

## File Format Guidelines

Rapid changes in technology mean that file formats can become obsolete at a rapid pace, causing problems for records management. The software in which a file is created usually has a default format, often indicated by a file name suffix (e.g., \*.PDF for portable document format). Most software allows authors to select from a variety of formats when they save a file (e.g., Microsoft Word Document [.docx] or Portable Document Format [.pdf] in Microsoft Word). You can also convert a record to another, more durable format.

The following table represents the digital formats that the Whitney Museum of American Art recommends for in-house preservation and long-term records retention. The record types included in this document are not exhaustive. Staff members producing specialized records may find that certain types of records are not covered by this document. Please contact the [Archivist](#) to discuss potential preservation strategies for such media. These guidelines classify formats into three categories:

1. **Recommended for long-term retention:** File formats that meet the minimum requirements for long-term retention, including documentation, wide adoption, transparency, self-containment, and use within the archival community. In most cases, these are the formats the State Archives itself uses to preserve electronic records.
2. **Acceptable for long-term retention:** File formats that do not meet the minimum requirements for long-term retention, but which come near to meeting the requirements and, for practical reasons, may be appropriate for long-term retention at some agencies. These formats are more likely to require frequent review and maintenance than formats recommended for long-term retention.
3. **Not recommended for long-term retention:** File formats that are not appropriate for long-term retention. Files saved in these formats should not be relied on to last more than five years. Electronic records whose retention periods are over five years should not be stored in these formats.

<b>Type of Record</b>	<b>Recommended for long-term retention</b>	<b>Acceptable for long-term retention</b>	<b>Not Recommended for long-term retention</b>
<b>Word processing documents</b>	PDF/a [.pdf]	Open Document Format [.odt]	WordPerfect [.wpd]Office 2003 and older [.doc and .xls]
		Open Office XML [.docx and .xlsx]	
		Rich Text Format [.rtf]	
<b>E-mail</b>	MBOX (entire folders or accounts) [.mbox]	Outlook individual email formats [.msg, .ost]PDF Portfolio [.pdf]	Outlook Personal Storage Table files [.pst]Opera Mail
		HTML [.htm, .html] EML [.eml]	[.mbs]Entourage [.rge]
<b>Audio</b>	WAV [.wav]	Audio Interchange File Format [.aif, .aiff]	Windows Media Audio [.wma]
		MP3 [.mp3]	RealAudio [.ra, .rm, .ram]
		Advanced Audio Coding [.aac, .mp4, .m4a]	Protected AAC [.m4p]
		Ogg Vorbis [.ogg]	
		Free Lossless Audio Codec [FLAC]	MIDI [.mid, .midi]
		MIDI [.mid, .midi]	
<b>Video</b>	AVI (Uncompressed, motion JPEG) [.avi]	Windows Media Video [.wmv]	Protected MPEG-4 [.m4p]
	Quicktime Movie (Uncompressed, motion JPEG) [.mov]	MPEG-2 (wrapped in AVI or MOV) [.avi, .mov]	RealVideo [.rv]
	MPEG-4 AVC [.mp4]	MPEG-4 (wrapped in AVI or MOV) [.avi, .mov]	Extended MIDI (.xmi)
<b>Presentation Files</b>	Convert to PDF/a	Open Office [.sxi, .odp]	PowerPoint 2003 or earlier [.ppt]
		Office Open XML [.pptx]	
<b>Images</b>	TIFF (Uncompressed) [.tif]	PDF/A or PDF/X (Graphic exchange format) [.pdf]	RAW digital camera images [.raw]
		PNG [.png]	Photoshop images [.psd]
		JPEG2000 (lossless) [.jp2]	Encapsulated PostScript [.eps]
		TIFF (Compressed) [.tif]	
		GIF [.gif]	
<b>Spreadsheets</b>	Comma Separated Values [.csv]	Open Office Format [.sxc, .ods]	Excel 2003 or earlier [.xls]
	Tab Delimited [usually .txt]	Office Open XML [.xlsx]	Microsoft Access Database [.mdb]
		dBASE [.dbf]	Lotus 1-2-3 [.wks]
<b>Structural Markup text documents</b>	SGML with DTD/Schema		XML without DTD/Schema
	XML (.xml) with DTD/Schema		SGML without DTD/Schema

SAP