

Preservation Action Plan: Multimedia Formats

National Archives and Records Administration (NARA)

Plan Date: 20200629

Template: 201907

Electronic Record or Digital Surrogate Types and Associated Formats

Multimedia formats are those that contain various forms of media--text, audio, and video. The most common multimedia formats are digital slide shows created using presentation software such as Microsoft PowerPoint or Apple Keynote. The key aspect of this format, however, is that the software that creates it has the ability to present information in multiple formats and therefore files may have characteristics not present in similar record types, such as containing audio and video.

Essential Characteristics of Multimedia Formats

Presentation formats are used to convey graphical information to audiences in the form of a slide show. Presentation formats are not acceptable for use as transfer containers for permanent digital still images. The record's Appearance and Structure are vital for the preservation of Multimedia; depending on the type(s) of component media files, it may not be possible to fully capture Behavior.

Appearance

Name	Definition	Function Description
Language	A language identifier specifying the natural language(s) used in the document.	
Fonts	<p>A list of fonts used in the document, and an indication of whether or not a font is embedded in a document.</p> <ul style="list-style-type: none">• Typeface (Arial, Times New Roman, etc.)• Size (10 pt, 18 pt, etc.)• Pitch	Stores the names of all fonts (embedded and non-embedded) used in a document to assist in rendering the content and identifying the documents with potential long-term preservation risks. There will always be a font in order to have text, however, there are other variables that

	<ul style="list-style-type: none"> • Spacing • Emphasis (bold, italic, strikethrough, underline, etc.) 	determine whether a distinct font is a core characteristic. For example, where appearance is fixed, such as in a publication, then font is core.
Size	Size ratio (e.g. standard, widescreen).	Physical or virtual dimensions. These characteristics are conditional. If the appearance is important or information is lost with resizing then this must be maintained. However, if only the content ("words absent their formatting") of the document is important, such as meeting minutes, then these dimensional characteristics are not core.
Layout Orientation	Orientation of the "pages," e.g., portrait versus landscape.	Needed for understanding content and context of record.
Text Orientation	Orientation of the text, e.g., horizontal, vertical, custom axes.	Needed for understanding content and context of record.
Color	<p>Identification of the use of color in text and layout elements, e.g. borders, boxes.</p> <ul style="list-style-type: none"> • Hue: color family or name • Saturation: purity or sharpness • Brightness: shade or tint • Contrast: range of optical density or tone 	Color is essential if it bears meaning and/or value, or if the appearance is fixed.
Annotations	Annotations included in the document.	
Tables	Includes content structured in a table feature, such as a dynamically generated table of contents.	A table is an arrangement of <u>data</u> in rows and columns, or possibly in a more complex structure. (Wikipedia)

Graphics	Images embedded in the document, or graphics created using a word processing graphics feature.	Used to point readers and viewers to particular information. They are also used to supplement text in an effort to aid readers in their understanding of a particular concept or make the concept more clear or interesting. (Wikipedia)
Outline	Includes content in an outline structure.	An outline, also called a hierarchical outline, is a list arranged to show hierarchical relationships and is a type of tree structure. An outline is used to present the main points or topics of a given subject. Each item in an outline may be divided into additional sub-items. (Wikipedia)
Transparency	The use of transparency features for graphics or text in the document.	
Layers	The presence of layered graphic elements in the document.	

Structure

Name	Definition	Function Description
Schema	<ul style="list-style-type: none"> • Lists • Arrays • Models • Trees • Tables • Linkage: connection between or within records. If connections exist, then they are core • Template 	
Character Encoding	Encoding schema, e.g., US-ASCII, EBCDIC, UTF-8.	Required for the proper parsing and rendering of the record content.

Word Count	Total number of words in the document.	Valuable for evaluating the completeness of the content after transformations, but can be highly variable between tools.
Character Count	Total number of characters in the document.	Valuable for evaluating the completeness of the content after transformations, but can be highly variable between tools.
Line Count	Total number of lines in the document.	Valuable for evaluating the completeness of the content after transformations, but can be highly variable between tools.
Paragraph Count	Total number of paragraphs in the document.	Valuable for evaluating the completeness of the content after transformations, but can be highly variable between tools.
Table Count	Total number of tables in the document.	Valuable for evaluating the completeness of the content after transformations, but can be highly variable between tools.
Graphics Count	Total number of graphics in the document.	Valuable for evaluating the completeness of the content after transformations, but can be highly variable between tools.

Behavior

Name	Definition	Function Description
Macros and Linkages		Essential characteristics are currently dependent on the export format.
Audio	Sounds is an audio waveform that has been created to be heard during playback.	In multimedia formats, there may be audio associated with specific slides (e.g. voiceovers).

Animation		The two most common forms of animations (in presentations specifically) are slide-to-slide transitions and animations of text/images on a slide (source).
-----------	--	---

Context

Name	Definition	Function Description
Series	Identification of the document's inclusion in a series, and enumeration if present.	
Descriptive Metadata	Descriptive metadata embedded in the record file header.	Information contained within the record (intrinsic) that refers to the intellectual content of material and aids discovery of such materials, for example, Caption/Subject/Date/Event/Transaction can all add value to the record.

Current NARA Transfer Guidance for Multimedia Formats [Bulletin 2014-04](#)

- Preferred:
 - OpenDocument Presentation Format (ODP)
 - Portable Document Format Archival (PDF/A-1)
- Acceptable:
 - Microsoft PowerPoint 1997-2007 Binary Format (PPT)
 - Microsoft PowerPoint Office Open XML Format (PPTX)
 - PDF/A-2

Current NARA Format(s) for Public Access and Reference for Multimedia Formats

Formats for Public Access are those made available online through the National Archives Catalog. Formats for Reference are defined as those made available to researchers upon direct requests for digital copies.

Formats Available for Public Access: Content created or delivered for public access in the Catalog is delivered primarily in the following file formats: PDF (Textual and Image), JPEG (Textual and Image), MP3 (Audio), and MP4 (Audio/Video) and ASCII (Datasets). Other file formats may be present depending on when they were added to the Catalog.

Format(s) Available for Reference: PDF, PPT, PPTX, ODP