

DataverseLV Preferred data formats

To ensure the long-term preservation and accessibility of research data, DataverseLV strongly encourages researchers to deposit data in preferred file formats. Choosing the right format is essential for future readability, allowing data to be reused long after the original software used to create it becomes obsolete.

Characteristics of Preferred Formats

We prioritize formats that support the FAIR data principles (Findable, Accessible, Interoperable, Reusable). Specifically, we recommend formats with the following features:

1. **Non-proprietary and Non-commercial:** Formats should be freely available and usable without requiring purchased licenses or specific commercial software. This ensures access regardless of changes in the commercial software market.
2. **Open and Documented Standards:** Formats should be based on publicly available, standardized technical specifications. This allows various systems to process the data without restriction.
3. **Standard Character Encoding:** Files should use standard encoding, preferably Unicode (UTF-8). This eliminates incompatibility issues and ensures text is displayed correctly across different languages and platforms.
4. **Uncompressed:** Whenever possible, data should be uncompressed to prevent potential corruption and remove dependencies on specific decompression tools. Uncompressed data facilitates easier processing and safer long-term storage.

Adhering to these standards ensures your data remains secure, easily accessible, and widely usable for future research.

File type	Preferred formats	Acceptable formats	Acceptable formats
Text documents	• PDF/A (.pdf) • ODT (.odt) • <i>Markdown</i> (.md) • XML (.xml)	• RTF (.rtf) • PDF (.pdf) • <i>OpenOffice XML</i> (.docx)	• <i>Microsoft Word</i> (.doc) • <i>Google Docs</i> (.gdoc)

File type	Preferred formats	Acceptable formats	Acceptable formats
	• <i>Unicode Text (.txt)</i>		
Plain text	• Unicode text (.txt)	• <i>Non-Unicode text (.txt)</i> • HTML (.html)	
Presentations	• PDF/A (.pdf)	• <i>OpenOffice PowerPoint XML (.pptx)</i> • <i>OpenDocument Presentation (.odp/.sxi)</i>	• <i>Microsoft PowerPoint (.ppt)</i> • <i>Google Slides</i>
Data tables	• CSV (.csv) • TSV (.tsv) • <i>Delimited text (.txt)</i> • TAB (.tab)	• ODS (.ods) • <i>OpenOffice (.docx)</i> • <i>XML Workbook (.xlsx)</i> • SPSS (.sav) • R (.RData, .rds)	• <i>Microsoft Excel (.xls)</i> • PDF (.pdf/.pdfa) • SPSS Portable (.por) • SAS (.sas7bcat) • <i>Stata (.dta)</i> • <i>Matlab (.mat)</i>

File type	Preferred formats	Acceptable formats	Acceptable formats
Database	• SQL (.sql) • SIARD (.siard)	• SQLite (.sqlite) • <i>dBase</i> (.dbf) • CSV (.csv)	• <i>Microsoft Access</i> (.mdb, .accdb) • HDF5 (.hdf5, .h5)
Data from statistical analysis	• <i>Delimited Text</i> (.txt) • CSV (.csv)	• R (.R, .RData, .rds) • JSON (.json)	• SPSS (.sav, .sps) • <i>Stata</i> (.dta) • SAS (.sd2, .7dat, .tpt) • JASP (.jasp)
Audio	• FLAC (.flac) • BWF (.bwf) • MXF (.mxf) • WAVE (.wav) • <i>Matroska Audio</i> (.mka)	• AIFF (.aif) • MPEG-4 <i>Audio</i> (.mp4, .m4a) • <i>Ogg Vorbis</i> (.ogg) • MP3 (.mp3)	• AAC (.aac, .m4a) • <i>Monkey's Audio</i> (.ape) • WMA (.wma)
Video	• <i>Matroska Video</i> (.mkv) • MXF (.mxf)	• MPEG-4 (.mp4, .m4v) • <i>QuickTime</i> (.mov, .qt)	• WMV (.wmv)

File type	Preferred formats	Acceptable formats	Acceptable formats
	• AVI (.avi)	• MPEG-2 (.mpeg, .mpg) • <i>WebM</i> (.webm)	
Pictures	• TIFF (.tif, .tiff) • PNG (.png) • SVG (.svg) • DICOM (.dcm)	• JPEG (.jpg, .jpeg) • GIF (.gif) • BMP (.bmp) • JPEG 2000 (.jp2)	• <i>Adobe Photoshop</i> (.psd) • <i>Apple Picture</i> (.pct) • RAW formāti (piem., .cr2, .nef, .arw, .raw)
Vector files	• SVG (.svg) • PDF/A (.pdf)	EPS (.eps)	• <i>Adobe Illustrator</i> (.ai) • WMF/EMF (.wmf, .emf) • <i>CorelDRAW</i> (.cdr)
Geographic Information Systems (GIS)	• <i>GeoPackage</i> (.gpkg) • GML (.gml)	• <i>Esri Shapefile</i> (.shp + datnes) • GeoJSON (.json)	• <i>MapInfo</i> (.tab) • KML (.kml, .kmz)

File type	Preferred formats	Acceptable formats	Acceptable formats
	• GeoTIFF (.tif, .tiff)	• MIF/MID (.mif/.mid) • WKT (.wkt)	• <i>Esri Geodatabase</i> (.gdb) • MXD/WOR/QGS (.mxd, .wor, .qgs) • <i>ESRI Interchange Format</i> (.eoo) • CAD (.dwg)
Archives	• ZIP (.zip) • TAR (.tar)	• gzip (.gz) • 7z (.7z) • CPIO (.cpio)	• BZIP2 (.bz2) • RAR (.rar)
Analysis of qualitative data	• XML (.xml) • <i>Unicode text</i> (.txt) • CSV (.csv) • RTF (.rtf)	• JSON (.json)	• <i>Atlas.ti</i> (.atlpri) • NVivo (.nvp, .nvpj)