DOI: https://doi.org/10.14324/000.learn.01

Introduction

In April 2016, Jisc issued an information leaflet on *Managing research data in your institution*.[1] They concluded that 'data needs to be selected, curated, retained and stored, using appropriate metadata'. The call was timely and aimed at research performing institutions. Similarly, the Research Data Alliance (RDA) also makes an important offering in the research data space.[2] The RDA is an international organisation focused on the development of infrastructure and community activities aimed to reduce barriers to data sharing and exchange, and to promote the acceleration of data-driven innovation worldwide. With over 4,500 members globally, the RDA comprises individuals, organisations and policy makers representing multiple industries and disciplines, who are committed to building the social, organisational and technical infrastructure needed to reduce barriers to data sharing and exchange, and to accelerating data-driven innovation worldwide.

From 11-17 September 2016, more than 850 data professionals and researchers from all disciplines around the globe convened in Denver, Colorado, for the first edition of International Data Week (IDW). This landmark event, organised by CODATA, the Committee on Data of the International Council of Science (ICSU), the ICSU World Data System (WDS) and the Research Data Alliance (RDA), brought together data scientists, researchers, industry leaders, entrepreneurs, policymakers and data stewards from all disciplines to explore how best to exploit the data revolution in order to improve science and society through data-driven discovery and innovation.[3]

In the UK, the Digital Curation Centre (DCC) provides access to a range of resources including How-to Guides, case studies and online services. Their training programmes aim to equip researchers and data custodians with the skills they need to manage and share data effectively. The DCC also provides consultancy and support for issues such as policy development and data management planning.[4]

Clearly, research data management is a topic of wide interest. The *Research Data Curation Bibliography* by Charles W. Bailey lists over 620 selected English-language articles, books, and technical reports that are useful in understanding the curation of digital research data in academic and other research institutions.[5]

What issues are current for those involved in RDM? For decision makers, the primary issue is probably that of the associated costs. The 4C project offers an overview of relevant cost models.[6] One of these is the LIFE model – Lifecycle Information For E-literature – for which one of the LEARN project partners (UCL) was a joint lead.[7] The LIFE costing model is:

$$\mathbf{L = C + Aq_T + I_T + M_T + BP_T + CP_T + Ac_T}$$

where
L = Complete lifecycle cost over time 0 to T.
C = Creation
Aq = Acquisition
I = Ingest
M = Metadata Creation
BP = Bit-stream Preservation
CP = Content Preservation
Ac = Access
T = Period of time over which identified activity lasts

However, there is an elephant in the room with regard to RDM costing. As 4C says, 'There is a sizeable canon of research into cost modelling for digital curation but the research is in many ways preliminary and there has been little uptake of the tools and methods that have been developed. For example, tools to manage and estimate costs have not been integrated into other digital curation processes or tools.'[8] This has made it extremely difficult for research performing institutions to take RDM forward locally when total costs are unclear, for decision makers do not write blank cheques. Even when costs are known, many institutions are unable or unwilling to reveal their costing activities.

Another issue which is setting the agenda for RDM in Europe is the recent publication of the Commission's High Level Expert Group Report on the European Open Science Cloud.[9] This Report has at its kernel the benefits which Open Data can bring to research communities. It bemoans the current fragmentation in the European research data landscape and states starkly, 'There is no dedicated and mandated effort or instrument to coordinate EOSC-type activities across Member States'.[10]

1. Jisc: https://www.jisc.ac.uk/guides/research-data-management; last accessed 8 February 2017.
2. RDA: https://www.rd-alliance.org/; last accessed 8 February 2017.
3. CODATA: http://www.codata.org/; last accessed 8 February 2017.
4. DCC: http://www.dcc.ac.uk/about-us; last accessed 8 February 2017.
5. Bailey, C.W.: *Research Data Curation Bibliography,* version 7: 01/24/2017: http://digital-scholarship.org/rdcb/rdcb.htm; last accessed 8 February 2017.
6. 4C: http://www.4cproject.eu/summary-of-cost-models/; last accessed 9 February 2017.
7. LIFE3: http://www.4cproject.eu/summary-of-cost-models/16-community-resources/outputs-and-deliverables/105-life3-costing-model-life3/; last accessed 9 February 2017.
8. 4C: http://www.4cproject.eu/about-us/; last accessed 9 February 2017.
9. European Commission: http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud; last accessed 9 February 2017.
10. *Ibid*., p. 6.

