Training subject librarians in Research Data Management

Authors: June Hedges (Head of Liaison and Support Services, Library Services, University College London) and Myriam Fellous-Sigrist (Research Data Support Officer, Library Services, University College London)
Email: m.fellous-sigrist@ucl.ac.uk

19.1. INTRODUCTION

University College London (UCL) ranks among the top twenty universities in the world and is one of the most successful British research institutions at attracting funding. Almost all academic disciplines are represented in its 380 research departments, units, institutes and centres. UCL is home to 12,000 research staff and research students.

UCL Library Services run eighteen libraries which support UCL’s teaching and research activities, including one in the award-winning School of Slavonic and East European Studies building and several that provide services to both UCL and the National Health Service. The combined staff in UCL Library Services totals 263 FTE (full-time equivalents). Amongst this number are around 30 subject liaison and site librarians who have responsibility for supporting the research and teaching of the institution. These librarians are the primary points of contact for academics, researchers, UCL staff and students. They provide subject-specific support and advice on resources and collections, offer training to staff and students, and promote and provide training on the various teaching and research support services that the Library offers, including open access services.

Two Research Data Support Officers work as part of the same team as well as in close collaboration with the UCL Information System Division (IT Services) and several other central services. These officers coordinate Research Data Management (RDM) advocacy and support across the institution. To ensure the long-term sustainability and scalability of the RDM support service, as well as sufficient subject discipline support, the RDM team aims to foster several support networks of subject-specific experts across the university. The subject liaison and site librarians form one of these networks.

19.2. INTRODUCING LIBRARIANS TO RDM THROUGH WORKSHOPS

The first UCL Research Data Policy (launched in August 2013) was accompanied by introductory presentations on RDM and related service developments for Library Services’ staff. A programme of three day-long workshops was subsequently planned to inform and train library staff about current issues in Research Data Management, from key definitions up to the review of Data Management Plans. These workshops took place in 2015; they gathered between 30 and 35 participants each. The sessions were designed and delivered by Data Management experts from the Information School of the University of Sheffield. The outline for each workshop was as follows:

**Session 1:** This workshop provided an introduction to research data and its management in the context of UCL and the Library’s role. Topics covered included the nature of data and research data services; data management planning; information security; and issues around Text and Data Mining. At the end of the session, librarians were tasked with talking to a researcher about RDM in advance of the next workshop.

**Session 2:** This workshop started with presentations on what participants had learnt from their conversations with their researchers. This was followed by a discussion about the survey approach to gathering further information about data management in the university. Librarians looked at identifying key choices in planning training and at the issues around selecting, describing and citing data. At the end of the day, participants were tasked with group exercises to prepare the last workshop.

**Session 3:** In this workshop, participants heard from each group what their ideas and plans were for addressing the various aspects of RDM support identified in the first two sessions. Librarians presented group reports on practical RDM, requirements gathering (and in particular the Data Asset Framework method), data sharing, data sources and RDM websites.

The first two sessions comprised presentations by the workshop leaders introducing a new area, approach or issue coupled with group activities which led to the final project of creating a plan on how to respond to key aspects of RDM. This was done in small groups working together outside the workshops to prepare their response and also a brief presentation to be delivered at the final workshop.

Feedback from the workshops was positive and librarians welcomed the thoroughness of the programme, which did cover all the key aspects of RDM. They also welcomed the opportunity to work collaboratively with colleagues from across Library Services. The only negative feedback received was that the workshops had been spread over a 6-month period with 2-3 month intervals between each.

In 2016 a fourth workshop focused on central RDM services run by UCL Library Services and the UCL Information Systems Division jointly. The 30 participants discussed the roles and interaction of these different services, and how these are explained to researchers and research students across all disciplines. This event was also planned as a networking opportunity which gathered together for the first time librarians, Research IT staff and departmental data managers.

Feedback received from this event showed that it helped to meet colleagues across the university and to understand how the different services join up. Theoretical presentations were considered to be less useful, and several participants suggested that having small group activities to put into practice what was said in the presentations would enhance their learning; this would include creating flow-charts to explain data storage processes within the university and drafting discipline-specific guidance for researchers and students.

19.3. INVOLVING LIBRARIANS IN RDM THROUGH PARTICIPATION IN A WORKING GROUP

Subsequent to formal RDM training, librarians were given the opportunity to apply their knowledge by actively contributing to a new Working Group. The Library RDM Working Group was created in 2015. It supports the two Research Data Support Officers with discipline-specific knowledge and essential staff resource for short-term projects. This Group is made up of thirteen volunteers (Librarians, Records Manager, Digital Curation Manager, Research Data Support Officers) who work on a specific project each summer; not all Working Group members are required to participate in all projects.

