



Preferred formats in DTU Data

Open file formats

DTU Data accepts data in all file formats. However, to ensure long-term accessibility and readability¹, you should choose open file formats for deposited files.

Open file formats are recognised by the source code being open and therefore, often in non-proprietary file formats. If you save data in open file formats this will

- 1) make it more easy for you and others to use the files in the future, if there is no longer access to the proprietary software with which to process the files and
- 2) make the file content more accessible to others that don't have access to proprietary software.

If you aim for long-term preservation, data must be uploaded in an open or preferred file format. All files must have a valid file extension, e.g. .txt, .pdf. If your data cannot be stored in a preferred format, they can still be published in their original format, but in that case, DTU Data only commits to preserve the data at bit-level² in the long term. If appropriate, the file may also be archived in their original file format **in addition to** preferred format(s).

Proprietary formats

When saving and publishing data in proprietary formats, consider including an explanatory readme.txt file with the data. Include information about name, version and original use of software used to generate the files. This may be necessary information to handle files in the future.

Preferred formats

File formats suitable for full preservation ³ (examples)	
Containers	TAR, GZIP, ZIP
Databases	XML, CSV
Geospatial	GeoTIFF, NetCDF
Sounds	WAVE, AIFF, MP3
Statistics	ASCII, DTA, POR, SAS, SAV
Video	MPEG-4
Images	TIFF, JPEG, PDF/A, PNG
Tabular data	CSV, tab-delimited values
Text (slides, illustrations)	PDF/A (and original file)
Text	Plain text, XML, PDF/A
Array data	netCDF

¹ See "Guide to preservation in DTU Data" at DTU Inside

² Bit-level preservation: Access to the file in its submission format is provided

³ Full preservation: Usability of files will require actions such as migration, normalization and conversion. Preferred formats is a prerequisite for such actions.



The table gives examples of preferred formats for long-term preservation in DTU Data. It is not exhaustive and exclusive.

Contact datamanagement@dtu.dk before you prepare your research data for long-term preservation.

Version date: 2022-11-04

Text and table is adapted from: <https://guides.library.stanford.edu/data-best-practices/format-files>,
https://data.4tu.nl/info/fileadmin/user_upload/Documenten/Preferred_File_Formats_2019.pdf,
<https://site.uit.no/dataverseno/deposit/prepare/#preferred-file-formats>