SHARING DATA

tips and best practices

ArcheoLogica Data publishes datasets from any type of archaeological investigation. We believe that the data lifecycle should not end with their publication, and that accessible and reusable data are a valuable resource for research.

The papers accompanying the datasets explain how the data were created or digitized, what they represent, what their content and structure are, and any manipulations performed. *What does a new user need to know when using the data for the first time?*

checklist for an irresistible dataset

- The type of data is clear and defined. Datasets may contain raw data (data as originally collected); processed data (unified and reorganised, or analyses, visualisations); documentation, programming codes.
- There are no empty folders; files and folders are clearly organised, without excessively long paths.
- File nomenclature is uniform; individual file names are unique and contain no special characters, with the exception of underscores and hyphens ('_' and '-').
- In the case of highly specialised data, the specific software required for reading and analysis is specified in the accompanying documentation.
- Some file formats are more accessible than others; in general, it is preferable to use standardised, documented and, when possible, non-proprietary formats. The table outlines the best solutions and acceptable alternatives depending on the type of data.

	Best solution	**Acceptable solution**
Tabular data	CSV, TAB, TXT, JSON	XLS MDB/ACCDB, DBF, ODS
Textual data	XML, RTF, TXT	HTML, DOC, ODF
Documentation	RTF, HTML, ODT	DOC, PDF
Geospatial data	SHP, GEOTIFF, DXF, GeoJSON	MDB, KML
Image data	TIF	JPEG, TIFF, RAW