Sustainable Services for Managing Higher Degree Research Data

Arif Shaon, Eamon Smallwood, Maude Frances
UNSW Library, UNSW Australia
Shane Cox
USNW IT, UNSW Australia
Lucian Hiss
Graduate Research School, UNSW Australia

Effective management of research data is increasingly being recognised as an important aspect of successful research projects. This is driven by the growing need to publish good quality research data, for which efficient data management is essential. Publishers, funding agencies and researchers are paying more attention to the chain of intermediate data results and their associated metadata (including provenance), in order to determine the authenticity and quality of research data [1].Although researchers in Australia are not generally expected to provide planning documentation, major funding bodies require compliance with the Australian Code for the Responsible Conduct of Research [2], which stipulates responsibilities of institutions and researchers in relation to record-keeping, storage, preservation, ownership, access, confidentiality, retention and disposal of research data. Applicants for funding from the Australian Research Council are asked to "[o]utline plans for the management of data produced as a result of the proposed research, including … storage, access and re-use arrangements" [3].

As underlined in the Australian Code for the Responsible Conduct of Research, researchers alone cannot be expected to bear the full responsibility for managing their data. The Code is designed to prevent research misconduct by providing a detailed set of best practices to comply with responsible research. However, this is accomplished by describing principles and practices for responsible research, backed up with a framework to manage research misconduct if it does occur. To meet the requirements of the Code, researchers need to be supported by their institutions through an institutional framework, which includes policies, infrastructure and support services.

Developing an institutional infrastructure for research data management requires institution-wide collaborative efforts with contributions from the key stakeholders. An important requirement for such an institutional infrastructure is to provide support for Higher Degree Research (HDR) candidates who usually represent a significant percentage of the research community an academic institute.

At UNSW Australia, there are currently over 3000 enrolled HDR candidates that represent a large portion of the university's research community. Unsurprisingly, the research conducted by HDR Candidates forms the foundation of much of the research progress at UNSW. This makes appropriate planning and management an important requirement of UNSW HDR. Additionally, the HDR practices at UNSW have a number of special attributes that place further importance on effective management of HDR data.