1 As listed in the UCL Departments A to Z (at www.ucl.ac.uk/departments/a-z/); accessed 4 August 2016.
2 Figure from UCL Human Resources as of 1st October 2015 and Registry Services as of 1st December 2015.
3 This method is explained by the Digital Curation Centre at www.dcc.ac.uk/resources/tools/data-asset-framework/; accessed 20 December 2016.
In 2015, the Group first concentrated on building the new RDM website, and second on designing and promoting a cross-university RDM survey. For the first activity, all Group members were trained to improve their knowledge of the Library Content Management System and to write for the web. They worked in pairs to draft, edit and publish online the webpages that they had chosen to work on. The website was completed in two months and launched at the start of the academic year 2015/2016. It featured nine how-to guides, a section about the university’s and research funders’ policies on research data, key definitions about RDM, a searchable list of Frequently Asked Questions, and a selection of resources and tools to learn more about RDM. The website is regularly updated since its launch and new resources have since been added. The survey was designed, tested and promoted by five members of the RDM Working Group between the summer 2015 and winter 2015-2016. The exercise was primarily aimed at finding information about awareness, practices and needs related to RDM across all faculties. Analysis of the results helped assess what support was needed by researchers with regard to RDM, and how this should be prioritised by UCL Library Services and other research support services across the university.

In summer 2016, the Group worked on creating discipline-specific resources to help researchers throughout their research projects; such resources include RDM guidance, metadata standards, data repositories and ethics guidelines. A second completed project was the design of a course template to introduce research students to RDM. The template consists of a series of presentation slides, a lesson plan and guidance to deliver the course. The course was tested in autumn and winter 2016 by three members of the Group with cohorts of Masters and PhD students. It serves as an essential basis to develop future courses on RDM at a more advanced level and aimed at further communities across the university.

RDM Working Group members cite their primary reason for volunteering to take part as being the opportunity to extend their knowledge of RDM both to full personal interest, but also to provide extended research support to the departments with which they work. In the case of subject liaison librarians, a greater knowledge of RDM has been a means to establish new points of contact within the academic communities that they support.

20.1. MOTIVATION AND BACKGROUND

Modern research requires new types of specialists that are capable of supporting all stages of the research data lifecycle – from data production and input to data processing, storage, and the publishing and dissemination of scientific results, which can be jointly defined as key components of the emerging profession of Data Science (DSP).

To address this demand from research and industry, the Horizon 2020 Programme is funding the EDISON Project (Grant 675419, INFRAIUPP-4-2015: CSA), the goal of which is to build the Data Science profession for European research and industry. This includes the definition of Data Science and data handling-related professional profiles (or occupations), corresponding core competences and skills, the Data Science Body of Knowledge and a Model Curriculum that together comprise the EDISON Data Science Framework. This work is done with the involvement of the main stakeholders from the research community, industry, data preservation and handling community, universities and professional training organisations.

The University of Amsterdam is coordinator and a base organisation for the EDISON Project; other partners include the University of Stavanger (Norway), the University of Southampton (UK), Engineering Italy, EGI.eu, FTK (Germany), and Inmark Europe (Spain). The project benefits from multiple Data Science-related initiatives and academic activity and effective cooperation between Computer Science and multi-disciplinary departments, University Library and IT departments. It is also supported by such external initiatives as the Amsterdam Data Science Centre and Amsterdam School of Data Science (ASDS). On the other hand, all project recommendations find their practical pilot implementation at the University of Amsterdam and in cooperating organisations. This includes four Data Science and Big Data programmes, Research Data Management (RDM) training (together with the University Library), training for researchers, programmes and course catalogue services for universities and students, and advice for companies.

20.2. STAKEHOLDERS AND THEIR ROLE IN DATA SCIENCE EDUCATION

To create a foundation for the sustainable education and training of future Data Science professionals and Core Data Experts to support present and future data-driven research, the EDISON Project involves and is cooperating with multiple stakeholders, relevant bodies and communities. This includes but is not limited to the following:

- **Academic and research departments** are key for developing and teaching educational courses on Data Science and Research Data Management: four different Data Science programmes have started in the 2016-17 academic year, targeting different demand sectors in research and industry.