

1st LEARN Workshop, Embedding Research Data as part of the research cycle

London, 29th January 2016

Breakout Session Group 4

Chair: **Susan Reilly** (Stichting LIBER)

Rapporteurs: **Gema Bueno de la Fuente** (Stichting LIBER), **Imola Dora Traub** (University of Vienna)

Breakout group initial participants list:

Adrian Stevenson	JISC
Ben Mollitt	University of Liverpool
Benjamin Veasey	University of Nottingham
Betty Woessner	University of East London
Chris Browne	(Not known)
Danny Kingsley	University of Cambridge
David Tomkins	University of Oxford
Fiona Kyle	City University London
Gareth Knight	London School of Hygiene and Tropical Medicine
Gema Bueno de la Fuente	Stichting LIBER
Imola Dora Traub	University of Vienna
Jasmine Hunter Evans	Spa Bath University
Jennifer Basford	St George's University of London
Jenny Bunn	University College London
Kirsten Paterson	Kingston University
Laurence Horton	London School of Economics
Léa Croissant	University College London
Nadine Lewycky	The Open University
Sabina Leonelli	University of Exeter
Samantha Oakley	University of Swansea
Susan Reilly	Stichting LIBER
Tim Brooks	Anglia Ruskin University
Veronica Lawrence	University of London
Wayne Peters	King's College London

The **four topics** from the eight given themes were voted on by the participants. The topics were prioritised and discussed in the following order:

1) Engagement

Engagement with researchers. Are researchers aware of the issues around research data management – e.g. complying with research funder requirements, the need to construct research data management plans? How would you describe the level of awareness in your organisation and what can be done to raise that awareness amongst researchers, decision makers, funders, academic support staff (e.g. Library, IT)?

The group commented on the main drivers that could contribute to improving engagement with researchers:

- **High level policies** on research data management which include statements and an explanation of its importance, as well as clear requirements for research data management plans, could be helpful for developing understanding at the academic community.
- It is absolutely necessary to promote the **benefits** for individual researchers, general benefits or advantages for the whole community are not enough to convince and engage them.
- Requirements from funders to include an RDM plan could act as a driver, but these requirements are not common practice for the Humanities and Social Sciences.
- Foster **conversation** between researchers and funders.
- Creation of **online materials**, like a **FAQ**, blog, website to address researchers' concerns and questions on RDM.

Best Practice Example 1: One institution arranges open meetings between their researchers and funders. It provides researchers with the opportunity to ask direct questions of the funders regarding their policies. This increases clarity for researchers as often funder policies are not very clear. The outcomes of these sessions are also documents via a dedicated blog.

"What is your data?"

What about engaging with others librarians, IT services, managers... not just with researchers?

Also, the main barriers preventing effective engagement of researchers were discussed:

- **RDM is considered to be time-consuming**, and researchers are not willing to add an extra task to their workload.
- **Academic freedom** is usually presented as a reason for not engaging with RDM; researchers refuse to be obliged to do so.
- **Communication issues**: communication is not the same as engagement or awareness. But even communicating effectively is a challenge. Explaining what RDM is (and what it is not), why it is important, and how it should be carried out, is not always easy.
- There is a need for **funders** to improve their recognition **for RDM**. It is the funders' role, not only to mandate RDM as a requirement for funding, but also recognise it as a value, and to take into account the time and other resources needed to undertake it.
- In some disciplines, it is difficult to **identify their data**, what is the value of their data, or even if they are using data (as they may not consider it to be data).
- RDM is **not obvious for researchers**, not only may they not be aware of RDM practices, but they may also not recognise its usefulness and relevance for their research.
- Moreover, there are other concerns coming from researchers that prevent them from publishing and sharing their data, such as the **health and quality of data**.

*"Maximize the value of your data!
Optimize your data!"*

Research libraries or institutions: are they taking too much responsibility?

Researchers need to know why their data should be submitted to a repository, be aware of the objectives and the services offered. Repositories are not data dumps, where researchers can 'throw' their datasets, and then forget about them, because the repositories do all the work for them. That is why they need to think first about their data, the value of it, the workflow. A necessary approach is to **encourage researchers to reflect on their data and** work on their own data management plans.