- Intellectual Property (IP) from HDR Candidate research is retained by the candidate, unlike research conducted by University employee,
- Varied privacy requirements are associated with data sharing between students and supervisors.
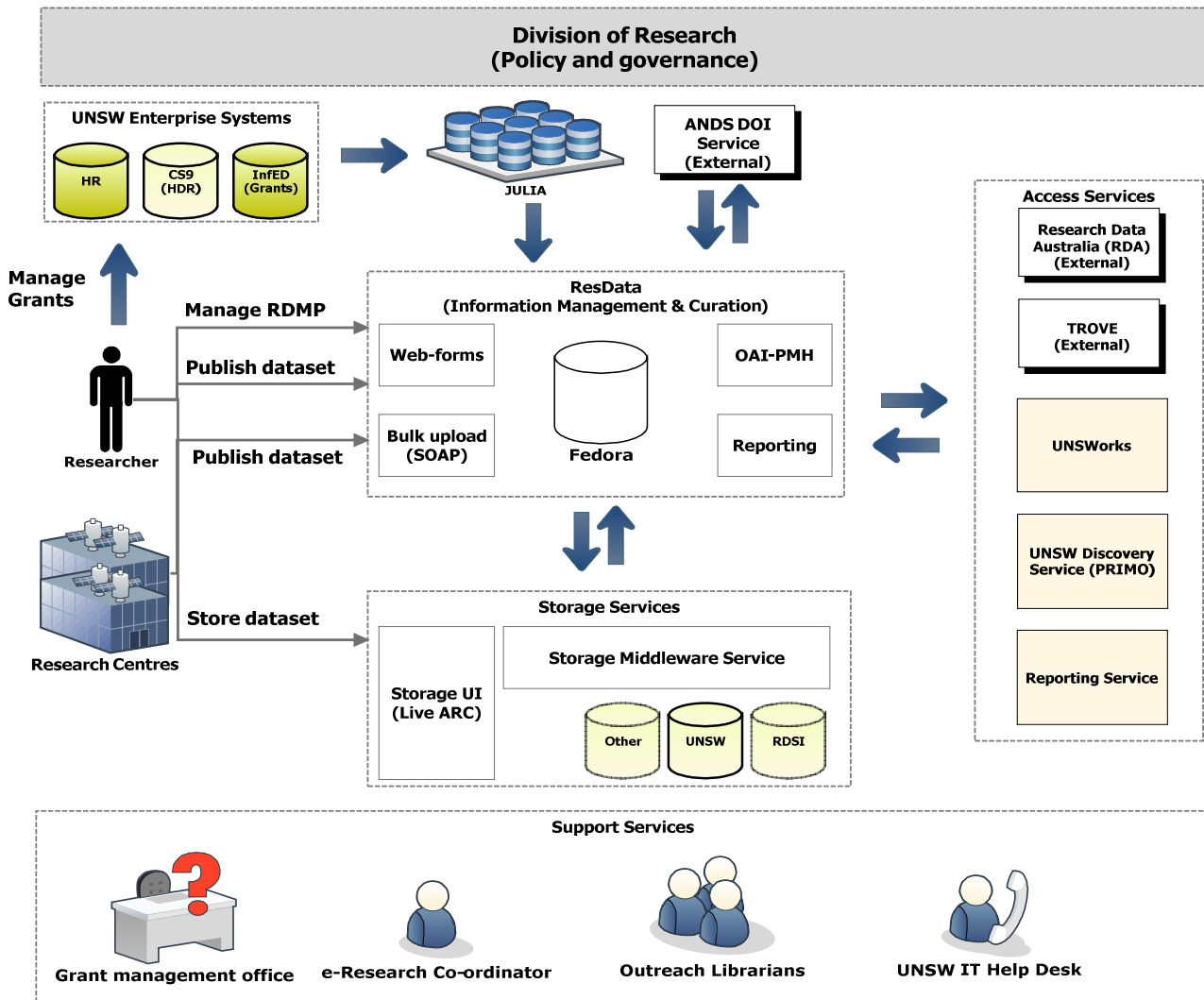
Fig1. UNSW Research Data Management Services

In order to support research data management at UNSW Australia, the UNSW Library and UNSW IT, governed by the Division of Research, are developing a sustainable research data management (RDM) service (fig. 1). In this presentation, we highlight how the UNSW RDM service infrastructure supports the data management needs of the university's HDR candidates by establishing and leveraging important relationships between the key stakeholders. In brief, the infrastructure aims to support provision of the following services.

- Long-term data storage with support for different sharing permissions for HDR Candidate research - allowing candidates to store their data privately if they wish.
- A web-based tool to enable creation of research data management plans as a prerequisite for obtaining access to data storage. It is anticipated that by encouraging new researchers through the RDMP to engage with good data management practices will improve overall research outcomes from these researchers.

The requirements of these services were collected through a number of workshops with the UNSW research community that included the HDR candidates, and the UNSW Graduate Research School (GRS). The infrastructure also provides a foundation for establishing future relationships between new and existing stakeholders, thereby developing sustainable services for the wider UNSW research community.

References
[1] A. Shaon, S. Callaghan, B. Lawrence, B. Matthews, A. Woolf, T. Osborn, and C. Harpham, "A linked data approach to publishing complex scientific workflows," Proc. 7th IEEE International Conference on e-Science (eScience2011), Stockholm, Sweden, December 05-08, 2011. - http://doi.ieeecomputersociety.org/10.1109/eScience.2011.49

[2] Australian Code for the Responsible Conduct of Research, Australian Government, National Health and Medical Research Council, Canberra, ACT, 2007.

[3] Discovery Projects: Instructions to applicants for funding commencing in 2015, Australian Government, Australian Research Council, Canberra, ACT, 2014.