



## Guidance for Databases and Spreadsheets

When creating databases staff should consider long-term preservation and accessibility of the data to others. The data will be archived converted into delimited text (comma separated value, tab separated value or XML.)

### Creation

- Do not use common delimiters in the database (e.g. commas, tabs or pipes), as this could cause difficulties in migrating the data to delimited text. If they are used this should be identified and used to guide migration.
- Assign columns, rows, worksheets and tables meaningful names.
- Use controlled vocabularies and established word lists wherever possible.
- Clearly describe and define any keys for data description and entry.
- Always define any data values used in the dataset or metadata:

Height (Metres)	not	Height
1.2		1.2

- Try to use codes as little as possible. If they are required; ensure they are documented.
- Use data controls, such as look up tables, in the software to assist standardise data entry as much as possible and assist to ensure data quality.
- Should you be creating a similar datasets regularly (personally or within your team) you should consider ensuring that there is a standardised recording structure and its use should be added to internal procedures.
- Standardization of datasets (structure, names for tables etc and terminology) and the data within them will make creating metadata easier, because the initial metadata form created will become a template for future datasets.
- Do not embed objects into databases (e.g. images). These should be archived separately.
- Do not create databases for which the understanding is dependant on styles or formatting (e.g. cells filled pink mean objects have been photographed). This information will be lost on migration to delimited text.
- Layout and formatting (merged cells) can also cause problems on migration; if these are required it better preserved as a PDF/A1-A.
- Migration will only preserve the results of complicated formulae, not the formulae; if these need to be preserved they will need to be additionally documented.

### Archiving

- It is policy to convert databases and spreadsheets to delimited text, with the structure documented in the metadata.
- Any graphics, figures or charts created **within** the dataset should be exported as image files if required for archiving.
- On migration of the data to a delimited format a number of checks should be undertaken:
  - Check delimiters have not been used

For further guidance see: Archaeology Data Service, 2013, *Caring for Digital Data in Archaeology: A Guide to Good Practice*, Oxford: Oxbow Books



## Guidance for Databases and Spreadsheets

- Column count – check the use of common delimiters has not caused the addition of columns
- Truncation – check not fields have been truncation on migration

### ***File names***

File naming conventions for databases and spreadsheets exist in the *File Naming Convention*.