

A research data management plan

Dasapta Erwin Irawan & R. Willem Vervoort

5 January 2018

Research data management plans

What we have learned so far:

- ▶ Increasing data and methods publishing can really move science forward
 - ▶ It will help you as a researcher (you can find data and use methods easier)
 - ▶ It will help the whole science community
 - ▶ It might resolve some of the issues around peer review and costs of publishing
- ▶ Good meta data are the key for data re-use and discovery
- ▶ Writing meta data is quite a job, especially after the fact
- ▶ We need to develop strategies and methods to tackle this when we start!
- ▶ **Research data management plans**

What are the main components of a research data management plan?

Unsurprisingly: **Metadata**

- ▶ Who are involved (who owns the data)
- ▶ How you collect the data
- ▶ How you consistently describe the data content
- ▶ Where you store the data and how you store the data (format and file structure)
- ▶ How you plan to curate the data (where is the long-term home of the data)

Schematic of a research data management plan

The University of Sydney

► The University of Sydney Research Data Management Plan

The screenshot shows the University of Sydney Library website. The top navigation bar includes links for University, Library, Current students, and Staff intranet. Below this is a secondary navigation bar with Home, Study, Research (highlighted), Collections, Our libraries, Help, and Contact us. A third bar contains links for Book a study space, Book equipment, MyLoans login, Ask Us Now, Meet with a librarian, and Feedback. A large red search bar with the text 'Search Library site' and a magnifying glass icon is prominent. Below the search bar, the breadcrumb trail reads 'Library / Research / Research data management /'. The main content area is titled 'Research Data Management Plans' and describes it as 'Essential for all research projects'. It explains that creating an RDMP helps meet funder requirements and University policy, and guides users on how to actively think about research data management. A red box on the right side of the page contains the text 'RDMP tool' and 'Create your RDMP' with a right-pointing arrow. A vertical 'Share' link is visible on the far right.

- The RDMP at USYD is currently being revised to be more focussed on the actual research data management
- A living document, a “readme.txt” for the project.

Examples of research data management plans

- ▶ Best practice dataone.org: Go through website
- ▶ Simple Research Data Management Plan: [Tier Protocol](#)

Vocabulary and data dictionary (again)

- ▶ Choosing the right keywords for your project is crucial
- ▶ For the research data management plan: this is the source
 - ▶ feeds into all the further “downstream” metadata
- ▶ The keywords should originate from a **vocabulary** or an **ontology**
 - ▶ Ontology: set of controlled terms for keywords with a hierarchy: example [FOR codes](#)
 - ▶ A good ontology would have related terms, Wikipedia is an example of a system that uses an ontology
 - ▶ A vocabulary is a more simple list of keywords, for example, most journals require you to choose from specific keywords when you submit a paper

Vocabulary and data dictionary (again)

- ▶ Data dictionary is simpler:
 - ▶ describes columns in a data sheet
 - ▶ describes layout of code structure
 - ▶ describes files and folders
- ▶ example data dictionary: [Readme file in example project](#)