Research data management for you

Andrew Cox, Lecturer, Information School, University of Sheffield & Barbara Sen, Lecturer, Information School, University of Sheffield
An Open Educational Resource of materials for taught and CPD learning about Research Data Management tailored for Information Professionals

Umbrella conference
July 2013

Funded by

JISC
WHAT IS RESEARCH DATA MANAGEMENT?
NAME SOME TYPES OF RESEARCH DATA!
A list we came up with earlier...

- Weather measurements
- Photographs
- Results from experiments
- Government records
- GIS data
- Simulation data
- Log data
- Field notes
- Software

- Images (e.g. brain scans)
- Quantitative data (e.g. household survey data)
- Historical documents
- Moving images
- Physical objects: such as bones or blood samples
- Digitised photos / born digital photos
- Social media data: tweets
- Metadata
What is data?

• Some researchers use other terms, eg “sources”
• Complex: data can be produced from other data
• Massive
• Fragile
Complexity of information practices

- Information flow maps for life science research (RIN, 2009) e.g. in neuroscience illustrate
  - Multiple data sources, of different types
    - Visual images, quantitative data, secondary data
  - Storage devices
  - Multiple analytic tools
    - Some requiring grid power
  - Supporting complex scholarly communication
Context

• “Data deluge”
  – eScience, cyberscholarship, e-research
  – Collaborations
  – Multiple forms of complex data
  – Huge cost of research

• National policy
  – Data Protection Act, Freedom Of Information Act compliance

• Funders’ mandates
RDM: definition

• “Research data management concerns the organisation of data, from its entry to the research cycle through to the dissemination and archiving of valuable results.” (Whyte & Tedds, 2011)
Drivers for RDM

- Open access
- Compliance
- Preservation and curation
- Data storage and security
- Good research practice
Force field analysis of RDM

- Direct benefits to researchers
- Public good obligations
- Compliance
- Data preservation
- Data storage and security
- Other priorities
- Nature of data
- Academic culture & lack of reuse culture
- Lack of RDM knowledge & skills
- Legal, ethical & commercial exceptions

Good Research Data Management practices
Why do librarians have something to contribute?

• Open access leadership role
• Liaison, negotiation skills and contacts with academics
• Knowledge of information management, collection development, metadata skills and practices
• Understanding of research data management as a form of information literacy (IL)
• Established LIS networks for sharing best practice across the profession
• Librarians are good at explaining complex things in accessible ways
IF RDM WERE AN ANIMAL, WHAT WOULD IT BE?
What some librarians drew when asked “if RDM were an animal what would it be”?

• A rather malicious looking spider in a web
• A girly octopus
• A dragon: (“a mystery animal I cannot defeat”)
• A dung beetle
• Ants that cling together to form a living raft to save themselves in a flood
• A creature just coming into being
What librarians need...

• Confidence raising... demystification of a complex social world
• Increased knowledge and competencies
• A change of identity – ability to take risks, operate in undefined contexts
• Prompts to get started with RDM, rather than waiting till policy or infrastructure is clear
The liaison role

- Collection development
- Information literacy training
- Enquiry handling
- Marketing
- Committee work
- Informal networking
- Strategic influence
- Management roles

- Facilitator
- Collector
- Liaison
- Collaborator
- Service provider
- Adviser
- Expert
- Gatekeeper
- Change agent
- Leader
- Maven
- Teacher
- Helper
- Counsellor
Library roles to support research

• Offering advice on funding sources
• Embedded or support roles conducting literature reviews or current awareness alerts for research projects or groups
• Information literacy training
• Supporting REF
• Bibliometrics and measuring impact
• Bibliographic software training
• Advocacy for open access / institutional repository
• Offering data analysis advice
• Offering advice on copyright issues
• Offering advice on archiving of research records (e.g. correspondence)
RDM in UK HEI libraries survey results

• 83 (c 50%) UK HEIs responded to our survey (with Stephen Pinfield) conducted in November 2012

• Low level of service development; high priority for next 3 years
Current RDM services

Few well-developed or extensive services currently being offered by libraries but some basic services

- web portal
- advisory service
- RDM plan advice
- early career awareness
- open access and policy
- data citation
- metadata
- repository
- audit RDM
- awareness of reusable sources
- external data sources
- copyright
- licensing
- data impact
- data analysis
- PGR training
- PGT training
- UG training

Legend:
- Basic service
- Well-developed service
- An extensive service
- No service
- Don’t know
Priorities for the next 3 years

Libraries see RDM services as a priority, with a particular emphasis on advisory, policy support and training services.

