Known unknowns: Towards a strategy for preserving digital research data

Owen O’Neill
The University of Melbourne Research Data Preservation and Archiving strategy:

• Why do we need a strategy?
• How are we approaching it?
• What have we found so far?
• Where to next?

About me:
Owen O’Neill, Data Archiving Analyst
ITS, Research Services team
University of Melbourne


Image attribution: http://www.flickr.com/photos/mkriegsman/2852232536/
Why do we need a strategy?

Supporting research & researchers

Internal/push factors
- Enable wider dissemination of the University’s scholarship
- Cultivate an environment that supports researchers and research projects
- Meet legal, ethical, moral obligations

External/pull factors
- Funding body and government expectations
- Open data movement
- Community/social expectations
- Technology driven change
- International trends

Image attribution: http://www.flickr.com/photos/x1brett/1530622309/
How are we approaching it?

Broad objectives

• Improved discoverability and access to University research data
• Support cross-disciplinary collaboration
• Meet legal and ethical requirements for data preservation
• Utilise international standards and best practice
• Define and clarify recommended workflows, infrastructure and roles and responsibilities

Desirable outcomes

• Build communities around data
• Support new research and new collaborations
• Facilitate new ways of utilising/working with data
How are we approaching it?

Our recipe

Expert advisory group
- Researchers, library, technical and policy focused staff
- Meeting every 6 weeks for the 18 month project
- Discussing a running list of agreed topics
- Will provide endorsement of the strategy

Project steering group
- Overseeing progress
- High level representation from University Library, Information Technology Services (ITS) and Melbourne Research Office
- Final sign-off of strategy and associated roadmap

Targeted trials
- 3 small trials working with researchers and data collection custodians
- Documenting workflows and good practice, helping participants to improve processes as required
- Three themes: preparing for preservation, managing access, maintenance

Image attribution: http://www.flickr.com/photos/15957761@N05/2230432135/
Our recipe continued…

Stakeholder consultations
- Researchers and data custodians
- Academic Board IT Committee

General investigations
- Aligning with/supporting key University objectives (The University Plan, Research Strategy, Melbourne’s Scholarly Information Future)
- Identifying strategies, policy, guidelines, infrastructure and related projects at other research-focused Universities
- Identifying relevant standards and good practice examples

Image attribution: http://www.flickr.com/photos/15957761@N05/2230432135/
Four key principles

- **Organisation**: Support services better configured towards supporting research
- **Culture**: Data management and preservation is part of what we do
- **Infrastructure**: High quality infrastructure with support for uptake and use
- **Policy**: Requirements are clearly articulated and contextualised
What have we found so far?

There’s a diversity of opinion…

- Data should be stored locally/centrally/externally for long term preservation
- Who’s going to pay?
- Our needs are unique
- Data management plans should be encouraged
- How can I keep my data safe?
- What do I need to do to maintain data?
- How to manage access to data more effectively?
- The University needs to [take more responsibility] / [stop meddling]
- We’re not convinced there’s a problem
- This sounds like an additional layer of bureaucracy
- What skills do my team need?
What have we found so far?

In relation to our 4 key principles…

Culture
• What is the role of training and support?
• How to acknowledge research data management and preservation skills?

Organisation
• How to be smarter in gathering and utilising administrative information about data collections?
• How to provide cohesive support for researchers that is easy to access?
• What would a digital archive look like?
What have we found so far?

In relation to our 4 key principles…

**Infrastructure**
- What is the best mix of storage options (Central, Local, External)?
- What funding and costing models are going to work for storage infrastructure?
- How should we optimise storage infrastructure for preserved data collections as well as active data collections?

**Policy**
- How to provide better guidance on good practice interpretation of policies?
- Is there scope for streamlining related policies to make them more accessible?
Our to do list

• Completion of trials and stakeholder consultations
• Progressively wider input to Strategy draft
  • Key University groups
  • Input from the wider University research community
  • Opportunities for external feedback/input
• Development of a business case to support the strategy

Strategy due to be finalised: July 2013
Questions and comments?

More information

Owen O'Neill
03 9035 4882
ooneill@unimelb.edu.au
https://twitter.com/osoneill