



Grant agreement no. 654139

LEARN**Leaders Activating Research Networks: Implementing
the LERU Research Data Roadmap and Toolkit***Coordination & support action*

H2020-INFRA-SUPP-2014-2

Topic: e-Infrastructure policy development and international cooperation

D1.5 Workshops (5)

Work Package: WP1

Due date of deliverable: month 20

Actual submission date: 06 / April / 2017

Start date of project: 1 June 2015 Duration: 24 months

Lead beneficiary for this deliverable: *UCL*
Contributors: *LIBER, UB, UNIVIE, UN ECLAC*

Project co-funded by the European Commission within the H2020 Programme (2014-2020)		
Dissemination Level		
PU	Public	✓
CO	Confidential, only for members of the consortium (including the Commission Services)	
CI	Classified, as referred to in Commission Decision 2001/844/EC	

Disclaimer

The content of this deliverable does not reflect the official opinion of the European Union. Responsibility for the information and views expressed herein lies entirely with the author(s).

Table of contents

1.	Version log	4
2.	Definition and acronyms	5
3.	Introduction.....	6
4.	Activities carried out and results	7
	4.1 Organisation	7
	4.2 Programme.....	7
	4.3 Evaluation and feedback	8
5	Conclusions	9
	Appendix A.....	13
	Appendix B.....	14

1. Version log

Version	Date	Released by	Nature of Change
First version	08/03/2017	Ignasi Labastida (UB)	
Revision	09/03/2017	Rodney Amis (UCL)	English review
Reviewed	06/04/17	Paul Ayris (UCL)	Final review

2. Definition and acronyms

Acronyms	Definitions
ICT	Information and Communication Technology
LEARN	LEaders Activating Research Networks: Implementing the LERU Research Data Roadmap and Toolkit
LERU	League of European Research Universities
LERU Roadmap	Roadmap for universities and research organisations on how to tackle the challenges which research data poses. It also has a series of messages for researchers, support services, research institutions and policy makers.
LIBER	Association of European Research Libraries
RDM	Research Data Management
UB	Universitat de Barcelona
UCL	University College London
UNIVIE	Universität Wien
UN ECLAC	United Nations Economic Commission for Latin America and the Caribbean

3. Introduction

Work Package 1 “Stakeholder Engagement” (Task 1.1) states that the LEARN Project will organise five Workshops to be organised in Europe and Chile. These Workshops form one of the main project activities. Speakers who are leaders in the field of RDM will address the Workshops. Attendees will come from the various sectors of the research community. The communities from which attendees will be sought will be identified by the project partners. The Workshops will be targeted to encourage attendance by all stakeholder groups in the research community – researchers themselves, data scientists/stewards, library and IT staff, faculty and University managers, research funders. The project partners will ensure that attendance comes from all stakeholder groups by using their influence in their own communities to encourage attendance. The mix of stakeholder groups will ensure that the Workshop presentations and discussions are broad and cover all relevant aspects of RDM. The Breakout Groups in the Workshops will help identify best practice and case studies which are significant to these sectors. These will then be gathered and published in the Toolkit.

This deliverable is the report on the 5th LEARN Project Workshop entitled ‘*Make Research Data Management policies work*’, held in Barcelona on the 26th of January 2017.

4. Activities carried out and results

4.1 Organisation

The venue for the Workshop was the Aula Magna of the University of Barcelona. The logistics for the event included catering and professional photographer support. The Workshop was publicised and promoted via the LEARN Project's website and Twitter feeds as well as through direct mailings to relevant mailing lists and identified contacts supplied by all the Project partners. Registration for the Workshop was possible at the website using the Eventbrite booking system. A maximum ceiling of 75 registrants was set in order to limit the day's three break-out groups to no more than 25 people in each.

The Workshop attracted 69 registrants of whom 50 attended on the day, plus 2 on-site registrations, and therefore the total number of participants reached 52. The range of institutions represented by the registrants is listed in Appendix A.

A welcome pack was assembled providing each registrant with the Workshop programme, a copy of the *Draft Model Policy v 1.5 for Research Data*, the first page of the Evaluation Grid for RDM Policies in Europe, and details of the afternoon break-out groups (see 4.2.2 below).

4.2 Programme

The Workshop programme was developed to allow for five keynote speakers (four external and one from within the Project (UCL)), three break-out discussion groups and a panel session with a moderator (UNIVIE) and six panellists (the chairs of the break-out groups along with three of the morning speakers).

