GRIIDC Preferred File Formats

GRIIDC encourages data providers to submit data in formats that promote re-use. Data should be submitted in non-proprietary (open) file formats whenever possible to facilitate long-term data storage and curation. In general, text-based file formats are preferred over proprietary binary formats, as text files will remain usable in the future, while the knowledge and/or software to work with binary formats can be lost.

A dataset can include the same data in both proprietary and non-proprietary formats. Presented below are the preferred, acceptable, and not recommended file formats for common data types submitted to GRIIDC.

Tabular/Spreadsheet

Tabular or spreadsheet data are the most common data types submitted to GRIIDC. While Microsoft Excel is an acceptable format, a CSV file is preferred. To convert an Excel spreadsheet to CSV, simply click Save As, and Save as type: CSV (Comma delimited). Please note that this will only save one worksheet of data. If the file contains multiple worksheets of data, each worksheet will need to be saved as a separate CSV file.

Preferred File Formats	Acceptable File Formats	Not Recommended File Formats
Comma Separated Values (.csv)	Microsoft Excel (.xls, .xlsx)	Adobe Portable Document Format (.pdf)
Tab Separated Values (.tsv)	Origin (.opj)	Microsoft Word (.doc, .docx)
		Microsoft PowerPoint (.ppt, .pptx)
		Microsoft Excel Macro-Enabled Spreadsheet (.xlsm)
		Microsoft Excel Open XML Format Spreadsheet (.xlsl)
		Microsoft Access (.mdb)
		SPSS (.sav)
		Sigma Plot workbook (.jnb)

Text

Many software packages save data in text files, which may or may not use the .txt extension. Common examples include .dat files, used by many programs, and CTD data in .cnv files; it is not necessary to rename these file extensions to .txt. Text files are also often submitted to GRIIDC as ReadMe files or files containing supplementary information about the data. GRIIDC encourages the submission of ReadMe or supplementary files to provide additional information about the dataset.

Preferred File Formats	Acceptable File Formats	Not Recommended File Formats

Text (.txt)	Text, ASCII (.txt)	Microsoft Word (.doc, .docx)
eXtensible Mark-up Language (XML) (.xml)		Rich Text Format (.rtf)

Image Files

Image files can be submitted as data, providing that additional explanatory information is also submitted describing the contents of the image files. For example, use supporting documentation to describe a specific place, time, or depth an image was taken, or any experimental conditions used.

Preferred File Formats	Acceptable File Formats	Not Recommended File Formats
TIFF file (.tif)	JPEG (.jpeg, .jpg)	Graphic Interchange Format (.gif)
	JPEG 2000 (.jp2)	
	Adobe Portable Document Format (PDF) (.pdf)	
	Portable Networks Graphic (.png)	
	Bitmap (.bmp)	

Video Files

Video files are like image files in that additional explanatory information are required to be submitted describing the contents of the video file within a ReadMe text file.

Preferred File Formats	Acceptable File Formats	Not Recommended File Formats
QuickTime Movie (.mov)	MXF (.mxf)	Windows Media Video (.wmv)
AVI (.avi)		
MPEG (.mpg, .mp2, .mp3, .mp4)		

Audio Files

Audio files are like image and video files in that additional explanatory information are required to be submitted describing the contents of the audio file within a ReadMe text file.

Preferred File Formats	Acceptable File Formats	Not Recommended File Formats
WAV (.wav)	MPEG-3 (.mp3)	Windows Media Audio (.wma)
AIFF (.aif, .aiff)		
Free Lossless Audio Codec (FLAC)		

Geospatial Data Files

Geospatial data files submitted as data should include the spatial reference/coordinate system definition so the data can be properly located. The attribute table, if present, should include documentation of field types, values, and units as with a non-spatial tabular dataset.

Vector

Preferred File Formats	Acceptable File Formats	Not Recommended File Formats
Shapefiles (.shp/.shx/.dbf)	KML/KMZ	ESRI file geodatabase
	GeoJSON	e00
	GML	MapInfo (.tab/.map)

Raster/Gridded

Preferred File Formats	Acceptable File Formats	Not Recommended File Formats
GeoTIFF (.tif, .tiff)	JPEG (.jpeg, .jpg)	MrSID
NetCDF (.nc)	ESRI ASCII Grid	ECW (.ecw)
Hierarchical Data Format (HDF)		ERDAS Image (IMG) (.img)
		ESRI Binary Grid

CTD Data Files

Provide the raw data files (.hex, .hdr, .cnv, .XMLCON, and .bl). Processed data can also be provided in any of the preferred tabular, text, or raster/gridded formats.

ADCP Data Files

Provide the raw data files. Processed data is accepted, only as non-proprietary.

Do not submit formats such as .PO0 unless they are accompanied by a process file in any of the preferred formats above (e.g., netCDF or ASCII).