)
- [About Us](#)
- [Collections Care](#)
- [Conservation](#)
- [Emergency Management](#)
- [Preservation Science](#)
- [Training Opportunities](#)
- [Resources](#)
- [FAQ](#)
- [En Español](#)
- [Contact](#)

Related Links

- > [Donate](#)
- > [Digital Preservation](#)
- > [Audio-Visual Preservation](#)
- > [Building Digital Collections: A Technical Overview](#)
- > [Preservation Metadata Maintenance Activity](#)
- > [National Film Preservation Board](#)

Recommended Formats Statement

Print Subscribe Share/Save Give Feedback

[Resources Home](#) | [Preservation Research Reports](#) | [Preservation Supply Specifications](#) | [Recommended Format Specifications](#) | [Staff Publications](#) | [Web Resources for Paper and Book Arts](#)

[Main](#) | [Table of Contents](#) | [Introduction](#) | [Textual Works and Musical Compositions](#) | [Still Image Works](#) | [Audio Works](#) | [Moving Image Works](#) | [Software and Electronic Gaming and Learning](#) | [Datasets/Databases](#)

Library of Congress Recommended Formats Statement 2015-2016

Introduction

Categories of creative content:

- I. [Textual Works and Musical Compositions](#)
 - i. [Textual Works - Print](#)
 - ii. [Textual Works - Digital](#)
 - iii. [Textual Works - Electronic Serials](#)
 - iv. [Digital Musical Compositions \(score-based representations\)](#)
- II. [Still Image Works](#)
 - i. [Photographs - Print](#)
 - ii. [Photographs - Digital](#)
 - iii. [Other Graphic Images - Print](#)
 - iv. [Other Graphic Images - Digital](#)
 - v. [Microforms](#)
- III. [Audio Works](#)
 - i. [Audio - On Tangible Medium \(digital or analog\)](#)
 - ii. [Audio - Media-independent \(digital\)](#)
- IV. [Moving Image Works](#)
 - i. [Motion Pictures - Digital and Physical Media](#)
 - ii. [Video - File-Based and Physical Media](#)
- V. [Software and Electronic Gaming and Learning](#)
- VI. [Datasets/Databases](#)
 - i. [Datasets](#)
 - ii. [Geospatial Data](#)
 - iii. [Databases](#)

[Back to Top](#)

Strategies and Resources

Professional Guidance

Sustainability of Digital Formats Planning for Library of Congress Collections

[Introduction](#) | [Sustainability Factors](#) | [Content Categories](#) | [Format Description](#)

The Digital Formats Web site provides information about digital content for Fleischhauer, and Kate Murray invite [feedback](#) on the content.

[Introduction](#)

Background information and overview: What is a format? How shall we evaluate formats? What projects in other organizations are addressing these questions? >>

[Sustainability Factors](#)

What affects the ability of the Library to preserve content in a given format? These sustainability factors apply to all formats. >>

[Content Categories](#)

The evaluation of formats must take into account quality and functionality. These factors vary according to the type of content under consideration and the categories will be expanded as time passes. >>

[Format Descriptions](#)

Documents with more information about specific formats. >>

LoC Sustainability of Digital Formats

- Useful tool for thinking about long-term format sustainability, and building corresponding institutional format policies.
- Discusses sustainability factors, content categories, and provides descriptions for individual content formats.
- <http://www.digitalpreservation.gov/formats/>

Strategies and Resources

Utah State Archives Practice

2.3

Acceptable Formats for Electronic Records

POLICY: Certain file formats are recommended for transferring to the Utah State Archives to ensure the best results for long-term preservation. File formats are typically identified by their filename extensions (e.g. .pdf, .jpg), which are allocated by the software that created the file. Preservation formats are preferred, but the Archives will accept file formats labeled as "Acceptable Formats" under this policy.

SELECTING FILE FORMATS:

"Some formats present greater risks to the continued accessibility of records than others. For example, when companies develop a file format, they can choose to keep the code closed (proprietary) or allow others to access it (open, non-proprietary). Open formats are less at risk of becoming inaccessible; with an open, published specification, anyone can develop a tool to open those files in the future if the original software becomes unavailable. Very common file formats (such as Microsoft formats), even if proprietary, are also at less risk of becoming inaccessible due to market demand."¹

The ideal preservation format is:

- stable and standardized
- free of patent rights (public domain)
- as simple as possible
- fully self-descriptive, not dependent on external sources
- suitable for containing structured and self-selected metadata
- fully documented (open specification)
- not application or platform based, but easily exchangeable
- used and implemented widely
- easy to implement and to check for errors (validation)²

RECOMMENDED FILE FORMATS FOR TRANSFER:

The following table has been adapted from recommendations of the Indiana State Archives, the State Archives of North Carolina, and the Library of Congress Sustainability of Digital Formats. Whenever possible, governmental entities should use preservation formats for each record type. Acceptable formats listed in the tables below are additional methods to use in the event that preservation types are not achievable. Any question(s) concerning formats should be directed to the records analyst assigned to the governmental entity.

