

Bard Graduate Center Archives
Preferred file formats for long-term preservation

Updated May 24, 2016



Decorative Arts, Design History, Material Culture

Overview

The following table outlines the types of files that are managed by the digital records repository. File types that have poor long-term viability, such as files types that proprietary and not openly documented, are recommended to be normalized into formats that are well-documented and/or have better long-term viability. The original file should always be retained.

Note that this table should be updated as new file types are incorporated into the digital archives. The archives should not incorporate any files that it cannot make accessible to a researcher.

Media type	Submission file format	Normalized version for preservation	Notes
MS Office (2007 and after): Word, Powerpoint, Excel	DOCX, PPTX, XLSX	<i>Use original format</i>	These files are zipped-up XML files, so they work well for preservation.
MS Office (pre-2007): Word, Powerpoint, Excel	DOC, PPT, XLS	<i>Use original format</i>	Proprietary binary format, but MS has openly documented them, and they show no signs of being dropped from MS Office, so they continue to be acceptable.
Plain text	TXT	<i>Use original format</i>	
Rich Text Format	RTF	<i>Use original format</i>	
Filemaker	FMP, FP5, FP7	<i>Use original format, PDF, CSV</i>	Use options in File Maker to export PDF version and CSV versions. Put CSV files in a subfolder called: mydatabasename_normalized_csv

Portable Document Format	PDF	PDF/A or High Quality PDF	<p>Embed fonts; avoid using “Smallest file size” save option, which does not embed fonts.</p> <p>Use Adobe Acrobat’s Preflight to convert to PDF/A. If you have questions, please consult archivist. Prioritize conversion for visually important PDFs.</p>
Graphics - sustainable	JPG, TIFF, PNG	<i>Use original format</i>	
Graphics - low sustainability	PSD, BMP, GIF, PSD	TIFF	Flatten any layers
Adobe Illustrator	AI	PDF/A or High Quality PDF	<p>Save as “high-quality” PDF.</p> <p>After saving as high quality PDF, use Adobe Acrobat’s Preflight to convert to PDF/A. Prioritize conversion for visually important PDFs.</p>
Adobe InDesign	INDD	PDF/A or High Quality PDF	<p>Save as “high-quality” PDF.</p> <p>After saving as high quality PDF, use Adobe Acrobat’s Preflight to convert to PDF/A. Prioritize conversion for visually important PDFs.</p>
Audio	AC3, AIFF, MP3, WAV, WMA	<i>Use Original Format, see note.</i>	<p>Monitor formats for obsolescence issues. Migrate to WAV LPCM if format becomes endangered.</p>
Video	AVI, FLV, MOV, MP1, MP2, MP4, SWF, WMV, QT, M4V	<i>Use original format, see note</i>	<p>Use original format, but register the CODEC and Wrapper in the Video Codec Registry. Migrate to FFV1/LPCM in MKV if format</p>

			becomes endangered.
Camera raw files	CR2, CRW, RAW, DNG	TIFF	Use Photoshop or other tool that reads these files.
WordPerfect	WPD	PDF	Convert to PDF using MS Word for Windows or script wpd_to_pdf.vbs script (requires MS Word for Windows).
QuarkXPress	QXD	PDF/A or High Quality PDF	Convert to PDF using InDesign. Save as “high-quality” PDF. After saving as high quality PDF, use Adobe Acrobat’s Preflight to convert to PDF/A. Prioritize conversion for visually important PDFs.
SketchUp (2-dimensional drawing)	SKP (SKB is the backup of the SKP file)	PDF/A or High Quality PDF	Open in Sketchup, and Export as PDF. After, use Adobe Acrobat’s Preflight to convert to PDF/A.
SketchUp (3-dimensional model)	SKP (SKB is the backup of the SKP file)	<i>Use Original format</i> , PDF/A or High Quality PDF	For 3-dimensional models, a copy of SketchUp is required to read the original SKP file (although the PDF/A should still be created for easy reference to what is in the model).
Websites	HTML, various web assets	WARC, screenshots	Create WARC files using Webrecorder and use Webarchiveplayer to play back. Webrecorder should also be used for creating screenshots. If webrecorder screenshots are not sufficient, create additional screenshots as PNG files

			(e.g., use Awesome Screenshot)
Website files, local	HTML, various web assets	<i>Use original format, see note</i>	Common files used on the web, such as HTML, JPG, GIF, can be maintained as is. May need to normalize individual files if significantly endangered.
Gmail	ZIP file of inbox via Google Takeout	MBOX	Google Takeout uses MBOX, so nothing to do other than download the file from Takeout.

SAA Museum Archives Section Working Group Example