



Workshop Report: Skills for Supporting Research Data Management

On 29 June 2016 the LIBER Steering Committee for Scholarly Communication & Research Infrastructures¹ held a workshop² focusing on how libraries provide training on Research Data Management (RDM).

Libraries across Europe are rolling out research support services, including for RDM. This involves a range of activities, for example: establishing a data policy, support for its implementation, providing advice regarding data management planning and storage.

This year's workshop focused on transferring skills and knowledge through training. Typical target groups are young researchers, project coordinators and library staff. Main questions addressed by the workshop included if libraries feel ready to teach RDM, what target groups and topics are addressed, with whom to collaborate, what benefits for trainers are provided, and how this can be linked to data management services provided by libraries.

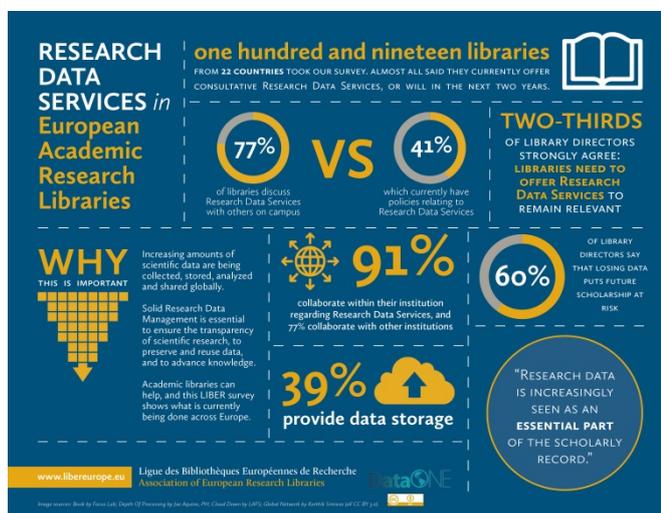
Wolfram Horstmann, State and University Library Göttingen, opened the workshop and emphasized that skilling for RDM is essential to one of LIBER's core strategic objectives: Enabling Open Science.

Carol Tenopir, University of Tennessee, and Wolfram Horstmann presented findings of a recent study on Research Data Services (RDS) in libraries³. The study was commissioned by LIBER and builds on a similar study conducted for the Association of Research Libraries (ARL)⁴ and DataONE⁵. A forthcoming paper summarises the outcomes of the study and will be published by LIBER Quarterly⁶.

Key Findings

Overall, LIBER has about 400 members, with 15% located in the North, 16% East, 17% South, and 53% West regions. The survey responses were representative for these regions. Key findings of the study are:

1. LIBER members are very much involved in RDS but consultative services are most common, fewer libraries offer technical services.
2. There are some regional differences.



Infographic created LIBER, October 2016.

¹ <http://libereurope.eu/strategy/strategic-direction-1-enable-open-science>

² <http://www.slideshare.net/libereurope/presentations>

³ <http://libereurope.eu/blog/2016/10/13/research-data-services-europes-academic-research-libraries>

⁴ <http://www.arl.org>

⁵ <https://www.dataone.org>

⁶ <https://www.liberquarterly.eu>

3. Libraries use a variety of strategies for developing staff skills.

Other key figures from the survey included the following:

- RDS are already offered by most libraries: over 3/4 of all libraries are discussing RDS with others and 2/3 are involved in policy development and planning.
- Half of libraries are involved in training colleagues on RDS.
- Nearly half (46%) indicated that they offer consulting on data management plans, 44% offer consulting on data and metadata standards, and 43% work with others for outreach.
- Of all respondents, 38% provide technical support for RDS. Reference support for finding/citing data is offered by 37%. In addition, 35% have created web guides, 32% directly participate with researchers, and 26% offer support in identifying data.

There were some regional differences. For example:

- Training colleagues on RDS and discussing RDS with others is slightly more common in the West (60-65%) and North (86-90%), compared to the East (29-36%) and South (46-50%).
- Identifying datasets seems to be well-established in the South (50%) and the West (33%), while emerging in the other regions.