E-Archives Master Draft – Version 2

May 2016

WORD PROCESSING DOCUMENTS	
PRESERVATION FORMATS	ACCEPTABLE FORMATS
PDF/A-1a ⁱⁱⁱ or PDF/A-2a (compatible with screen readers) ^{iv} ;	PDF/A-1a (.pdf) (ISO 19005-1 minimally compliant);
Open Document Text (.odt, .odt) ^v	PDF
For publications, use EPUB format (also an access format widely used by book readers)	Microsoft Open XML Document (.docx); Microsoft Word Document (.doc); Rich Text Format (.rtf)
PLAIN TEXT	
PRESERVATION FORMATS	ACCEPTABLE FORMATS
Plain Text (.txt) US-ASCII ^{vi} or UTF-8 or 16 encoding ^{vii} ;	Other delimited text files (space-delimited, colon-delimited, etc.) where the delimiting character is not present in the data.
PDF/A-1a or PDF/A-2a (.pdf) (ISO 19005-1 compliant PDF/A)	
Comma-separated file (.csv) US-ASCII or UTF-8 encoding;	
Tab-delimited file (.txt) US-ASCII or UTF-8 encoding	
TEXT WITH STRUCTURAL MARKUP	
PRESERVATION FORMATS	ACCEPTABLE FORMATS
XML (Extensible Markup Language 1.1) ^{viii} or SGML (Standard Generalized Markup Language) ^{ix} using standard or well-known DTD or schema, XSD/XSL presentation stylesheet(s) and explicitly stated character encoding.	PDF Rich text format Plain text

E-Archives Master Draft – Version 2

May 2016

Acceptable Formats at the Utah State Archives

Utah State Archives Practice

<p>2.2.1</p>  <p>Service Agreement# _____</p> <p>UTAH STATE ARCHIVES ELECTRONIC ARCHIVES TRANSFER INFORMATION MANIFEST (TIM)</p> <p>GOVERNMENTAL ENTITY: _____</p> <p>Section A – Records Identification</p> <p>1. Please provide the following information (attach additional pages if necessary):</p> <p>a. Record Series Name: _____</p> <p>b. Record Series Number: _____</p> <p>c. Inclusive Dates: _____</p> <p>d. Brief description of records being transferred:</p> <p>_____</p> <p>_____</p> <p>_____</p> <p>e. Please identify and cite any statutory restrictions on public access to these records:</p> <p>_____</p> <p>_____</p> <p>_____</p> <p>E-Archives Master Draft – Version 2 May 2016</p>	<p>Service Agreement # _____</p> <p>Section B – Security, Connectivity, and Data/Metadata</p> <p>1. Transfer of the records does not substitute for disaster recovery or standard backup. Materials transferred to the Utah State Archives must adhere to the Acceptable Formats Policy adopted by the Division (see the Acceptable Formats for Electronic Records Policy for further information).</p> <p>2. What is the estimated size of the initial records transfer (i.e. total volume and number of files – for example, 100MB, 100 files)?</p> <p>_____</p> <p>3. How will the records be transferred to the Utah State Archives? For information on available options please see the Transfer Policy section in the Utah State Archives Electronic Archives Policies and Procedures Manual.</p> <p><input type="checkbox"/> Open Records Portal Upload <input type="checkbox"/> Hard Drive <input type="checkbox"/> Other: _____</p> <p>Note: The Utah State Archives reserves the right to request a second data capture if the initial capture fails.</p> <p>4. Are there any special instructions for handling, opening, or accessing the records being transferred?</p> <p>_____</p> <p>_____</p> <p>_____</p> <p>Note: Partners may be required to submit sample electronic records to the Utah State Archives for beta testing before a complete transfer of electronic records is permitted. Samples will be submitted in a format and medium stipulated by the Utah State Archives.</p> <p>5. Please include any additional metadata that will aid with the preservation and/or access of these records as a separate attachment.</p> <p>E-Archives Master Draft – Version 2 May 2016</p>
--	---

Transfer Information Manifest



James Kichas
jkichas@utah.gov
801-531-3844