It was also highlighted that there is a need for the institutions to consider and respect the **disciplinary angle**, regarding the **use of institutional repositories** and the **deposit mandate**. Researchers are usually more involved in their own community, and may prefer to use a disciplinary-specific repository.

The fact that researchers do not use the institutional repository for information retrieval is a known problem and recalls the debate on institutional vs. disciplinary repositories.

Change the discourse! It was argued that there is a need to change the discourse towards researchers, not only discussing the management of data, but using other keywords that highlight the benefits and the notion that RDM adds value to their research.

Best Practice Example 2: One institution has branded its RDM support service portal for researchers as “How to Add Value to Your Research?”

2) Incentives

What incentives do you think would help encourage researchers to share their research data – salary increase?; making data available for sharing and re-use becoming a recognised criterion for promotion?; encouragement to make data shareable by academic professional bodies?; research funders making requirements on research data as a condition of grant? Other?

Some key elements that could be incentives are:

- **Publishers’ mandates** (if there is a clear requirement)
 - **Research funders’ mandates:** but just useful if a Research Data Management Plan is required, and it is also recognised and then funded.
 - Time and other resources required for RDM should be granted, integrated into project management.
 - There should be different approaches depending on disciplines (**discipline level**)
 - It is important to show the risk of losing data, and how RDM could help prevent this problem.
 - **Training in RDM.** This training should start early, as it would be easier for **early career researchers** (PhD students) to assimilate new approaches. New generations may be more open to RDM.
- “If you want to be a top researcher, you have to learn how to do it!”*
- “How will data you collect contribute to your field?”*
- A **cultural shift** is needed, at all levels, and at all careers stages.
 - Evaluation, **recognition**, acknowledgment.
 - Understanding how it works.
 - For recognition, **re-use of datasets** could serve as a value indicator for researchers. A good strategy is to show researchers who is using their data.
- **Statistics, Altmetrics** are needed.
 - Researchers want to have **strong visibility, reputation** in their communities, so if sharing their data is recognised they would be more willing to do it.

Another key question has arisen on the limits and responsibilities of **data journals / repositories** - these difficulties should be clarified to help researchers understand how to proceed.

3) Training

What are the training and CPD (Continuing Professional Development) needs of researchers for research data management in your organisation? Training to help describe data so that it can be discovered for re-use? Support in writing and monitoring a research data management plan? Should universities include RDM training as a pre-requisite for research Masters and Doctoral degrees?

Barriers:

- Researchers do not want generic training, but **disciplinary specific training**
- General **lack of awareness**: it is not recognised as a problem by researchers
- In some disciplines (**Humanities** mainly) they do not think they use data, or create data for their research. And if they do, they have difficulties understanding how this data can be re-used.
- **Terminology barriers**: there is a need for terminology alignment, among disciplines

So, **different approaches** are needed, **depending on disciplines**, their methods, what they consider evidence and how data support this evidence.

RDM is necessary for **postgraduate research programmes**, but it is not always easy to embed into existing programs.

"Include RDM in academic practice!"

Methods training courses addressed to researchers at any stage of their career should include RDM.

Early adopters (data champions) can act as catalysts and champions for RDM, mentoring young researchers and spreading skills and awareness of benefits.

4) Costs

The costs of providing research data management infrastructure and support are significant. Do institutions have Case Studies on costing the provision on RDM support that they would be willing to share? The [LERU Roadmap for Research Data](#) contains 2 Case studies from UCL and Oxford in the UK. Do you have more Case Studies you can add?

"Indirect costs are difficult to measure"

A large research university costed its RDM support services at £160,000 pounds a year, around £460 / per project.

This cost analysis takes

"How much will it cost?"

into account not just the workflow of curation, ingest, and storage, but also the costs of personnel, and all the costs related to advocacy, training, promotion, analysis, etc. (**direct / indirect costs**). Projects may opt out of availing themselves of RDM support services. The current challenge to implementing such a cost model is that not all funders allow for RDM support to be charged against overheads or as a direct cost.

It is important to make funders aware of RDM costs and to integrate them in their funding models.