ARC applications: some tips for writing about data management

Management of Data is a new addition to ARC applications, and is important for most types of research. Consider, for example, that appropriate data storage management is a part of the Australian Code for the Responsible Conduct of Research, adherence to which is a requirement of receiving ARC and NHMRC funding. Sound data management supports the integrity, credibility and repeatability of research, avoids duplication or loss, and equips future data sharing and accelerated research. Data management will only continue to grow in importance in future, so a genuine plan is a necessity.

Presently, under the Management of Data section, ARC applications require you to:

Outline plans for the management of data produced as a result of the proposed research, including but not limited to storage, access and re-use arrangements.

Suggested length
Approximately a quarter to a third of a page should suffice in most cases.

General principles
Writing this section of an application should be straightforward if you have already worked on your Data Management Plan. Macquarie has created a template to help you with this.¹

Note that data can include physical specimens and a variety of other sources, not merely numerical or digital datasets.

See Macquarie’s Research Data Management Toolkit for a further introduction to data management.

Resources to consider in writing about your management of data
Make sure that your description of data management for the project covers at least the minimum required by the ARC; i.e. storage, access and re-use.

Data storage
When deciding how and where to store your data, consider questions such as:

- what is the project’s anticipated data volume over the period of the project? (e.g. 1Gb per week for 3 months)
- who requires access to the data, when and via what means? (e.g. file sharing via your desktop, connectivity via your web browser)
- what data formats will be used and can these be readily migrated to newer/other formats? (e.g. Excel, CSV, text, NetCDF files,.doc,.jpg,.plt)
- how will your chosen storage facility meet privacy and security requirements? (e.g. are you sharing the data with Macquarie and/or external colleagues? Does the data contain information on patients or subjects in health or medical projects?)
- is the storage medium suitable for long-term storage or will you need to migrate it for archiving purposes?
- will you deal differently with working data and final datasets?
- where will data be backed-up? Multiple copies in different locations and frequency of back-ups?

¹ For data-intensive projects, note that Intersect provides dedicated expertise as part of Macquarie University’s membership. Please contact Stuart Allen for specialist advice on the preparation of data management plans and/or the inclusion of Intersect eResearch services in your grant.
Your description of your management of data should address these and any other relevant points.

For storing digital data, there are various internal and external stores that you could potentially use, ranging from the Faculty of Science’s network drive or Macquarie’s Truth system to external systems such as NeCTAR’s Research Cloud, Cloudstor+ or the Research Data Storage Infrastructure (RDSI). These and other options are described, together with details of the growing sharing and data citation culture, as part of the Toolkit.

For physical samples safe storage will depend on their nature, but options may include specialised storage in one of Macquarie’s museums, storage in a biobank and so on. Local discussions would need to be held within your department or faculty as a starting point.

Data access

The key considerations are who will be given access to the data and the means by which they will access it. The storage resources described in the Toolkit allow different access: some only allow access by Macquarie staff; other stores, such as the Research Data Storage Infrastructure (RDSI), are designed expressly to allow access to data for re-use from anywhere, provided the intended use is appropriate. For that reason, inclusion of data within the RDSI is by nomination and considers potential value for future research; see Intersect’s RDSI website for more details.

Data re-use and sharing

A key reason to safely and accessibly store data is to allow suitable re-use of data, especially data collected during the course of publicly-funded research. Increasing data re-use is important for the ARC and NHMRC, so the more accessible data can be, within the constraints listed below, the better.

When deciding on how the data could be re-used, please consider:

- The likely level of interest in data from the wider research community (or what is termed being “valuable for future research” by RDSI), which should help you choose the most appropriate data store too.
- Whether privacy, security or other ethical considerations and/or protection of intellectual property (IP) would limit the ability to share the data for re-use. In the case where considerable intellectual effort has been invested in creating and/or analysing the data, so that it represents valuable IP for Macquarie, perhaps the data may be able to be shared after a suitable period has elapsed.
- How you will ensure during the course of the project that all the necessary metadata (e.g. details of measurements, any pre-processing done etc.) are captured and made available to allow suitable re-use.

One final point that you may want to consider mentioning is any strategy for promotion of the data, once available, to ensure uptake and re-use. Research Data Australia is an online directory for that purpose. Macquarie University Library can help you with promoting your research data collections by describing them and ensuring the description is included in ResearchOnline and in Research Data Australia. Please contact the Library at researchdata@mq.edu.au for more information.