Data Management Planning: Exercise

A Data Management Plan (DMP) outlines how data will be created, managed, shared and preserved. It helps you save time and effort, check that necessary support is in place, enables sound decisions, demonstrates awareness of good practice and reassures funders that the proposal is in line with their data policy.

1. Read the research project scenario.
2. Identify the potential data management issues (use the left-hand column on page 2 of this handout to list and number each issue you identify)
3. Begin to develop a DMP (by mapping each issue number to the DMP template headings in the right hand column on page 2 of this handout).
4. The ESRC guidance may help you consider the issues (p.3-4 of this handout).

Research project scenario

You are a Principal Investigator at the University of Manchester and plan to lead a 5 year research project involving the Universities of Leeds, Sheffield and Manchester, along with international collaborators from the Universities of Geneva, Bonn, Melbourne, Michigan, and Oslo.

The programme of work will study how children learn to communicate. The first four streams will address four key questions in language and communicative development. The fifth stream will provide the interconnection between the other four streams. It will follow 50 English-learning children from 6 months to 5 years to deliver a data-intensive, cohort study on language development in UK children.

The project will involve a number of analytical and research methods and methodologies that will be used to collect and analyse data. These will consist of neurophysiological measures, artificial grammar learning, eye tracking and looking-while-listening/preferential looking measures, forced choice pointing and hidden object search tasks, corpus analysis, elicited production, standardized tests, and computational models. Data and analyses need to be shared between all organisations throughout the five year programme.

Existing data provided by project partners will be used in some work packages, as well as data that is available under a Creative Commons CC-BY-NC-SA (Attribution-NonCommercial-ShareAlike) License. Existing data will be re-analysed and the new analyses stored as Excel and SPSS datasheets in the project archives.

All data will be analysed, stored and documented, and will include MS Excel, SPSS, R Studio and video files. All data will be collected where possible in a paper-free method directly onto laptops or hand-held devices. Video-data will be recorded using the hard drive of the digital video-camcorder and then transferred to the devices. Paper copies and manuals generated from the research will be stored.

Intellectual property and copyright are held across the partner institutions of Leeds, Sheffield and Manchester.

Supplementary resources: Jones, S. (2011). 'How to Develop a Data Management and Sharing Plan'. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: http://www.dcc.ac.uk/resources/how-guides

Summary of issues for Data Management Plan

List of issues	DMP template headings
	Assessment of existing data
	Information on new data
	Quality assurance of data
	Backup and security of data
	Expected difficulties in data sharing
	Copyright/intellectual property right
	Responsibilities
	Preparation of data for sharing and archiving

ESRC data management plan: guidance for peer reviewers (extract)

Assessment of existing data

You may want to consider the following questions:

- Is there evidence that secondary sources of data have been considered and evaluated?
- Is there evidence presented that the project is not creating new data when there are existing resources that could be re-used?
- If existing data are used, have issues such as copyright or IPR of such data been considered and possible copyright clearance obtained to be able to share data or data derived thereof?

Information on new data

You may want to consider the following questions:

- Is the information on data to be produced adequate and realistic and according to the research and methodology proposed in the application?
- Is there evidence that the plan covers all data that is planned to be generated from the research?
- Is sufficient information given on how data will be collected and in which formats (eg Open Document Format, tab-delimited, Excel etc) data will be analysed and stored, as well as an indication of how they will be documented?

Quality assurance of data

You may want to consider the following questions:

- Is information given on procedures for quality assurance that will be carried out on the data collected? (Please refer to the Case for Support for full information on quality control of the proposed research.) This could include methods for data validation or standards applied during data collection and data entry, codes of research practice adhered to, transcription templates used, etc.
- Are no quality assurance procedures mentioned when there is a clear need from the proposed research that there should be? Please note that quality issues are to be addressed at the time of data collection, data entry, digitisation or data checking.

Backup and security of data

You may want to consider the following questions:

- Is the data back-up procedure described fit for purpose? eg considering back-up procedures for all institutions involved in research and considering back-up frequency
- Are multiple media and multiple copies considered for back-up?
- Are measures considered to check the usability of back-up copies?
- Is information given on an institutional and/or local centre back-up policy?
- If sensitive data (ie detailed personal data) are collected, is there evidence that appropriate security measures in line with the Data Protection Act are considered when handling and storing data? eg encrypting data, anonymising data, care when transmitting data
- Is there evidence presented that proposed measures reflect existing best practices?
- Are methods of version control described? (ie making sure that if the information in one file is altered, the related information in other files is also adopted, as well as keeping a track on a number of versions and their locations)

Expected difficulties in data sharing

You may want to consider the following questions:

- Have all obstacles to sharing data been considered?
- Have strategies been considered for dealing with these issues? For example by:
    - o discussing data sharing and re-use with interviewees and gaining specific consent from participants to share research data
    - o anonymising data to remove personal and disclosive information
    - o regulating access to data

If there are ethical issues which may cause difficulties in data sharing, strategies for dealing with these issues should be discussed in the relevant section in the Je-S form. In assessing this part of the application you may want to refer to the requirements of the ESRC Framework for Research Ethics (www.esrc.ac.uk/researchethics).

If newly generated data cannot be shared, adequate justification should be given. It may be a case that parts of the data that are sensitive cannot be shared, but this should be considered critically and the plan should provide evidence that it has been assessed from all angles.

We regard a waiver of deposit as an exception, and reserve the right to refuse waivers where there is insufficient evidence that the applicant has fully explored all strategies to enable data sharing and archiving.

Copyright/intellectual property right

You may want to consider the following questions:

- Is copyright of research data (both existing sources of data used or created) agreed or clarified, especially for collaborative research or if various sources of data are combined?
- Are plans in place for copyright clearance for data sharing (if possible)?

Responsibilities

You may want to consider the following questions:

- Have data management responsibilities been allocated to named individuals?
- Is there evidence that data management will be followed throughout the course of the project?
- Has consideration been given to the variety of data management tasks that may be required for the research?
- For collaborative research, are data management responsibilities allocated at each partner organisation (if needed for the research) or has the coordination of data management responsibilities across partners been considered?

Preparation of data for sharing and archiving

The following questions may be considered when assessing this section of the plan:

- Are the plans for preparing and documenting data for sharing and archiving with the UK Data Service appropriate?
- Is there evidence that data will be well documented during research to provide high-quality contextual information and/or structured metadata for secondary users? eg documenting the method of data collection, origin, circumstances, processing and analysis of data