5th LEARN Workshop
Shaping and aligning research
data management policies
Barcelona, 26 January 2017

Outcomes from the Breakout Session

Group 1 | Chair: **Paul Ayris** (University College London)
Rapporteur: **Martin Moyle** (University College London)

Breakout Group 1 focused on issues arising from the draft Model Policy and the 'evaluation grid', as provided in the delegate packs. Discussion was wide-ranging, but may be grouped under the following headings:

Usefulness of LEARN Model Policy

It was agreed that the Model Policy is very useful. It covers the main points that delegates expected to see. The supplementary definitions (e.g. of 'research data') are helpful. The roles and responsibilities section capture all the important points. A statement about metadata levels and commitments was suggested as possibly worthy of inclusion.

Getting policies agreed

Most of the delegates were members of a Catalan consortium (CSUC) who were jointly preparing a policy for the region. It was agreed that a consortial/group approach was helpful in securing agreement for policies. Having said that, policy negotiation is still a complex process, involving dialogue with many parties: vice-rectors, research teams, central IT teams, libraries…the process could still take a long time.

Recommendation: universities should establish Research Data Management (RDM) offices to collate and co-ordinate activity across all the many stakeholders in RDM policy and practice. The UCL RDM team is a possible model for this.
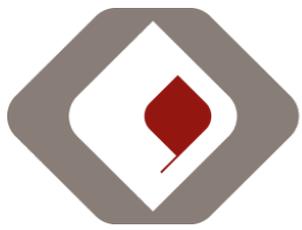
Evaluation criteria

It was agreed that the evaluation criteria used by University of Vienna were relevant. Data Management Planning (DMP) was identified as the single most important component of an RDM policy.

Retention, destruction

It was noted that the Vienna review showed that most of the 20 policies examined were weak in the areas of Ownership, Retention and Destruction. It was further noted that these areas were all linked; a clear understanding of ownership being a prerequisite for the execution of decisions about retention. The Model Policy does not explicitly address ownership, but it does make reference to the fact that different rights holders might have a stake in any RDM dataset; it was agreed that more definite assertions about ownership probably could not be made at policy level. However, it was suggested that the authors should look at this again to make sure that ownership issues are clearly signposted in the policy and/or the guidance.
In general, the policy should take a 'layered' approach, allowing for disciplinary differences. These will be addressed at DMP level.

5th LEARN Workshop
Shaping and aligning research
data management policies
Barcelona, 26 January 2017

Responsibility for destruction needs to be clear: who can make a destruction decision? It is the researcher's responsibility, but the institution needs to be able to make decisions on the researcher's behalf, for instance if there is no response from the researcher. A researcher may have left. Other parties may also have rights in the data and make their own stipulations about destruction.

Recommendation: ensure that responsibilities for decision-making around destruction are clear in the Model Policy. Perhaps a flowchart or decision tree would be useful as part of the guidance?
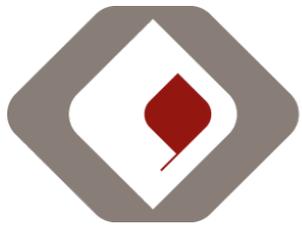
Costs

RDM does come with new costs. This fact is inescapable. However, it is impossible to generalise about cost. The Model Policy clearly commits institutions to meeting infrastructural needs, and therefore (implicitly) meeting any RDM costs. Open Science is often blamed for creating new costs: the important thing is that those trying to influence policy should have the arguments ready to ensure that an institution accepts the financial commitments necessary to underpin policy.

Internationally-produced data

One delegate asked about raw data whose production had been 'outsourced' to other countries (in this case, China and India). There was discussion about the scope of the policy and responsibilities in such situations. Where should such data be kept? The conclusion was that the person or institution who commissions the production of data has the responsibility to arrange for its curation. The Model Policy, as drafted, would apply in such situations.

Cultural heritage data; educational data; long tail

The group looked at paras. 23-25 of the evaluation template, noting that the policies under review had also all been very weak in these areas. RDM for cultural heritage data was agreed to be very important. It was recognised that RDM policies are often written in response to the sciences, and that they often disregard cultural heritage material, which was lamentable. It was noted that the LEARN Toolkit will have some guidance on material from the Arts and Social Sciences. In the context of educational data, it was agreed that the re-use of research data for educational purposes should be in the scope of the policy. The policy should be 'permissive', allowing any number of different approaches, including educational re-use. Additionally, it was agreed that educating students about RDM and FAIR data principles was very important. The LEARN Toolkit will have a study on Open Education. Finally, it was acknowledged that policies do need to address long-tail data. It was felt that policies tend to be led by those working in a 'big data' environment, but most of the data sets at UCL, for example, are thought to be small, 'long-tail', data.

5th LEARN Workshop
Shaping and aligning research
data management policies
Barcelona, 26 January 2017

Group 2 | Chair: **Wouter Schallier** (CEPAL/ECLAC)
Rapporteur: **Gema Bueno** (LIBER Europe)

Participants in this Breakout Session included: researchers, librarians in charge of repositories and research data management, people in charge of research support, and a representative of university management. The main points discussed can be summarised in the following sections

Institutional policies

It is not common to have a person devoted to research data at universities. Most institutions don't have a policy about research data either. However, there is currently a joint effort going on under the CSUC umbrella to implement RDM policies in the universities who are member of this consortium. A first document, on the need and how to develop a RDM policy, was approved by the (vice-)rectors, and now a working group has the commitment to develop the framework policy. The CSUC policy is essentially based on the LEARN Model Policy, which each university can adapt to its own context. Initially the objective was to develop a policy for each university this year, but it seems that this will probably take more time. This cooperative effort seems to be more successful than individual initiatives.

