Data Management Checklist

The Data Management Checklist can be used in a number of ways to assist with data management planning. The checklist identifies a series of issues and underlying questions that should be considered for the development of a data management plan as well as to raise awareness of good practice when planning for data management during the life cycle of a research project. The Checklist is aligned with the Data management Plan Template and either can be utilised. The Checklist is informed by the requirements set out in the Australian Code for the Responsible Conduct of Research, the UK Digital Curation Centre Checklist and current University Policy.

To use the Checklist effectively to develop a data management plan you should provide detailed and descriptive responses to the following questions.

1. Research Project

☐ What is the project title, project description and date commencing and estimated date for completion?
☐ What is the name of the Chief Investigator, Partner Investigators, or Supervisors?
☐ Who is the primary contact for the data?
☐ Who is the lead organisation and are there other partner organisations?
☐ What School or Centre is the project associated with?
☐ What is the name of the Chief Investigator and partner Investigators?

2. Funding Body

☐ What are the names of relevant funding bodies and their grant application numbers.

3. Data To Be Produced

☐ What type of data will be produced, collected, generated or captured during the project?
☐ How will the data be captured, collected or created?
☐ What tools, instruments, equipment, hardware or software will you use to capture, produce, collect or create the data?
☐ What are the expected file formats of the data that will be captured, produced or created?
☐ Are these file formats based on open standards, non-proprietary or widely used, documented and supported?
☐ Will the project use existing or third party data as part of the investigation?
☐ Are there any requirements for use of third party data such as licencing conditions?

4. Data Documentation/Metadata

☐ What supporting information/documentation will you create to enhance understanding of the data? e.g. codebooks, data dictionaries, data definitions, publications, websites.
☐ Are you using any metadata standards, controlled vocabularies or ontologies to describe the data?
☐ Are there any Quality Assurance processes that could be applied to your data? e.g. calibration, validation, transcription, peer-review etc.
☐ What processes will be established and followed to document and organise data? i.e. version control, filename conventions, directory structures etc.
5. Data Storage and Security

- How much data are you likely to collect/generate throughout the project?
- Where will the data be stored during the project?
- Will your data be backed up, by whom, how often, where?
- How will access to the data be managed during the project?
- Are there any commercialisation, ethical or confidentiality restrictions relating to accessing or storing the data during the project?
- Is there any non-digital data or outputs that the project will generate? Where will these outputs be stored?

6. Ethics, Copyright and IP

- Does/will the data contain sensitive, confidential or personal information? If yes, what methods will be used to protect the data e.g. encryption, password restrictions etc.
- Are there ethical/privacy considerations surrounding the ability to share/publish the research data outside the immediate research team?
- If intending to share any part of the data, do your participant consent forms include information about intentions for sharing, retention of data and steps taken to protect participants privacy and confidentiality?
- What steps will be taken to protect privacy and confidentiality? e.g. de-identification or anonymising data.
- Has an agreement about the ownership of research data and primary materials been reached between partner institutions?
- Are there likely to be any copyright restrictions that will apply to the data?
- Will the data be collected in or transported to another country or area outside of Australia? A
- Are there any legislative requirements to meet?

7. Access, Sharing and Reuse of Data

- Will part/all of the data be retained on completion of the project? Where will this data be stored?
- Where will non-digital data be stored post project?
- How will access to the data be managed post project?
- Do you plan to share some/part of the data post project? Will the data be deposited with an archive or repository or published on the web?
- Are there any restrictions placed on sharing/reuse of some/all of the data?
- Can these be managed by setting mediated access to the data? e.g. access to the data must be negotiated via Chief Investigator.
- What supporting information to assist with interpretation of the data will be made available? How will the information be made available?
- How will you ensure that identified processes or steps taken to protect privacy and confidentiality will be achieved prior to completion of the project and sharing of the relevant data.
- When will the data be shared post project? e.g. immediately, 3 months, 6 months, 1 year.
- Is there likely to be any costs associated with making the data available for sharing or re-use?

8. Data Retention and Disposal

- How long should the data be retained for? i.e. permanently, 5 years, 7 years, 20 years, etc
- If disposing of data, outline how will you handle the disposal of sensitive, confidential data

9. Preservation and Archiving of Data

- Will the final format of the data files be in a sustainable format supporting long term access? i.e. based on open source standards, non-proprietary.
- Will the supporting documentation be stored with the data to enable future interpretation?
- Who is responsible for maintaining the data after the research project is complete? e.g. Chief Investigator, data manager, research assistant.
- Are there likely to be any costs associated with the long term storage of the data?