Open access to research data

Opportunities, risks and implications

Martin Lewis
Director of Library Services & University Librarian
Open access to research data 2013-03-28
Some preliminary messages

• **RDM**: Research data management
• It’s still *early days* for RDM
• Researchers are the RDM **experts**
Why manage research data?

• Maximising the impact of data-intensive research
• Improving the likelihood of success in future grant proposals for data-intensive research
• Assurance of research integrity
• Enhanced data security and reduced risk of data loss
• Facilitation of data sharing and collaboration
Why manage research data?

- Maximising opportunities for new research based on reuse and recombination of data from multiple sources, including data mining
- The principle of open access to publicly-funded research outputs, recognised by RCUK and OECD
- Compliance with the requirements of research funders.

From: University of Sheffield Research Data Management policy, 2012
The research data challenge

• Research data volumes ("the data deluge")
• Policies and standards underdeveloped
• Lack of ownership
• Heterogeneity of formats, disciplines, cultures
• IP
• Skills, infrastructure, cash
• Local/regional/national/international provision
Science as an Open Enterprise

• Royal Society, 2012
Science as an Open Enterprise

Recommendation 1

Scientists should communicate the data they collect and the models they create, to allow free and open access, and in ways that are intelligible, assessable and usable for other specialists in the same or linked fields wherever they are in the world. Where data justify it, scientists should make them available in an appropriate data repository...
Science as an Open Enterprise

Recommendation 2

Universities and research institutes should play a major role in supporting an open data culture by: recognising data communication by their researchers as an important criterion for career progression and reward; developing a data strategy and their own capacity to curate their own knowledge resources and support the data needs of researchers; having open data as a default position, and only withholding access when it is optimal for realising a return on public investment.
“The Administration is committed to ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community.

Such results include peer-reviewed publications and digital data.”

John P. Holdren, Director, Office of Science & Technology Policy, Executive Office of the President, February 2013
“we are putting openness at the heart of EU research and innovation funding”

Neelie Kroes, Vice-President of the European Commission responsible for the Digital Agenda, at the launch of the Research Data Alliance, last week.
# Curation policies and support services of the main UK research funders

<table>
<thead>
<tr>
<th>Research Funders</th>
<th>Policy coverage</th>
<th>Policy stipulations</th>
<th>Support provided</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Published outputs</td>
<td>Data</td>
<td>Time limits</td>
</tr>
<tr>
<td>Arts and Humanities Research Council</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Biotechnology &amp; Biological Sciences Research Council</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Cancer Research UK</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Engineering and Physical Sciences Research Council</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Economic and Social Research Council</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Medical Research Council</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Natural Environment Research Council</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Science and Technology Facilities Council</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Wellcome Trust</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

**Terminology clarifications**

- **Published outputs**: a policy on published outputs e.g. journal articles and conference papers
- **Data**: a data policy or statement on access to and maintenance of electronic resources
- **Time limits**: set timeframes for making content accessible or preserving research outputs
- **Data plan**: requirement to consider data creation, management or sharing in the application
- **Access / sharing**: promotion of OA journals, deposit in repositories, data sharing or reuse
- **Curation**: stipulations on long-term maintenance and preservation of research outputs
- **Monitoring**: whether compliance is monitored or action taken such as withholding funds
- **Guidance**: provision of FAQs, best practice guides, toolkits, and support staff
- **Repository**: provision of a repository to make published research outputs accessible
- **Data centre**: provision of a data centre to curate unpublished electronic resources or data
- **Costs**: a willingness to meet publication fees and data management / sharing costs

See related guidance at: [www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies](http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies)

Version 2.2, compiled by Sarah Jones, DCC, February 2012
Research data curation spectrum

CERN  CCDC  EBI  NERC  UKDA  AHDS
What does research data management involve?

The DCC Curation Lifecycle Model
Reasons to be cheerful

• The Digital Curation Centre
  • World-leading centre of expertise and knowledge about RDM

• JISC Managing Research Data programme
  • Funding multiple UK RDM projects
  • DMSPpsych, RDMRose at Sheffield
  • Sheffield a partner in RECODE (EU)

• Distributed RDM expertise in UoS eg HRI, GIS, life sciences, iSchool

• University support for RDM developing
University support for RDM

• Research & Innovation Committee RDM project 2011/12
R&I Committee RDM project

- 12 months from September 2011
- Review of research funders’ requirements
- Interviews with over 20 researchers from PI to research support staff in 5 Faculties, and Prof Services staff
- Assessing gap between funders’ requirements and current practice
- Reported to R&I Committee in November 2012
R&I Committee RDM project

“Some researchers felt they had been doing RDM for years, whereas others felt that RDM was a new concept, viewed with some suspicion as an exercise imposed by research funders.”

“Researchers would support RDM initiatives if they were researcher-led & addressed researchers’ problems, did not involve unnecessary bureaucracy, and allowed flexibility.”
R&I Committee RDM project

Areas for action

- Awareness-raising and training: PGRs, ECRs, established researchers
- UoS portal for RDM advice and support
- Support for costing RDM requirements in grant applications
- Creating a network of RDM champions
- Data Management Plans (DMPs) and metadata
- IP: need for additional guidance
- Storage, back-up and security
- Develop data repository infrastructure
University support for RDM

- Research & Innovation Committee RDM project 2011/12
- RDM policy approved 2012, linked to new GRIP policy
- University RDM Coordinator to be appointed from June 2013 for 2 years to take forward project actions
- Professional services working together
Research data management

- Why manage research data?
- Major funders' research data management policies
- Develop a data management and sharing plan
- MANTRA: online research data management training
Final thoughts

• Not all research data needs to be shared
• Before we can open research data for sharing, we have to manage and curate it effectively
• What do **you** think the University’s RDM priorities should be?