At University Carlos III of Madrid (UC3M) they are planning to implement a more comprehensive policy of Open Science, that includes RDM but also Open Access.

All participants agreed that there is no perfect policy, but that research institutes are probably delaying for too long and waiting for others to take the lead.

It is a challenge to get the approval of the whole university and government communities, since there are different points of view on the principles of the policies and their implementation.
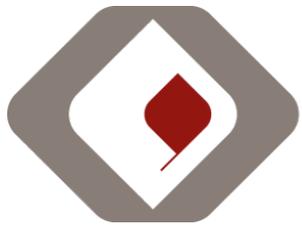
Incentives

Researchers need incentives to make RDM part of their workflows. It should be easy, fast, interesting and rewarding for researchers to make their data available. Adding more to the administrative burden is definitely not the way to go: on the contrary, an effective RDM plan should probably allow researchers to lessen it. Other incentives could be: financial, legal and visibility.

Just one university in the group is working on incentives. UPC, with its post prints initiative, ensures that being open has a direct, economic benefit for researchers. The participants agreed that, for the issue of incentives to be really successful, a legal mandate on the national level would be required.

Researchers

Researchers expect an easy technical solution ("magic button") for sharing their data. On the other hand, they also have worries about openness, in a sense that it could hinder their promotion, career development etc. "Young researchers" seem to be more willing to change current academic practices and act as advocates, on condition, of course, that an adequate scheme of incentives will be in place.

5th LEARN Workshop
Shaping and aligning research
data management policies
Barcelona, 26 January 2017

A cultural shift is also needed in the publication policies. But traditional publishers are still too strong. Some of them understand the enormous value of research data and are acquiring research data services to the level that there is a serious risk of data privatisation.

The differences among disciplines are still big. Humanities and social sciences usually are much more reluctant. The meaning of data for these researchers is different, and they often do not see themselves as being responsible for managing their data. Several of them also fear to lose competitivity, and are not really challenged to leave their comfort zone.

Support for researchers

Librarians and other research support staff should also be challenged to leave their own comfort zone. They should become more confident and comfortable in giving RDM support to their researchers. Research support staff should also be proactive in approaching researchers to offer their new services. One cannot simply wait until the researcher takes the initiative… These new RDM support services can consist of: help desk, FAQ, thematic portals, training, infrastructure, budget (for example for acquiring DOI's) etc.

Internally, libraries and other research support staff urgently need to acquire RDM skills. This can be done by re-training existing staff or even by attracting new job profiles (data scientists or similar). A lot of work stills needs to be done to clarify the precise skills needed for successful RDM at research institutes. Data scientist covers a large spectrum of skills, but not all of these are necessarily needed.

Conclusion

Participants thanked the LEARN project for this timely discussion on RDM. The LEARN deliverables are very useful, and it is now up to the institutions to take their responsibility for the implementation of RDM policies.

5th LEARN Workshop
Shaping and aligning research
data management policies
Barcelona, 26 January 2017

Group 3 | Chair: **Barbara Sánchez Solís** (University of Vienna)
Rapporteur: **Raman Ganguly** (University of Vienna)

During the session the group discussed the current Model Policy created by LEARN Work Package 3 "Policy Development and Alignment" and it ran through different aspects of the research data management process.

At the beginning the main stakeholders in the policy making process where identified: ICT, Libraries, research support units, legal services, institution's trade union representatives and funders.

The question "When should different stakeholders be involved in the decision making process?" was raised, focusing then on how to involve ICT departments. The general opinion was that the ICT should be involved at a very early project stage in order to assure common strategic alignment.

Concerning the aim of the creation of the policy, the main outcome can be summarised in the following sentence: "Keep in mind, the policy should be implementable!" This means managing a balancing act between some degree of granularity and a certain level of abstraction. In any case, the policy should be as close as possible to practicability. Research data management is not a service, it is a capability and it has to be implemented in a collaborative way.

A further aspect that was considered important was that policies should be machine-readable to allow comparisons between different policies and facilitate the monitoring of key performance indicators. In this way, the compliance with policies could be monitored as well.

Costs were also mentioned: How much does it cost to have a policy? And on the other hand, how much does it cost <u>not</u> to have a policy? It was generally agreed that implementing a policy would raise costs for a research institution because it implies the implementation of services (research support services, legal services, infrastructures, etc.).

The discussion also involved the issue of formats, for instance asking which formats should be managed? Should this be decided in the main policy or rather should this be decided by the different university schools? The promotion of FAIR principles should be facilitated by open formats. Attenders were not sure if the discussion concerning policies should also be related to the topic of formats. However, it was stated, that the handling of metadata should be part of a policy.

The scope of a RDM policy should also include educational data and cultural heritage.

Further steps concerning the implementation of a policy are: training activities, legal support, clear definition of roles and competencies and work on terminology.