

File Formats Guide

The following is intended to provide guidance for content creators and those depositing content into one of the Indiana University Libraries' repositories. Specific format requirements and policies may be defined at the repository level, so also check the [appropriate repository-level documentation](#) and/or contact staff prior to creating/submitting content.

For more information on specific file formats and guidance on selection, please see the [Library of Congress' Sustainability of Digital Formats](#) and [Archivem atica's Format Policies](#) page.

Content Type	Best Practice Recommendations for Preservation	Other Supported File Formats	Further Resources
Textual	Portable document format for long-term preservation (PDF/A) Plain text (TXT)	Microsoft Word file (DOC) (DOCX) OpenDocument text (ODT)	
Images	TIFF version 6 uncompressed (TIFF)	JPEG Portable network graphics format (PNG)	
Video	FFV1 in Matroska (MKV)	Audio video interleaved format (AVI) QuickTime (MOV) MPEG-4 (MP4)	
Audio	Waveform audio file format (WAVE)	Free lossless audio codec (FLAC)	
Data: Quantitative	Comma-separate values file (CSV)	SPSS file formats STATA file formats Microsoft Excel file (XLS) (XLSX) Tab-delimited file (TSV)	UK Data Archive Managing and Sharing Data
Data: Qualitative	eXtensible mark-up language (XML)		
Data: Geospatial		ESRI Shapefile (SHP, SHX, DBF) Geo-referenced TIFF (TFW) CAD data (DWG) Tabular GIS attribute data	
Web	Web ARChive file format (WARC)		
Email		Microsoft Outlook file (PST)	
Social Media	eXtensible mark-up language (XML)	JSON	Digital Preservation Coalition's Preserving Social Media
3D Objects	Contact staff		
e-Learning Objects	Contact staff		JISC Significant Properties of e-Learning Objects Report
Corresponding Documentation	See Textual Content Type		

The columns are arranged as follows:

Best Practice Recommendations for Preservation - These formats are agreed upon by the digital preservation community as being the best for ensuring long-term preservation of a digital object based on a variety of sustainability factors. If it is possible, content should be migrated and/or saved utilizing these formats. Indiana University Libraries' Digital Preservation Unit is highly confident in our ongoing ability to steward content in these formats.

Other Supported File Formats - These formats are well-known and highly utilized, but they are not preservation standard. If possible, content in these formats should be migrated into the best practice recommendations for the specified content type; however, we understand that in certain cases this might result in the lost of necessary functionality. Indiana University Libraries' Digital Preservation Unit is relatively certain that we will be able to provide continued access and stewardship of content in these formats.

File Formats Not Listed - The formats listed above are the main file formats for which the Indiana University Libraries' Digital Preservation Unit can provide ongoing stewardship. Formats not listed are therefore not recommended to content creators and managers for content intended to be preserved in one of the IU Libraries' repositories. If you have content in an unlisted format, please contact the repository administrator to determine whether content should be normalized into a more preservation-ready file format.