

Data Management Plan

The data management plan describes how the data will be managed throughout the lifecycle of the research. Providing this documented approach will aid in well organised documented data that will demonstrate accuracy and validity. The data may have a longer lifespan than the research project. A content checklist is provided by Donnelly *et al* (2009).

Data Documentation

The documentation will include:

- The aims and objectives for each type of collection (quantitative and qualitative)
- The research question
- The methods will be documented to include the instrument used, sampling and data set structure.
- Data validation, cleansing and any quality assurance tests carried out.

A catalogue of data will be produced.

Format	Name	Size in KB	Description
.csv	Dataset.csv	1	Data Set of xyz
.pdf	Questionnaire.pdf	1	Questionnaire questions delivered
.rtf	Report.rtf	1	Technical Report
.mdf / .ldf / .bak	Surveydatacollection1.mdf	100	SQL Server Database

Table 1 Data documentation amended from Van den Eynden *et al* (2011, p.10)

A metadata description for each data entry in the catalogue will include the origin, purpose, time created, creator and the location of the collection.

Storage Formats

The file formats to be used for this research recommended for UK data archiving for long term preservation of research data are

Type of Data	Recommended File Formats for Sharing, Reuse and preservation
Quantitative tabular data	Comma separated values (.csv)
Qualitative data textual	Rich Text Format (.rtf) Plain text data (.txt)
Documentation	Rich Text Format (.rtf) PDF (.pdf)
SQL Server Database (Not shared – internal use only)	Data (.mdf), log (.ldf) and backup file (.bak)

Table 2 File Formats amended from Van den Eynden *et al* (2011, p.12)

File Formatting and Version Control

To ensure order and logical structure is maintained basic processes will be followed to ensure short meaningful names are used and spaces and special characters are not used in the names. The naming of documents will include the classification type, a short description, date and time. This convention will represent the way version control numbering system works. An example of the naming structure is 'dataset_name_0711111956.csv'.

Quality Assurance

Quality control methods will be put in place to verify the calibration of instruments. One such method is the survey design being tested by a pilot survey. Careful checking of data and validation will help ensure the data recorded is of high controlled quality. For example transcripts of the semi structure interviews will have the following header:

Version Control Table for a data file	
Title:	
File Name:	
Description:	
Created By:	
Created:	
Last Modified:	

Table 3 Version Control for a data file amended from Van den Eynden *et al* (2011, p.16)

Data Storage and Backups

All critical files will be backed up in multiple locations to protect the data against hardware failure, virus infection and human error.

The data is initially stored during the research period on a local PC hard disk and local external backup disk. Also files are transferred into the cloud for additional backups through Googledrive <http://docs.google.com> and Dropbox <http://www.dropbox.com> . Weekly scheduled and manual backups are undertaken.

Data Security

The computer systems where the data is stored are protected by passwords and the local PC has up to date virus definitions, up to date security patches for the operating system applied and the firewall is enabled.

If any personal data is collect it will be encrypted using <http://www.truecrypt.org/> and any passwords stored using <http://keepass.info/> .

Data Disposal

The data will be disposed of on the hard drive and any flash drives via physical destruction when that point arises.

Ethics

When the data is initially collected consent will be obtained to provision data sharing and reuse. Also any items within the data which would identify the people who completed the research will be anonymised in the data. The data fields to be anonymised will be

documented. Legislation that may affect the data in the research project is the Data Protection Act and Freedom of Information Act. The types of data to be collected may be personal data, confidential data and sensitive personal data. A plan will be put in place to control access to the data.

On the quantitative survey a sample of the text to be used to inform participants of the data storage requirements, consent for data sharing and reuse and how the data will be anonymised is on <http://www.sqltoolkit.co.uk/survey.htm> and below taken from an OU forms survey:

Code of Practice for the Research

Please read this before completing the survey.

- The Data will be used only for the purposes of this research.
- No data in the final research will be personally identifiable.
- If articles are published in journals it may be necessary to send the anonymised data to the publisher to verify the results.
- If you enter your email address with your answers, your contributions can be identified and destroyed if you choose to withdraw consent. You have a right to withdraw consent at any time with no resultant adverse consequences.
- Partially completed data is saved and if you decide to not continue completing the survey part way through it will not be possible to identify that data and as such it cannot be deleted.
- You can answer all questions in this survey anonymously and although Survey Monkey can store the respondent's IP address, this survey is configured so the respondent's IP address will not be stored in the survey results to protect anonymity. Anonymous contributions are not personally identifiable so your contributions cannot be linked to you and so you won't be able to withdraw consent for anonymous contributions.
- By participating in this questionnaire you are agreeing to give consent for your contributions to be used for the study.
- The study will abide by the requirements of the Data Protection, Freedom of Information and Copyright and Licensing Acts.
- Once the research is concluded the anonymised research data will be available as an open data set on this website in order to make the data available to other researchers in line with current data sharing practices.

Survey Monkey

Please note that the online survey is hosted by "Survey Monkey" which is a web survey company located in the USA. All responses to the survey will be stored and accessed in the USA. This company is subject to U.S. laws, in particular, to the U.S. Patriot Act that allows authorities access to the records of internet service providers. If you choose to participate in the survey you understand that your responses to the questions will be stored and accessed in the USA. Survey Monkey can store the IP address of those accessing its site and so could determine the location of the originator's computer. However as mentioned above this option is disabled. If you would like more on Survey Monkey go to their [privacy policy](#).

Access Control

Any data collected will be accessed by the researcher and could be requested to be viewed by the Institution and the Sponsor during the research period. Once the research is concluded the anonymised research data will be available as an open data set on this website in order to make the data available to other researchers in line with current data sharing practices.

References

Van den Eynden, V. , Corti, I. , Woollard,M., Bishop,L. and Horton,L. (2011) Managing and Sharing Data [online]
<http://www.data-archive.ac.uk/media/2894/managingsharing.pdf> (Accessed 5 November 2011)

Donnelly, M. and Jones, S. (2009) Data Management Plan Content Checklist [online]
http://www.dcc.ac.uk/sites/default/files/documents/templates/DMP_checklist.pdf (Accessed 5 November 2011)