The keynote speakers' presentation slides can be viewed and downloaded at <http://learn-rdm.eu/en/events/5th-learn-workshop/presentations/>. They are linked to the Project's Slideshare account. Moreover, the morning session was recorded and the videos are available at <http://learn-rdm.eu/en/events/5th-learn-workshop/videos/>. The full programme is listed in Appendix B

4.2.1 Keynote speakers

Profiles of all the five keynote speakers can be viewed at <http://learn-rdm.eu/en/events/5th-learn-workshop/speakers-profiles/>

4.2.2 Break-out discussion groups

Registrants were distributed in three break-out groups at the registration. Each session discussed the current model policy elaborated by LEARN WP3 together with the grid used to evaluate some of the existing policies in Europe. This evaluation analysis had also been carried out by LEARN WP3. Each break-out group was serviced by a chairperson and a rapporteur provided from one of the Project partners. Rapporteurs took notes during the session and compiled full reports after the event. A summary report illustrating the main themes to come out of all the discussions was also produced and

made available on the Project's website <http://learn-rdm.eu/en/events/5th-learn-workshop/outputs/>.

4.3 Evaluation and feedback

All attendees were contacted in the week following the Workshop and invited to complete an online satisfaction survey. 34 of the 52 attendees completed the survey and the results were made available via the Project website <http://learn-rdm.eu/en/events/5th-learn-workshop/evaluation/>.

5 Conclusions

The fifth LEARN Workshop attracted mainly local participants and some from other Spanish institutions. There was a clear interest in knowing about the model policy drafted by the LEARN Project since the Catalan universities are committed to adopt an institutional policy during this year 2017 and they have already established a working group, which is using the model LEARN policy as a template.

The Project partners were interested to collect feedback from attenders in order to create the final form of the LEARN policy document.

Usefulness of a policy

It was agreed that a model policy is very useful. The current document 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.

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

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

Since most of the participants were members of Catalan institutions, 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.

Some institutions are planning to implement a more comprehensive policy of Open Science, that includes RDM but also Open Access to publications.

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.

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 (<http://www.ucl.ac.uk/library/research-support/research-data>).

Policy Implementation

Concerning the aims for creating 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 to facilitate the monitoring of key performance indicators. In this way, compliance with policies could be monitored as well.

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

Policy Evaluation criteria

It was agreed that the evaluation criteria used by LEARN WP3 were relevant. Data Management Planning (DMP) was identified as the single most important component of an RDM policy.

Data Formats

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 considered by the different university schools? The promotion of FAIR principles should be facilitated by open formats. Attendees 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.

Retention, destruction

It was noted that the evaluation showed that most of the 20 policies examined as part of WP3 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 LEARN RDM 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.

Responsibility for destruction needs to be clear: who can make a destruction decision and when? 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.

How much does it cost to operate a policy? And on the other hand, how much does it cost not 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.).

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.

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.

A cultural shift is also needed in publication policies. However, 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 are usually much more reluctant to engage. 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 competitiveness through sharing, 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 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 still needs to be done to clarify the precise skills needed for successful RDM at research institutes. Data science covers a large spectrum of skills, but not all of these are necessarily needed.

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 needs of 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, Humanities 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 case 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.

Final 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 forward.

Appendix A

Institutions represented at the Barcelona LEARN Workshop

Apliscience	Università di Torino
Centre For Genomic Regulation	Universitat Autònoma de Barcelona
CIDOB	Universitat de Barcelona
CSIC	Universitat de Girona
CSUC	Universitat de València
UN ECLAC, United Nations	Universitat Jaume I
INIA	Universitat Oberta de Catalunya
LIBER Europe	Universitat Politècnica de Catalunya
National Institutes of Health	Universitat Politècnica de Valencia
Open Knowledge International	Universitat Pompeu Fabra
Science Europe	Universitat Ramon Llull
Universidad Carlos III de Madrid	University College London (UCL)
Universidad de Navarra	University of Vienna
Universidad de Zaragoza	

Appendix B

5th LEARN Workshop

‘Shaping and Aligning Research Data Management Policies’

26th January 2017, Aula Magna Universitat de Barcelona

Programme

09.00 – 9.30 *Registration*

09.30 – 9.45 *Welcome remarks*

Adelaida Ferrer, CRAI Director, Universitat de Barcelona
Oriol Pujol, Vice-President for Digital Transformation, Universitat de Barcelona

09.45 – 10.10 Paul Ayris, Pro Vice-Provost & Director of Library Services, University College London, “*Delivering Research Data Management in the ‘Data Deluge’*”

10.10 – 10.50 Marie Timmermann, EU Legislative and Regulatory Affairs Officer, Science Europe, “*Developing a framework for Research Data Management Protocols*”

10.50 – 11.20 *Coffee break*

11.20 – 12.00 Daniel Mietchen, National Institutes of Health, “*Integrating policy and infrastructure in research data management: the perspective of researchers*”

12.00 – 12.40 Eva Méndez, Deputy Vice-President for Strategy and Digital Education of the Universidad Carlos III de Madrid, “*Opening research data in EU universities: Policies, motivators and challenges*”

12.40 – 13.20 Oriol Pujol, Department of Mathematics and Computer Science, Universitat de Barcelona, “*About data from a machine learning perspective*”,

13.20 – 14:30 *Lunch*

14.30 – 16.00 Three breakout sessions. Each led by a LEARN project partner

16.00 – 17.00 Feedback from the Breakout sessions and Plenary Discussion chaired by Paolo Budroni, Library and Archive Services, Universität Wien