

Perspectives on Research Data Management - 24th May 2012, Ron Cooke Hub, University of York

Summary of Breakout 1: Benefits, Challenges, We want to know more about..

In groups of seven or eight, participants were asked to brainstorm what they saw as the benefits of successful research data management, the challenges research data management might pose and whether there were areas they were keen to learn more about. Each group had at least one representative from Leeds, Sheffield and York and the groups were structured so they represented a mixture of different job roles – library, IT and research support. We allowed 35 minutes for the exercise, including round table introductions, which was about right.

This was a useful breakout which the groups found relatively straightforward to engage with. It was a good exercise to offer relatively early in the day and allowed participants freedom to start to exchange and explore ideas. It became clear that some of the potential benefits were also challenges – overall the challenges comments section was the most populated!

1. Benefits

Some benefits were recorded by most or all of the groups:

- Data reuse
- Increased visibility (of data, institutional, researchers)
- Data impact and citations
- Ethics - DMP is part of good research practice
- Deriving maximum value from publicly funded research
- Providing access for multiple stakeholders: the researcher, the university, other researchers, businesses, the public

Several potential benefits for the **research process** were identified:

- Access to data sets for research: to inspire new ideas, identify collaborators
- Avoiding duplication of effort
- Good DMP improving the quality of grant applications and chances of funding
- A robust DMP infrastructure potentially frees researchers' time to do more research
- Encourages reciprocal data sharing

RoaDMaP: Email: roadmap@leeds.ac.uk

Web site: <http://library.leeds.ac.uk/roadmap-project>

Blog: <http://blog.library.leeds.ac.uk/blog/roadmap>

- Opportunities to integrate data with other research outputs e.g. publications
- DMP provides a framework which encourages researchers to engage with good practice e.g. collection, storage, data analysis and options for sharing
- Valuable data sets preserved for future analysis
- Datasets as evidence to use in bids and in REF submissions
- Encourages continuity within the research team (as data is reused)
- Improves scholarly record (for validation)
- RSS feeds based on interest would be possible if data catalogues are set up effectively

Technical benefits identified included

- Managed backups and data storage; best use of storage
- Data security

Several groups recorded potential **efficiency gains**:

- Opportunities for shared services, particular large scale
- Potential for pooling RDM expertise
- Time savings - finding data more efficiently

The benefit of having high level **RDM champions** was also noted.

Various **reputational benefits** were noted, including avoiding bad press as a result of poor data practices.

RoadMaP: Email: roadmap@leeds.ac.uk
 Web site: <http://library.leeds.ac.uk/roadmap-project>
 Blog: <http://blog.library.leeds.ac.uk/blog/roadmap>



2. Challenges

Themes picked up by all the groups were:

- Cultural change (researcher, institutional, support services) – supervisors and, through them, PhD students were seen as important agents of change (or stasis). The slow pace of cultural change was seen as a challenge for changing RDM practice in the short term but was seen as a potential longer term benefit
- Engaging researchers, highlighting meaningful benefits to their work, highlighting data reuse possibilities
- Limited time and resources
- Understanding roles within the research data management lifecycle e.g. who leads on preservation?
- Ensuring staff have relevant data management training and skills

Meeting **research funder requirements** was mentioned several times:

- One group noted this is particularly tricky as there are mixed messages and differences in compliance requirements
- Funder should drive good practice e.g. promotion of open standards.
- Who monitors compliance?

Institutional issues were seen as a potential barrier/challenge by some groups, specifically:

- Institutional management structures may not foster cross team communication and collaboration and decision making may not involve the relevant stakeholder groups
- Challenge to get the relevant stakeholders together
- Striking the right balance between top down and bottom up service development
- Competitive drivers may be difficult to balance with openness
- Setting up a clear referral network – knowing where our expertise lies and communicating this
- Risk of too much 'stick' to enforce research data management (compliance with funder and institutional policies)
- Identifying a sustainable income stream to support RDM processes
- Approach must fit with academic freedom

Providing a suitable **technical infrastructure** was also mentioned by most groups including:

- Storage space - physical and virtual
- Long term storage – including funding post-project
- Variety of file formats across disciplines

RoaDMaP: Email: roadmap@leeds.ac.uk

Web site: <http://library.leeds.ac.uk/roadmap-project>

Blog: <http://blog.library.leeds.ac.uk/blog/roadmap>

JISC



D|C|C

White Rose
Research Online

- Data security in the Cloud
- Dealing with obsolescence: storage media, software, formats
- Knowing what data is already out there and what is deposited in local or disciplinary repositories
- Creating an appropriate institutional repository infrastructure; understanding what support systems are already in place at the institution
- Tracking changes to data and version control
- Staff skills

Ethical issues included:

- IPR and copyright e.g. understanding ownership in multi-partner projects
- Security of personal data
- Reassuring participants – making sure data management issues don't put them off

Dealing with data

- Discussions also noted the importance of a shared and commonly understood vocabulary for research data management.
- Agreeing good practice on file formats and metadata was picked up by a couple of groups.
- Data management requirements may change or become clearer as a project progresses; how do best estimate requirements at the beginning of the project, how do we make DMP an ongoing process and who takes responsibility for data at different points in its lifecycle.
- Make sure non-digital research data is in the mix
- How do we capture algorithms and processes etc – the data may be useless without knowing how to process it
- What do we keep? Who decides? How?

RoaDMaP: Email: roadmap@leeds.ac.uk
 Web site: <http://library.leeds.ac.uk/roadmap-project>
 Blog: <http://blog.library.leeds.ac.uk/blog/roadmap>



3. We want to know more about..

This yielded a diverse list – most of which is reproduced here – there was no single area which was common across the groups.

- Everything!
- How to motivate / engage researchers (some participants suggested researchers may not want to know more than the bare minimum to satisfy funder requirements)
- Where does responsibility lie: what are the obligations of the researcher, the institution?
- Roles in the data lifecycle – who does what. Library role. IT role. Research support role etc.
- Research data management skills
- Cost benefit analysis for RDM
- Short and long term funding models
- Data ownership
- How to avoid institutions having “fig leaf” policies to cover their backs rather than something that is of real benefit to the researcher community and other stakeholders
- How to get policies through the institutional hierarchy
- Handling non-digital data
- Storing sensitive data – appraising what we can do with it
- Supporting groups that produce large quantities of data
- Ethics
- What are the consequences of *not* managing research data? How big’s the stick!
- Where to go for help
- How does RDM interact with other research management systems?
- The role of institutional repositories
- What IT systems are out there to help?
- What are other institutions doing already? Case studies of prior projects.
- Is there an institutional/research council/national framework that would make implementing RDM easy?
- National data centres. More about what they do and whether they can help us?
- Potential for cross-institutional collaboration
- How do we cite data?
- Funder requirements
- Applying good practice to non-RC funded research
- Digital curation and preservation
- Data versioning – what is the published version?
- Feasibility of a regional resources discovery system

RoaDMaP: Email: roadmap@leeds.ac.uk
Web site: <http://library.leeds.ac.uk/roadmap-project>
Blog: <http://blog.library.leeds.ac.uk/blog/roadmap>

