Open data and research data management at the University of Edinburgh: policies and services

Robin Rice, R.Rice@ed.ac.uk, EDINA and Data Library
University of Edinburgh
Overview

• Edinburgh and data-driven science
• Drivers for Edinburgh (institutional) RDM Policy
• Research Data Management (RDM) Policy: key points, roles and responsibilities
• RDM Roadmap: implementing the policy
• RDM training for researchers and support staff
• Challenges in RDM support & open data advocacy
Uni. of Edinburgh and data-driven science

- School of Informatics, 1998-present
  - Doctoral Training Programme in Data Science, 2014
- EPCC, 1990-present
  - Supercomputing infrastructure: HECToR and Archer; RDF
- Information Services (Computing Service, ECDF)
- Data Library, 1984-present
  - EDINA, 1995-present
- National e-Science Centre & eSI, 2001-2011
- Digital Curation Centre, 2004-present
- University RDM Programme, 2012-present
- Edinburgh Data Science initiative, 2014
Int/external drivers for Edinburgh RDM Policy

- OECD Principles and Guidelines for Access to Research Data from Public Funding, 2007 (& 2004)
- Data Audit/Asset Framework (DAF), 2008-9
- Adoption of UK Research Integrity Office’s Code of Practice for Research, 2009-10
- UK funders requiring Data Mgmt & Sharing Plans
- Publishers requiring underlying data be available
- ‘Climategate’ E. Anglia University email review, 2009
  (Summer, 2011: EPSRC “expectations” published)
Edinburgh policy ‘worldle’
UoE RDM Policy: key points for open data

- “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.”

- “The University will provide mechanisms and services for storage, backup, registration, deposit and retention of research data assets in support of current and future access, during and after completion of research projects.”

- “Any data which is retained elsewhere, for example in an international data service or domain repository should be registered with the University.”
Policy: Roles, rights, responsibilities

- Who will support your researcher's planning?
- Who has responsibility during the research project? Who has archival responsibility?
- Who has rights in the data? (What about collaborative research, what about subjects of research?)
- What responsibilities fall to the individual researcher, PI, department, college, university?
- Are students considered in the policy?
Activities involved in RDM

- Data Management Planning
- Creating data
- Documenting data
- Accessing / using data
- Storage and backup
- Sharing data
- Preserving data

A Data Lifecycle

Data Management Planning
Active Data Infrastructure
Data Stewardship

Data Management Support

Involvement from across University and Information Services
Support for DMPs

IN VOLVES:
Research office
Research administrators
RDM Coordinator
DCC (DMPOonline tool)
Central IT consultants
Local IT/web support
Data Library
Academic Service
Librarians

before research data are collected or created

Principal Investigator
Data Infrastructure

INVOLVES:
- Central IT
- Local IT & data mgrs
- Info Services Helpdesk
- IT Consultants
- RDM & IT Committees
  (policy oversight)

- Half TB storage/researcher
- Networked, secure backup
- ‘Dropbox-like’ service
- Long-term, private vault (proposed)
Data Stewardship

INVOLVES:
(Data) librarians
Repository administrators
Archivists
Records managers
Local data managers
Disciplinary repositories
Standards bodies

• Ongoing curation of completed research data outputs
• Access & preservation
• Metadata records, DOIs
DataShare Repository

What is Edinburgh DataShare?

Edinburgh DataShare is an online digital repository of multi-disciplinary research datasets produced at the University of Edinburgh, hosted by the Data Library in Information Services.

Edinburgh University researchers who have produced research data associated with an existing or forthcoming publication, or which has potential use for other researchers, are invited to upload their dataset for sharing and safekeeping. A persistent identifier and suggested citation will be provided.

Deposit Your Data

- How to deposit

Research Communities in Edinburgh DataShare

Select a community to browse its collections.

- Business School
- Edinburgh College of Art
- Edinburgh Law School
- Information Services (IS)
- Moray House School of Education
- Royal (Dick) School of Veterinary Studies
- School of Biological Sciences
- School of Biomedical Sciences
- School of Chemistry
- School of Clinical Sciences

Spotlight

New to Edinburgh DataShare - The reduced genome of Ostreococcus tauri: core eukaryotic signalling components in a tractable model species: School of Biological Sciences.

Information for Depositors

- About Edinburgh DataShare
- Checklist for deposit
- Benefits of deposit
- Service background
- Our definitions
- Service policies

Latest Items

- OttoGwadoAyoker_songJalabias (13 Oct 2014)
- OttoGwadoAyoker_songFuneralSongForKingKur3 (08 Oct 2014)
- TupasLaaWoi_WarlnMalakal (29 Sep 2014)
DataShare progress

Number of Items (datasets)

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Latest month (September, 2014)</th>
<th>Of which are new depositors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>162</td>
<td>15</td>
<td>3</td>
</tr>
</tbody>
</table>

Total size of deposited items: 111 GB

Total sessions (end-users and depositors)

August, 2010 to 18 October, 2014 totalling 16,961 with 9,806 unique visitors and 86,698 page views in Google Analytics (chart shows sessions per week).
Data Support

Data Management Support

- General consultancy and support service throughout the research process
- Training & advocacy
- Specialised support (data collection, analysis, visualisation)

IN Volves:
Academic Service Librarians
Data Librarians
IT Consultants
Local IT & Data Managers
PhD Supervisors & trainers
Research Data MANTRA

• Openly licensed online learning self-paced course in RDM for postgrads and early career researchers
• Embedded in at least 3 graduate schools at UoE
• Video stories from researchers in variety of settings
• Interactive quizzes make it fun
• Data handling exercises in four software analysis packages
• “Continuous improvement”: Four releases in 4 years

http://datalib.edina.ac.uk/mantra
Academic Service Librarian Training

- Data Library & User Services Division piloted training with 4 liaison librarians – then all (12 more)
- Aimed to reuse MANTRA in a blended learning approach for academic liaison librarians
- Materials gathered together in the DIY Training Kit for Librarians on the MANTRA website

http://www.mcphee.com/laf/
Benchmark: DCC 2014 UK survey - RDM services in place

n=87 respondents at 61 institutions incl. 24 Russell Group.
Angus Whyte, DCC, http://www.dcc.ac.uk/blog/rdm-strategy-action-glass-half-full
Challenges in RDM support & open data advocacy

- Reaching critical mass of students and academics at point of need (timing, priorities, scheduling)
- ‘Turnaround time’ in writing data management plans
- Identifying ‘RDM’ requests at helpdesk; proper routing to expert support
- Costing intensive data services & in-depth support in grants
- Getting balance right for private and open data
  - Incentivising sharing
- Working ‘upstream’ in research process is hard
## DCC 2014 UK survey results

<table>
<thead>
<tr>
<th>Obstacles to RDM provision (Table 8)</th>
<th>% (n=87)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of appropriate staff resources and infrastructure</td>
<td>71%</td>
</tr>
<tr>
<td>Availability of funding</td>
<td>64%</td>
</tr>
<tr>
<td>Low priority for researchers</td>
<td>56%</td>
</tr>
<tr>
<td>Lack of relevant accepted standards</td>
<td>38%</td>
</tr>
<tr>
<td>Lack of knowledge of appropriate solutions</td>
<td>36%</td>
</tr>
<tr>
<td>Lack of appropriate skills and expertise to implement solutions</td>
<td>34%</td>
</tr>
<tr>
<td>Low priority for management</td>
<td>23%</td>
</tr>
</tbody>
</table>

- Pre-release Briefing 1, 23 April 2014. Angus Whyte, Diana Sisu, DCC.