

Smithsonian Institution Archives Digital Preservation Formats

In accordance with best practices, SIA prefers to preserve transferred electronic records in the formats described below. These formats tend to be open, standard, non-proprietary, and well-established.

If a file cannot be created or saved in a preservation format, then the file will receive bit-level preservation (maintained as is) by SIA. An example of this could be a document in a proprietary format more than 30 years old that cannot be identified by current format detection tools and cannot be accessed. SIA will maintain the file in its current state.

When creating your files consider using these formats below as your original document (create/shoot TIFF images) or saving them in these formats (save your Word file as a PDF when complete) when possible.

Type	Primary Preservation Format (preferred)	Secondary Preservation Format (acceptable)
Text/word processing applications	PDF/A PDF	RTF (text) TXT XML with schema
Spreadsheet applications or structured data	PDF/A (must capture entire workbook – macros disabled) PDF	CSV Tab-delimited TXT XML
Presentations	PDF/A PDF	Original
Images	TIFF (uncompressed)	JPG DNG PNG JP2
Graphics	TIFF	PDF
Video	Motion JPEG 2000, MOV, AVI	MPEG-4
Audio	BWF-Broadcast WAV (<i>.wav is the extension</i>)	WAV AIFF FLAC
Websites and social media records	WARC	ARC Files from Content Management System
Email messages/account	XML email preservation format - Consult SIA Digital Services Division	Consult SIA Digital Services Division
Database Management Systems (DBMS)	Keep original	XML with schema
CAD	PDF/A, PDF/E or PDF with original file	Original
Other	Consult SIA Digital Services Division	