ANDS Guide

Data management plans

Level: Awareness

Last updated: October 2017

Web link: www.ands.org.au/guides/data-management-plans

The following information is likely to be of interest to researchers and research administrators who are charged with preparing a data management plan for a research project or an institution.

What are data management plans?

A data management plan is a living document that describes:

- What data will be created
- What policies will apply to the data
- Who will own and have access to the data
- What data management practices will be used
- What facilities and equipment will be required
- Who will be responsible for each of these activities.

Many Australian Universities have Data Management Plan tools available for use by researchers needing to create a Data Management Plan (DMP) at the start of a research project.

Discipline specific DMPs

- Managing research data in Archaeology from the Archaeology Data Service in the UK.
- MyScience and Data Management Plans from the Australian Antarctic Data Center.
- Example DMP for a Biology study from Curtin University
- Example UK DMPs - from many disciplines
- NEH DMPs – example from the Humanities

Why do I need a data management plan?

The carrot: improvements to efficiency, protection, quality and exposure.

Data management in some form is an important part of research. In the past data management was typically done at the last minute and using the first method that came to mind. This approach is usually time-consuming and error-prone. Taking time at the start of a research project to put in place robust, easy-to-use data management procedures will usually pay off several times over in the later stages of the project. Inadequate data management can also lead to catastrophes like the loss of data or the violation of people's privacy.

The stick: basic data management is required by the Australian Code for the Responsible Conduct of Research. Compliance with the Code is already a requirement for the Australian Research Council (ARC) and the National Health and Medical Research Council (NHMRC) funding and is likely to be mandated by other funding bodies, Government and institutions in the near future.

What does a data management plan need to cover?

The following list of topics can be treated as a check-list:

Backups	This is probably the single most important item on this list. You must have a credible backup strategy of regular backups, and of course you must then follow it. Consider including an off-site backup so that your data will not be lost if a local catastrophe occurs. Consider an automated backup process.
Survey of existing data	What existing data will need to be managed?
Data to be created	What data will your project create?
Data owners & stakeholders	Who will own the data created, and who would be interested in it?
File formats	What file formats will you use for your data?
Metadata	What metadata will you keep? What format or standard will you follow?
Access and security	Who will have access to your data? If the data is sensitive, how will you protect it from unauthorised access?
Data organisation	How will you name your data files? How will you organise your data into folders? How will you manage transfers and synchronisation of data between different machines? How will you manage collaborative writing with your colleagues?

	How will you keep track of the different versions of your data files and documents?
Storage	Where will your data be stored? Who will pay for the hardware? Who will manage it?
Bibliography management	What bibliography management tools will you use? How will you share references with the other members of your group?
Data sharing, publishing and archiving	What data will you share with others? What license will you apply?
Destruction	What data will you destroy? When? How?
Responsibilities	Who will be responsible for each of the items in this plan?
Budget	What will this plan cost? Possible costs include hardware for backups, research assistant time for data curation, metadata creation, archiving etc.
Anything else	Don't restrict yourself to the items above. Stop and think. What is missing from this list? If you think of something, please let us know so that we can update this information.

Next generation approaches

While there is still an open question about the efficacy of data management plans (DMPs), work continues on multiple fronts to improve them and employ them in a way that truly supports the research enterprise. Institutions are beginning to move from those long early versions of DMPs to a next generation of DMP tools and approaches that consider whether DMPs can or should be:

- public not private documents
- machine readable as well as human readable
- flexible living documents that can be changed through the course of a research project
- measurable (i.e. did researcher X do what they said they would do in their DMP?)
- connected to at least one other system rather than standalone forms.

A report on Machine-actionable Data Management Plans (maDMPs) by Stephanie Simms, Sarah Jones, Daniel Mietchen and Tomasz Miksa reflects collective thinking on this next generation approach.

Other issues to consider

- Funding bodies and Governments are moving rapidly to require sound data management. You have a responsibility to make yourself aware of any relevant codes and to comply with them.

- Failure to comply with requirements from funding bodies like the ARC or NHMRC may jeopardise future research funding.
- Failure to comply with legal requirements, such as those that safeguard the privacy of participants in medical research, may lead to prosecution.
- Changes to ARC funding rules can be seen in the Funding Rules for schemes under the Discovery Program.
- Further guidance is also available through the Instructions to Applicants and Frequently Asked Questions for each scheme.

Different disciplines have different conventions. In order to facilitate cooperation, you should make sure that your data management is compatible with the prevailing standards in your discipline. This mostly applies to file formats and metadata standards.

Feedback?

We welcome your feedback on this guide. Please email contact@ands.org.au with any comments or questions.

About ANDS

The Australian National Data Service (ANDS) makes Australia's research data assets more valuable for researchers, research institutions and the nation.

ANDS is a partnership led by Monash University in collaboration with the Australian National University (ANU) and the Commonwealth Scientific and Industrial Research Organisation (CSIRO). It is funded by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS).

NCRIS
National Research Infrastructure for Australia
An Australian Government Initiative