

Managing your digital material: some good practice guidelines for donors and depositors



The best way of ensuring that digital material you create is suitable for long term preservation within an archive is to manage it effectively from the point that you create it. What follows are some simple data management dos and don'ts, designed both to help you to stay in control of your own data and to make that data easier for others to understand in the future.

Do weed out duplicate files or out-dated versions unless there is a good reason for these to be kept for the long term. It is useful to get into the habit of regularly clearing out duplicate, spurious or unnecessary files in order to keep your directories tidy.

Note: in some cases, it may be necessary to be able to look back at previous versions of a document in order to trace the development of a document, or study the creative thought process. In many situations however it will only be necessary to keep the final version. Where it is necessary to store earlier versions, ensure a version or draft number is included in the filename so that the order in which the files were created is clear.

Do ensure that you organise your files into a meaningful directory structure. Use as many directories and folders as you need in order to organise your digital content logically. Ensure these directories are named in a useful way to help give context to the files that they contain.

Note: if you have a selection of documents containing the minutes of meetings of a committee over the course of several years, you may wish to put these in a directory named after the committee, in a sub-directory called 'minutes', and further directories named by the year the meeting took place. There are no hard and fast rules here, so find a directory structure that makes sense to you and apply it consistently.

Do ensure that you name your files sensibly. When named well, your file names act as a short descriptive caption for your files, briefly explaining what the file contains. This will help you and your family or colleagues to locate and understand your files. It will also be invaluable in the future for the archivists who may have the job of curating and preserving your files.

Note: where you are using dates at the start of file or directory names, we recommend you use the format yyyy-mm-dd (e.g. 11th July 2014 = 2014-07-11) for recording dates. This ensures that your files can be listed chronologically by file management tools.

Use only alphanumeric characters (a-z, 0-9), the hyphen (-) and the underscore (_) in your filenames. Full stops should only be used to separate the filename from its file extension. Spaces can cause problems if your files are moved to a different operating system or made available online. It is good practice to replace them with the underscore (_) character.

*Note: Windows users may not realise that the case of the file names they select are significant. A Windows PC would see no difference between the file names **minutes.doc** and **Minutes.doc** but to a*

case-sensitive operating system these would refer to two different files. To avoid confusion, we recommend using lower case file names.

Ensure your files are *self-documenting* where possible. As a minimum a document should have a title, author and date, but other information such as document history, purpose and status are also very useful. If this information appears on the first page of a document this makes it much easier for you (and anyone who curates or uses the data in the future) to understand the content. For some types of digital material, extra documentation may be required, for example for a database or spreadsheet where codes or abbreviations need to be interpreted and relationships between data tables described. For images, captions or titles may need to be recorded separately. It is far easier to document your files at the time of creation as it is easy to forget important details about a file (for example where a photograph was taken) if you leave it until a later date.

Note: It is also possible to store data about your documents (metadata) within the file itself. Many desktop applications include a 'properties' option where you can record title, author, keywords and comments about your files.

Do save copies of your *original* digital files from the creating software or hardware. Derived files may be poorer quality or have lost some of the information that was in the original. From a digital archiving perspective, original files are generally easier to work with and preserve for the long term.

Note: We would prefer to receive a Microsoft Word document than a pdf version of the same file. If you are depositing photographs from a digital camera, we prefer to receive the image that was originally captured rather than a version which has been saved in a different format or at a lower resolution.

Do save your digital material in file formats that are widely used and widely supported. These files are easier for an archivist to work with than obscure or less well-known file formats. File formats that are open, well documented and uncompressed are also likely to be easier to preserve in the long term.

Do look after your data. Protect your data by backing it up regularly and ensuring anti-virus software on your computer is up-to-date.

Plan for the future of your digital data at the point of creation. Embracing good practice in your day-to-day work is the most effective strategy you have to ensure that your data is suitable for archiving in the long term. **Don't leave it too late!**

Please contact the Digital Archivist at the Borthwick Institute with any queries you have about preparing your digital data:

Borthwick Institute for Archives University of York Heslington York, YO10 5DD	E-mail: borthwick-institute@york.ac.uk Telephone: 01904 321166
--	---