- web portal
- advisory service
- RDM plan advice
- early career awareness
- open access and policy
- data citation
- metadata
- repository
- audit RDM
- awareness of reusable sources
- external data sources
- copyright
- licensing
- data impact
- data analysis
- PGR training
- PGT training
- UG training

Learning material produced by RDMRose
http://www.sheffield.ac.uk/is/research/projects/rdmrose
<table>
<thead>
<tr>
<th>Rank by current activity</th>
<th>Rank by top future priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open access and policy</td>
<td>1</td>
</tr>
<tr>
<td>Copyright</td>
<td>2</td>
</tr>
<tr>
<td>Data citation</td>
<td>3</td>
</tr>
<tr>
<td>Awareness of reusable sources</td>
<td>4</td>
</tr>
<tr>
<td>External data sources</td>
<td>5</td>
</tr>
<tr>
<td>Early career awareness</td>
<td>6</td>
</tr>
<tr>
<td>PGR training</td>
<td>7</td>
</tr>
<tr>
<td>Advisory service</td>
<td>8</td>
</tr>
<tr>
<td>Licensing</td>
<td>9</td>
</tr>
<tr>
<td>RDM plan advice</td>
<td>10</td>
</tr>
<tr>
<td>Web portal</td>
<td>11</td>
</tr>
<tr>
<td>Data repository</td>
<td>12</td>
</tr>
<tr>
<td>Metadata</td>
<td>13</td>
</tr>
<tr>
<td>Audit RDM</td>
<td>14</td>
</tr>
<tr>
<td>Data analysis</td>
<td>15</td>
</tr>
<tr>
<td>PGT training</td>
<td>16</td>
</tr>
<tr>
<td>Data impact</td>
<td>17</td>
</tr>
<tr>
<td>UG training</td>
<td>18</td>
</tr>
</tbody>
</table>
Survey results: challenges

• “The skill set of the library workforce, the costs of RDM and the difficult economic climate.”

• “Capacity and workload in a context of shrinking resources”
Do library staff have right skills to play a significant role in RDM?

• “A few library staff have some of the right skills”

• “Librarians have core skills regarding the organisation of information but these need extending to fully encompass the requirements of data management.”

• “They may not be aware that the skills they have are transferable however”
Challenges

• Librarians are already over-taxed with roles; they operate in a highly dynamic context
• They often do not have personal experience of research
• Translating library skills to research data issues
• Will researchers look to libraries for this support? “Being taken seriously”
• Complexity and scale of issues
• Resources, infrastructure, management structures have yet to be created in most institutions
Objectives of the project: RDMRose

- Create and evaluate learning materials about RDM tailored for liaison librarians

- Produce an Open Educational Resource (OER) for self supported Continuing Professional Development and full time Library & Information Management students
The URL...

http://www.sheffield.ac.uk/is/research/projects/rdmrose
Possible Library RDM roles

- Leading on local (institutional) data policy
- Developing researcher data awareness
- Teaching data literacy to postgraduate students
- Bringing data into undergraduate research-based learning, promoting data information literacy
- Providing researcher data advice, e.g. on writing Data Management plans or advice on RDM within a project
- Promoting data reuse by making known what is available
- Explaining the impact of sharing data, and how to cite data
- Advice on copyright and licensing issues
- Signposting who in the institution should be consulted in relation to a particular question
- Auditing to identify data sets for archiving or RDM needs
- Developing and managing access to data collections
- Developing local data curation capacity – building a data repository
Areas where libraries can contribute

- Policy
- Teaching appropriate literacies to PIs and early career researchers, PGR, taught students
- Advisory services on RDM; web sites
  - Awareness of data for reuse; data citation practices; copyright and licensing of data
- Signposting
- Auditing/asset review of data sets researchers have
- Data curation capacity, e.g. appraisal and collection management policy, metadata creation/advice

- In collaboration with other professional services such as computer services, research office and archives/records management staff
- In collaboration with researchers and research administrators
- In collaboration with other stakeholders, internal and external
Institutional Stakeholders

Extra-Institutional Stakeholders

The Researcher

Department

Research Project

The Researcher's perspective

Individual professional perspective

Commercial Partners and Customers

Research Office

PVC research

Human resources

Records unit and university archive

Library

Data repository manager

Researchers in other disciplines

Researchers in the discipline

Computing services

Other Researchers in the discipline

Other HEIs

Funding councils

The public and wider Society

Perspectives on RDM
Design approach to learning materials

- Participatory design process
- Process based or emergent curriculum
- Trialled with 40+ library staff at Leeds, Sheffield and York
Philosophy behind the curriculum

• Structured around different potential roles in RDM
• Open ended context demands exploration of issues, with individual professional reflecting on how issues relate to their own role and how the library organisation might change
• More than about lists of competencies/knowledge, also about professional identity
• Need to understand perspective of researchers
• Need to understand perspective of other professional services: especially research office, computing service, archives and records managers
• Not for specialist curators or data analysts
The learning materials

