Developing sustainable research data management services: an institutional collaboration

Maude Frances, Arif Shaon (UNSW Library)
Shane Cox, Luc Betbeder-Matibet (UNSW IT)
Outline

• UNSW Library and research materials
• Research Data Management service
  – Use cases
• Relationships
  – Research practice and assets
  – Institutional
  – Governance and support
• Future work and conclusions
UNSW Library Repository Services

- Library role in management and curation of research materials
  - Fedora for underlying storage technology
  - Ex Libris Primo for search and discovery service
  - Java EE6-based content management interface
- Web-based repositories for research material
  - Faculty-based (disciplinary) collections
- ResData - research data management service
  - Catalogue of datasets
  - Research data management plan
Research Data Publishing & Management

• Increasing importance of publishing research data as important resources
  – facilitates validation, reproducibility and attribution
  – requires efficient management

• Recommendation/requirements from funding bodies for research data management plans
  – the Australian Code for the Responsible Conduct of Research
  – Requires institutional policies, infrastructure and support services
Research Data Management

Research Data Management (RDM)

- Planning
- Curation
- Publishing

RDM Infrastructure (Services)

- Support
- Storage
- Discovery & Access

Research Policy

UNSW
UNSW RDM Services – an institutional collaboration

Library

Research

UNSW IT
UNSW RDM Services Architecture
Scenario: Researcher

I am a researcher with a current research project for which students are collecting data. I want to create a research data management plan to help them with their data management and also provide data storage for the project.
Scenario: Researcher

Step 1: Create a Research Data Management Plan (RDMP) through the web interface. The researcher will be prompted for information which aligns with requirements of the university and funding bodies.
Scenario: Researcher

Step 2: When submitted, the RDMP will be managed in ResData and will be used as a basis for subsequent data management activities for that project.
Scenario: Researcher

Step 3: Information about existing grants is available in a feed from the UNSW data warehouse (JULIA). The ResData record is prepopulated with these metadata.
Scenario: Researcher

Step 4: A report of the RDMP can be exported through the reporting service.
Scenario: Researcher

**Step 5:** Once a dataset has been collected the researcher can create a dataset description record through the Dataset Metadata Deposit interface or through other services that automate capture from analytical devices.
Scenario: Researcher

Step 6: These dataset records, when sent to ResData, are pre-populated with RDMP metadata for that project.
Scenario: Researcher

Step 7: A DOI can be requested at the time the metadata record is created or at a later date. The DOI, created through the ANDS Cite My Data service and DataCite, allows the dataset to be cited in publications.
**Scenario: Researcher**

**Step 8:** Information in the RDMP record will enable the Data Storage Interface to determine where the data can be located (either locally or externally), who may access it, and the associated retention periods, etc.
ResData Integration with UNSW Enterprise Databases

- JULIA Data Warehouse
- Links to Researchers' Profiles

- .csv (Grant)
- .xml (Party)

- ResData
  - Fedora/MySQL
    - Versioning
    - RDM Plans

- Research Data Storage
  - UNSW RDS
  - External RDS (RDSI, ADA)
ResData Publish Workflow

- ANDS Cite My Data Service
- Data Cite DOI Resolver Service
- Landing Page
- Dataset Record

- Request DOI for record
- Receive DOI
- DOI access resolves to landing page
- Landing page provides access to both the published record and other related resources
ResData Metadata Ontology

- Dataset
  - (1) generatedBy(1)
- Process/Methodology
  - (*) outputOf(1)
- Research Data Management Plan
  - (1) Has (?)
- Party/Contributor
  - (*) Contributors (1)
- Project/Grant
  - (*) Contributors (1)
  - (1) associatedWith (1)
Sustainability

• Two levels of relationships
  • Research projects, outputs (data), and the associated personnel
  • Roles and responsibilities associated with the provision and sustainability of the service components

• Flexibility
  • Future requirements
  • Changing organisational structure and relationships
Conclusions and future work

• Higher degree research candidates
  • RDMP template
  • Automatic allocation of storage
• Research Data Storage Infrastructure (RDSI)

• Data Documentation Initiative metadata (DDI)
• UNSWorks preservation policy and procedures
ResData URL

https://resdata.unsw.edu.au/pages/authenticate.faces
Acknowledgment

The work presented in this paper was supported by the UNSW Faculty IT Investment Plan fund and the Australian National Data Service (ANDS). ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy Program and the Education Investment Fund (EIF) Super Science Initiative.