At institutional level, a baseline was drawn by the LERU Roadmap for Research Data,[1] which was published in December 2013. This was the first document to look in detail at the opportunities and challenges which face European research performing organisations in the RDM space. LERU is the League of European Research Universities, comprising 23 members in 12 European countries. Two members of the LEARN project, UCL and the University of Barcelona, are also members of LERU.[2] The LERU Roadmap classified the issues facing their members under seven headings:

- Policy and Leadership
- Advocacy
- Selection, Collection, Curation, Description, Citation, Legal Issues
- Research Data Infrastructure
- Costs
- Roles, Responsibilities, Skills
- Recommendations

Whilst the Roadmap was written on behalf of LERU members, the issues it analysed were in fact generic and can be said to apply to any research performing organisation anywhere in the world – they are by no means exclusive to research-intensive universities. The Roadmap 'presents a series of blueprints which LERU members, indeed any European university, could use to begin to tackle the challenges which research data poses. It also has a series of messages for researchers, research institutions, support services and policy makers'.[3]

The EU-funded LEARN project took up where the LERU Roadmap finished. The five project partners – UCL, LIBER, Barcelona, Vienna and the UN Economic Commission for Latin America and the Caribbean[14] – identified what was needed to embed the principles of the LERU Roadmap into research performing institutions. Accordingly, the LEARN project set out to provide:

1. A model RDM policy for research performing institutions
2. A Toolkit of Best Practice Case Studies, illustrating the challenges and opportunities identified in the LERU Roadmap
3. 5 Workshops to examine the issues and to produce material for the Case Studies – in London, Vienna, Helsinki, Santiago in Chile, and Barcelona[15]
4. An Executive Briefing on RDM challenges and opportunities in six languages
5. A self-assessment survey to allow institutions to assess their level of preparation for RDM, with an analysis of the findings
6. Key Performance Indicators (KPI) to assess whether all elements of the LEARN template for a RDM policy are included in institutional policy work; and a set of KPIs to measure implementation of the policy
7. Lists of Further Reading and a Glossary of technical RDM terms[16]

11. LERU: http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf; last accessed 9 February 2017.
12. LERU: http://www.leru.org/index.php/public/about-leru/members/; last accessed 9 February 2017.
13. LERU: http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf, p. 3; last accessed 9 February 2017.
14 LEARN: http://learn-rdm.eu/en/partners/; last accessed 9 February 2017.
15 LEARN: http://learn-rdm.eu/en/events/; last accessed 9 February 2017.
16 LEARN: http://learn-rdm.eu/en/category/further-reading/; last accessed 9 February 2017.

