The LERU Roadmap for Research Data – a tool for institutions from Latin America and the Caribbean

The LERU Roadmap for Research Data plots a course which universities, and **indeed any research institution around from Latin America and the Caribbean**, can choose to follow in order to implement sound research data management practices at institutional level.

Research data, from the point of view of the institution with a responsibility for managing the data, includes:

- all data which is created by researchers in the course of their work, and for which the institution has a curatorial responsibility for at least as long as the code and relevant archives/record keeping acts require,

- and third-party data which may have originated within the institution or come from elsewhere.

Scientific research is increasingly data driven. The use and sharing of research data has the potential to fundamentally change the way research is undertaken and disseminated. Universities and research institutions with a **Research Data Management Plan** are able to make optimal use of this important revolution in science.

The Roadmap is divided into 6 chapters. The final chapter, chapter 7, contains a list of recommendations to different stakeholder groups.

Chapter 1 looks at the ideas of Policy and Leadership, showing that universities and research institutions have responded to a greater or lesser degree to (national) data policy directives. It argues that what is now needed are *institutional* data management policies and accompanying Roadmaps for their implementation.

The Data Publication Pyramid
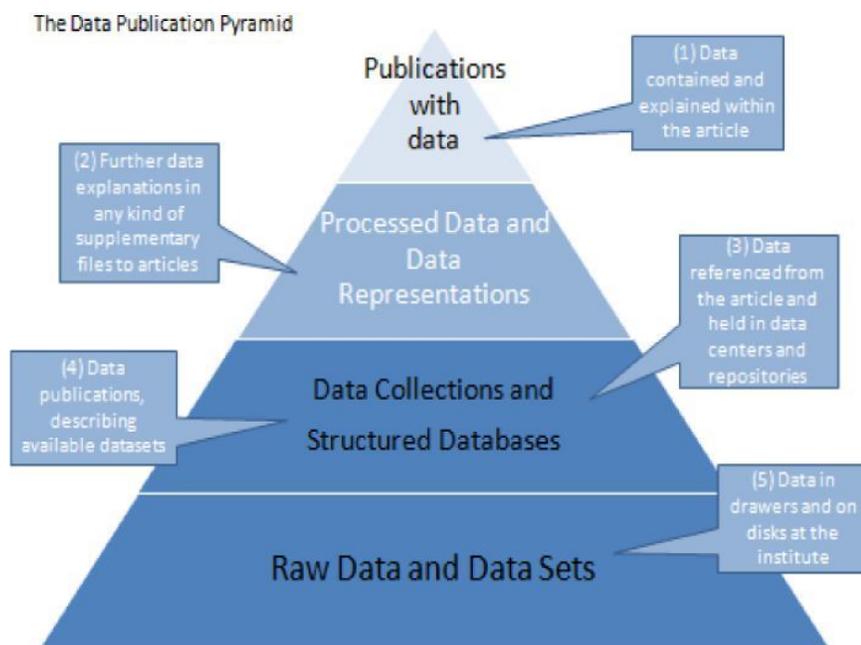
Chapter 2 looks at the issues of Advocacy, which the Roadmap identifies as crucial to successful data sharing. The Roadmap identifies incentives and barriers to data sharing, along with suggestions for how to overcome the reluctance of some researchers and/or institutions to share their data. Open research data is advocated as a goal for all researchers and institutions, where this is possible. This requires leadership at an institutional level. University and research institutions support services are well placed to advocate for best practice in research data management and data citation. Advocacy can underline the rewards inherent in data sharing, help to make data visible, increase collaboration and data reuse, and help to build the necessary trust to make all this happen.

