On being a cog rather than inventing the wheel: Edinburgh DataShare as a key service in the University of Edinburgh’s RDM Initiative

Robin Rice, Data Librarian
EDINA and Data Library
OR 2013 Workshop:
IR’s dealing with data
Charlottetown, Canada: 8 July, 2013
The Data Library and Edinburgh DataShare

• **Edinburgh Uni Data Library**: 1983-2013
  – Support for finding, accessing, using, teaching, managing data

• **Edinburgh DataShare**: 2009-2013
  – Built from Jisc-funded project to create exemplars of institutional data repositories: *DISC-UK DataShare*
University of Edinburgh
RDM Policy and Roadmap

• Data Asset Framework project: 2008

• Research Data Management Policy: May, 2011

• RDM Steering Group formed: May, 2012

• RDM Roadmap: November, 2012

• RDM funding: June, 2013
U of Edinburgh’s RDM Policy

- Passed Senate in May, 2011
- Ten brief principles
- Sets out roles and responsibilities
- Aims to enable researchers to achieve good practice
- Wide consultation organised by Library-led committee set up by the Vice Principal
What does the policy look like?
Institutional players

Info. Services Divisions
- Library & Uni Collections
- EDINA & Data Library
- IT Infrastructure
- User Services
- Digital Curation Centre

Rest of Uni
- 3 colleges, 12 schools
- ERI (Research Office)
- Records Mgmt
- ‘Champions’
Data Management Planning

Active Data Infrastructure

Data Stewardship

Data Management Support
The Data Lifecycle

Data Management Planning

Conceive

Design

Experiment

Analyse

Collaborate

Publish

Expose

Research Data Management Platform

National Repository or Institutional Repository or Electronic Journal or Community Portal
Data Management Planning

- Support and services for planning activities that are typically performed before research data is collected or created
  - Responsible: DCC & RDM Action Group
Data Management Planning objectives in the roadmap

1. Tailored DMP assistance for PIs submitting research proposals
   - Analyse plans and collate a set of successful examples
   - Consultancy service for researchers wanting advice

2. Customise DMP Online for optimal UoE use
   - Evaluate DMP Online with Edinburgh researchers
   - Brand / customise DMP Online for Edinburgh
   - Develop custom guidance with Schools
Facilities to store data that are actively used in current research activities, to provide access to that storage, and tools to assist in working with the data.

- Responsible: IT Infrastructure
Data infrastructure requirements

- Globally accessible cross platform file store with sufficient capacity to meet majority needs
- Support mobile devices and collaboration
- Mechanisms for backup and synchronisation
- Ensure integrity and long term retention of ‘golden copy’ of data
Data Stewardship

• Tools and services to aid in the description, deposit, and on-going management of completed research data outputs
  – Responsible: Data Library and Library & University Collections
Data stewardship services

- Data (dark) archive service - vault
- Data asset register
- Data repository (enhanced)
- PURE Current Research Information System (CRIS) integrated with other systems
The data repository as a key component of policy implementation

“9. Research data of future historical interest, and all research data that represent records of the University, including data that substantiate research findings, will be offered and assessed for deposit and retention in an appropriate national or international data service or domain repository, or a University repository.”

RDM Steering Group assigned a number of pilot deposits to test fitness for purpose.
Promoting Edinburgh DataShare

Deposit Your Data

• Increase impact of your research.
• Make your data assets discoverable and citable.
• Minimise the risk of data loss for your own future use.
• Meet funding body requirements & comply with University policy.
• Ensure your research data is preserved for posterity.

http://datashare.is.ed.ac.uk
Challenges from pilots

- Linguistics and English Language. School of Philosophy, Psychology, and Language Sciences. A Collection of Dinka Songs: [http://hdl.handle.net/10283/155](http://hdl.handle.net/10283/155) Also digitised audio files.
- Roslin Institute – lots of ‘omics data
- Informatics (2) – big data, software, licences
- College of Art – high resolution output
- Clinical Psychology – postgraduate work and “sensitive” data
Data Management Support

- General consultancy and support service throughout the research process
  - Responsible: Data Library and User Services
Supporting and training researchers

- Online guidance for academic staff
- Embedding RDM training into postgraduate programmes (MANTRA)
- Training librarians & IT staff (DIY Training Kit)
- Awareness-raising and service roll-out across schools and research units
- Data management consultancy; grant-costed effort such as in-depth planning or metadata support (AQMeN research project)
Online guidance for staff

Guidance and support for managing, sharing and preserving research data.

- **Why manage research data?**
  Benefits of managing your data.
  [Why manage research data?](#)

- **Data management planning**
  Ensuring all aspects of data management are fully perceived at the start of a project.
  [Data management planning](#)

- **Documenting data**
  Guidance for describing your data for future use.
  [Documenting data](#)

- **Data storage and backup**
  Storing and backing up research data to prevent data loss.
  [Data storage and backup](#)

- **Data security, protection & confidentiality**
  Keeping data safe and controlling access to sensitive data.
  [Data security, protection & confidentiality](#)

- **Benefits of sharing data**
  Reasons to share or not to share data.
  [Benefits of sharing data](#)

- **How to share your data**
  [How to share your data](#)

- **Support for data management**
  [Support for data management](#)

[search-support/data-library/online-data-mgmt/storage-backup](#)
It is important to identify and distinguish versions of research data files consistently. This ensures that a clear audit trail exists for tracking the development of a data file and identifying earlier versions when needed. Thus, you will need to establish a method that makes sense to you that will indicate the version of your data files.

- A common form for expressing data file versions is to use ordinal numbers (1, 2, 3, etc.) for major version changes and the decimal for minor changes e.g. v1.1, v2.6
- Beware of using confusing labels: revision, final, final2, definitive_copy as you may find that these accumulate
Wordle from one MANTRA unit
URLs

• Research data guidance
  – http://www.ed.ac.uk/is/data-management

• MANTRA online training
  – http://datalib.edina.ac.uk/mantra/

• Edinburgh DataShare
  – http://datashare.is.ed.ac.uk/

• Research Data MANTRA & DIY Training Kit for Librarians:
  • http://datalib.edina.ac.uk/mantra
  • Edinburgh Data Blog: http://datablog.is.ed.ac.uk
Thank you

THE UNIVERSITY of EDINBURGH

R.Rice@ed.ac.uk