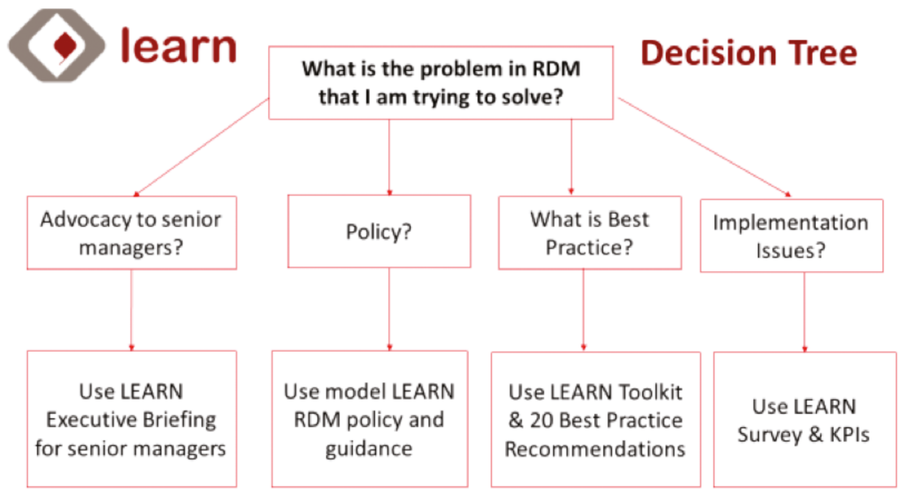
This Toolkit is a major deliverable of the LEARN project. It takes the issues identified above in the LERU Roadmap and provides templates and Case Studies on how success in implementation can be achieved. Many of the formal LEARN outputs and Deliverables are therefore embedded in the Toolkit.

Part 1 consists of 23 Case Studies. These are drawn from issues in the original LERU Roadmap, enlarged by discussions and interactions in the five LEARN Workshops. Briefings in six languages, summarising the main findings in the Roadmap, were prepared for the Workshop participants.[17] Workshops were attended by a range of stakeholders: researchers, research funders, librarians, IT specialists, publishers, senior institutional decision makers, and RDM specialists. The sessions consisted of formal presentations, with discussion and breakout sessions, allowing for wider discourse. In Santiago, the breakout sessions were replaced with panel sessions. Latin America and the Caribbean supplemented their formal Workshop with three mini-Workshops to gain more feedback. LEARN also held a café session at the Amsterdam Open Science Conference in April 2016. Feedback from all these sessions was analysed, resulting in the Case Studies contained in Part 1 of the LEARN Toolkit.

Part 2 of the Toolkit contains the Model RDM Policy produced by the University of Vienna, accompanied by Guidance and an Evaluation Grid of 20 European RDM policies which helped the Vienna Team to formulate the Model LEARN Policy. Part 3 consists of an Executive Briefing in six languages aimed at senior institutional decision makers. It takes the main points identified in the LEARN project and explains how senior decision makers can interact with these issues.

How is the LEARN Toolkit best used? It is important to note that the Toolkit is not itself a Roadmap, plotting a particular route. The LERU Roadmap for Research Data provided the original roadmap, which was particularly relevant to research performing organisations. Rather, the Toolkit is a collection of tools and services, which allows the user to tackle particular challenges on the journey to deliver sound research data management practice at institutional level.

The LEARN Decision Tree is the key to unlocking the treasures in the Toolkit.



17 LEARN: http://learn-rdm.eu/en/outputs/project-materials/; last accessed 9 February 2017.

To navigate through the Toolkit successfully, it is important that the user clearly articulates the problem they are trying to solve. Having defined the question, the user can then start looking for answers. For advocacy to senior managers, use the Executive Briefing. For policy development, use the model LEARN RDM policy and guidance. To identify what is Best Practice, use both the 20 Recommendations on Best Practice[18] emanating from the LEARN Workshops and the Best Practice Case Studies in the Toolkit. The key message in each of the Case Studies is summarised in the final section of Part 1, the Conclusions. Want to measure your success in implementing RDM practices? Use the LEARN self-assessment survey[19] and the Key Performance Indicators[20].

The LEARN Tookit provides an armoury of best practice for all research performing organisations who wish to develop a persuasive RDM offering. We live in an era of data deluge and institutions who remain unprepared to tackle these challenges/seize these opportunities do so at their peril.

[18] LEARN: http://learn-rdm.eu/wp-content/uploads/20-RDM-Policy-Recommendations.pdf; last accessed 14 June 2017.
[19] See Case Study 23 in this Toolkit, pp. 125-127.
[16] LEARN: http://learn-rdm.eu/wp-content/uploads/FinalKPITable.pdf; last accessed 14 June 2017.