Platforms for Research Data Management
Lessons Learned

Anthony Beitz
Monash e-Research Centre

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
Look, half the work is done!
All you need to do is fill in the top part so we can legally say the bottom part.

DATA:

CONCLUSION: Eating chocolate will make you look younger and thinner.
Lessons in Research Data Management (RDM) Platforms

1. What is a RDM platform?
2. Selection
3. Development
4. Deployment & Support
Conventional Research Data Management (RDM)
Evolving Research Practises

Conduct Experiment

Collect Data

Analyse

Publish

Conduct Experiment

Collect Data

Analyse

Publish
Research Data Management Lifecycle

Data Management Planning

Conceive

Design

Experiment

Analyse

Collaborate

Institutional Repository

Publish

Community Portal

Expose

Research Data Management Platform
RDM Platform Landscape
About Researchers

• Focused on research outcomes
• Work in an interpretive mode
  – Characteristics include: an iterative process; open-ended; and thrive on ambiguity
• Their requirements and goals may change over time
• Generally more loyal to research community than their institution
• May require an ICT capability for only a short period of time
• Very resourceful and driven
Selecting a RDM Platform

- Fit-for-purpose
- Cultural fit
- Fit with researchers’ tools, environment, and workflows
- No single institutional RDM platform will fit all researchers' needs
- Adopt, customise, or develop??
  - Developing a new product may be expensive, may be costly to support, and may split researchers from their community
  - Institutes can’t afford to develop/customise RDM platforms for each of their research disciplines
  - If a research community already has a RDM platform (or there is an emerging one) and it meets the needs of your researchers in that community, then adopt it
RDM Platforms @ Monash

Adopted

Derived

New
Being Agile

We're going to try something called Agile programming.

That means no more planning and no more documentation. Just start writing code and complaining.

I'm glad it has a name. That was your training.
Objectives when Developing a RDM Platform

- *Fit-for-purpose*
- *Good adoption*
- *Sustainable*
Ensuring Fit-for-Purpose

• Solution must be led by researchers
• Adopt Agile software development methodologies, with
  – at least one researcher as a Product Owner
  – sufficient time at start to acquire decent list of desirable functionality
  – a roadmap, to ease tensions if Scrum being driven by Waterfall from above
• Don't develop an RDM platform to be as discipline-generic as possible, as no one will find it useful
• It is unrealistic to expect any RDM platform to suit every discipline's needs without specific development or customisation
Promoting Good Adoption

• Deliver early and often
• Promote sense of ownership by the researchers
• Support researchers in raising awareness of the RDM platform to their research community
Ensuring Sustainability

- Promote sense of ownership by the researchers
- Support researchers in raising awareness of the RDM platform to their research community
- Software should be easily deployable at other institutions
- Encourage other developers to embrace and extend the solution
Deployment & Support

- Institutional IT groups can assist in the hosting of a RDM platform by providing VMs, a range of popular databases, large amounts of secure and reliable data storage; and assisting in the deployment
- As there may be many hosted RDM platforms at an institution, support may need to be a partnership between the researchers using a platform and the institute’s ICT staff
- For newly developed RDM platforms, establish governance and model for shared maintenance
RDM Platform Landscape @ Monash

• Growing range of Research Data Management (RDM) platforms with good fit-for-purpose, research community support, and high level of adoption

• Two versatile RDM platforms now available, which can be easily customised to particular research needs
  – MyTARDIS
  – Eddy (E-research Distributed Data sYstem)

• General RDM platform available for research disciplines without an RDM platform – Squirrel
Summary

- Research institutions will be hosting a range of RDM platforms
- If a research community already has an RDM platform and it suits their needs, then adopt it.
- If not, then look at customising an existing RDM platform
- Institutes can help in the hosting of RDM platforms by providing VMs, databases, data storage, and installation assistance.
This project is supported by 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
For more information…

Contact:
Anthony Beitz
Manager, Monash e-Research Centre
Ph: +613 9905-8681
Email: Anthony.Beitz@monash.edu
Web: www.monash.edu/eresearch
Twitter, Skype: a_beitz