• Desire for practical hands on experience needs to be balanced by a grasp of strategic issues
• Problem Based Learning (PBL)
• Inquiry Based Learning (IBL)
• Eight sessions equivalent to about 4 hours of study each
• Slides
• Readings
• Learning activity ideas
• Audio files of interviews with researchers
• A fictional case study
A sample of the activities

• An introduction to the theory of disciplinarity
• A discussion of DCC curation lifecycle and alternative models
• Carrying out an interview with a researcher about their work
• Design of a guidance web site
• Reading Data Management Plans; reading the local RDM policy
• Writing potential collection policies for an institutional repository
• Analysing recordings of interviews with five researchers
• Discussing research and professional staff stakeholders perspectives through a complex fictional case study
Learning outcomes

• Explain the diverse nature of research across academic disciplines and specialities and discuss different conceptions of research data
• Analyse the context in which research data management has become an issue
• Discuss the role of a range of support services, including libraries, in RDM
• Reflect for themselves as individuals and for information professionals in general on the role and priority of supporting research data management
• Explain and apply the key concepts of research data management and data curation to real world case studies and professional practice
• Understand how to keep knowledge acquired on the module up-to-date

Material can be studied systematically or followed by theme
How can you use the learning materials?

• Gain a systematic grounding in RDM, through self-directed CPD

• Undertake targeted learning about an RDM topic that is key for your role

• Reuse material or ideas for teaching your library colleagues and others – join an on-going informal RDMRose user group

• Come to Sheffield to take RDM as a module
THE PROJECT
About the partnership

• White Rose libraries:
  – Leeds, Sheffield and York,
  – iSchool, University of Sheffield

• Thank you to the library staff who participated in the delivery of the materials f2f at Sheffield, then and Leeds and York

• Advisory Board
  – Neil Beagrie – expert on digital preservation
  – D. Scott Brandt (D2C2) - Associate Dean for Research, Purdue University
  – Sheila Corrall – University of Pittsburgh, formerly Sheffield
  – Martin Lewis – Director of Library Service, University of Sheffield
  – Liz Lyon - Director of UKOLN and Associate Director , DCC
  – Andrew Thompson (DMSPpsych) University of Sheffield
  – Madeleine de Smaelen (3TU.Datacentrum)
Project Management Team

• Led by the Information School, University of Sheffield
  – *Eddy Verbaan and Jen Smith*, Research Associates/Curriculum Developers; *Andrew Cox*, *Barbara Sen*, Lecturers, Information School, University of Sheffield.
  – *Brian Clifford*, Deputy University Librarian (Head of Learning and Research Support), University of Leeds and Project Director, RoaDMaP
  – *Denise Harrison*, Head of the Learning and Research Services Team, Library, University of Sheffield.
  – *Liz Waller*, Deputy Director Information and Head of Information Services, University of York.
Project

• Starting July 2012– ending Summer 2013
• Focus groups (White Rose libraries) will scope the knowledge base, skills gaps, and training requirements
• Development of a core of material built round existing work
• Case study material
• Delivery and evaluation with Sheffield librarians (Autumn Semester 2012-13)
• Iteration with Leeds and York and FT Sheffield masters students (Spring semester 2012-13)
• Dissemination via website, blog and publications and event for LIS educators
Deliverables

• A module within Sheffield’s taught masters courses (but also available for stand alone study)
  – MSc Digital Library Management
  – MA Librarianship
  – Also MSc Information Management, Information Systems

• Curriculum material (lesson plans, case study, assignments etc) for other Information Departments as OER in Jorum

• A self study CPD version, with support forum
Strengths of the Project

• Addresses a real need
• Brings together practitioners, researchers, experts, teaching staff, managers, students
  – Sharing expertise, and promoting communication and understanding
• Draws on expertise across organisations exploiting the strengths within the consortium and beyond
• Participative curriculum development tailored for librarians
• Provides outputs to support best practice, teaching, and research
• Supported mutual learning, and engagement with reflective techniques
• Continuous evaluation built into the project design
• The resources will be made openly accessible
Evaluation and quality control

• Participating learners’ qualitative and quantitative formative and summative evaluation:
  – Is the material suitable in terms of content, level and approach?
• Tutors’ reflections: success of learning, also impact on personal practice
• Interviews with project stakeholders: wider expectations and impacts
• Response of wider LIS community
• Advisers to check quality against pedagogic principles
Timeline

- Current status: Designed and delivered half day sessions @ Sheffield
- Version 1 of CPD material launch January 2013
- Second iteration with Leeds and York, completed March 2013
- Final Version 2 launched 1st May 2013
If RDM were a movie...

Carry on Research Data Management!

data

actually
The URL...

http://www.sheffield.ac.uk/is/research/projects/rdmrose
Sources and further reading

- JISC. URL: http://www.jisc.ac.uk/
- Jorum. URL: http://www.jorum.ac.uk/