Chapter 3 looks at a range of issues involved in the Management of Research Data: Selection and Collection, Curation, Description, Citation and Legal Issues. For selection and curation, the Roadmap takes as its starting point the ODE Data Publication pyramid[1], and recommends that the research community should undertake further work to identify which

---

[1] See http://www.alliancepermanentaccess.org/wp-content/uploads/sites/7/downloads/2011/11/ODE-ReportOnIntegrationOfDataAndPublications-1_1.pdf

of the strata in the pyramid can be made available for sharing and reuse, and which can be open. For data curation, the Roadmap analyses the research workflow and then suggests how the necessary infrastructures can be created. For Description, the Roadmap underlines the difficulties inherent in encouraging researchers accurately to describe their data. For Citation, examples of best practice are given. The section on Legal Issues analyses the copyright frameworks in Europe and suggests that important changes are required to allow Data Science to flourish.

Chapter 4 looks at Research Data Infrastructures. These infrastructures can be classified into 4 types:

- research data itself
- data management
- data management tools
- technical components staffing

Research data infrastructure needs to offer a generic framework to accommodate the wide variety of research activities which will make use of it. An overview of research data management tools is provided and the chapter highlights that the 'long tail' of research data residing on local desktops, hard discs and servers might well comprise a bigger challenge than 'big data'. In terms of technical components, the chapter outlines how these components are distributed across the institution and that, ideally, support services should be organized as a coherent whole.
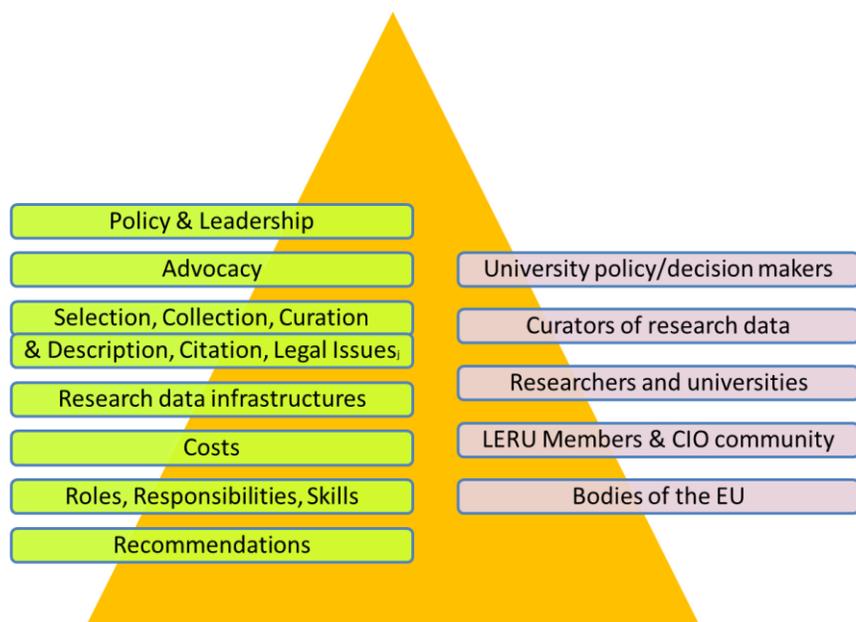
Chapter 5 tackles the difficult issue of Costs. There is no one single model which can be used to calculate costs. It provides 2 case studies, for the University of Oxford and UCL (University College London) to give indicative costs for service provision. The chapter shows that cost benefits sometimes provide a framework for judging the cost effectiveness of research data curation. It also shows who is likely to meet the costs – research funder, national collaborative service, or the university and research organization itself.

Chapter 6 looks at Roles, Responsibilities and Skills. The chapter undertakes an analysis of the different roles needed/involved in research data management and the responsibilities that these post holders have. It suggests that a new concept of Data Scientist has the potential to become a new role in its own right. The chapter also identifies the training requirements needed for a range of participants such as postgraduates/PhD students, senior researchers, librarians and data scientists.

The final chapter, **Chapter 7**, brings together 44 Recommendations drawn from the Roadmap and allocates them to specific audiences: institutional policy and decision makers, all those involved in the curation of research data, researchers and their institutions, LERU members and the LERU CIO community and the bodies of the European Union (most of which could apply to any university and research organisation in Latin America and the Caribbean wishing to engage in a regional collaboration with other institutions).