Libraries use a variety of strategies for developing staff skills. When asked if they have hired staff for RDS in the last 12 months, 31% responded to the affirmative. Of all respondents, 84% offer opportunities for developing skills. This includes attending conferences and workshops (78%), to take courses (60%), to join working groups (59%), in-house staff workshops (47%), and collaborate with others (7%). Some have collaborated with others to develop a skills-related MOOC, moved staff from other support unit (IT services), hired staff for projects (not permanent), refocused an existing role, etc.

Pre-Workshop Survey

Birgit Schmidt reported on the outcomes of a Pre-workshop Survey (developed with Rob Grim, co-chair of the working group).

Although there were only 15 responses, several useful comments and pointers to useful tools were collected which helped to prepare the workshop. For example, when asked for success stories, the following approaches were mentioned: mandatory PhD courses on responsible research conduct includes module on RDM; a basic RDM training for graduate students is very easy to deliver, trainings related to scientific production and IPR are usually highly valued, and a national training of multipliers involving librarians, IT staff, research support staff.



Foto: Linda Tammisto

RDM Training Study

Jens Nieschulze, University of Göttingen, presented results of the Knowledge Exchange's Study on RDM Training⁷. A survey was sent out via mailing lists, libraries and third-party funded projects (1/3 of contacted projects responded).

In addition, a workshop with infrastructure providers was used to validate the findings and to collect further information. Trainers primarily are curators/librarians (55%) and researchers

⁷ http://repository.jisc.ac.uk/6379/16/Training_for_RDM_-_Comparative_european_approaches_May_2016.pdf

(40%), while only a few from IT services are involved in training for RDM. Matching the indicated target groups, trainees are mainly PhD students (90%) and postdocs (70%). Core objectives of the training are to change practices and to raise awareness. However, there seems to be some discrepancy between trainer's expectations and feelings expressed by those being trained. Areas for improvement include the applied approaches to presentation. In addition, more focus on tools and concrete problems seems advisable. In general, researchers tend not to perceive a need for training but respondents observed a trend for more demand. Therefore, it is important to make clear that RDM is relevant and important: offer carrots not sticks, and focus on the researcher, based on bottom-up real-life examples and exercises. Regarding curricula and methods the study highlights that a registry of information on training resources is needed. Knowledge Exchange is currently exploring the possibilities.

RDM Training and Education in Denmark

Jesper Boserup and Filip Kruse, (on behalf of Katrine Hofmann Gasser), State and University Library Aarhus, presented the Training and education on RDM in Denmark. The Data Management in Practice (DMIP) project involves 7 universities, national archives, and the national library. It targets RDM use cases and themes. One use data management use case is the LARM archive⁸. It covers radio data and enables researchers to enrich the metadata. The focus is on Humanities but also large quantitative studies (e.g. coverage of national elections). Training activities in relation to cases are, for example, to work with them in creating a DMP. In addition, training for researchers and library staff is offered, based on the Essentials 4 Data Support⁹, and some Danish examples have been added to the curriculum. The experiences so far reveal that services must be developed in very close cooperation with researchers; otherwise the services will not be used. It is essential to meet researchers earlier, at PhD schools and other agencies helping researchers with applications and funding.

Open Science Training in Finland

Sari Räisänen, National Library of Finland, reported about the strategy and activities in the context to Finland's Open Science Training programme. The initiative coordinates the training programme for the whole country. Services which are referred to in trainings include data repositories, discipline-specific repositories, e.g. language sciences, and services to visualize and process data. Main steps of the programme are to provide Open Science expert training, to facilitate an expert network, a web portal for sharing training materials, an online course (in Finnish), and an Open Science course for doctoral programs (in English).

Expert training has been organised since autumn 2015, to support academic community and to adhere to OA policies and institutions to conduct training. Topics include:

- Why Open Science?
- How to plan an open study.
- Open data, open publishing and processes.

The training is conducted mainly through examples (e.g. speeches given by researchers, and panel discussions, workshops, hands on training to use a data management planning tool). Fourteen seminars with more than 2,000 participants, have so far been offered and materials are available on the web portal. Currently, the initiative is planning the development of an OS course for doctoral programmes which will be carried out by the Finnish doctoral programme network in close collaboration with the Open Science and Research initiative.

⁸ <https://larm.sites.ku.dk/about-larm>

⁹ <http://datasupport.researchdata.nl/en>

Some challenges have to be addressed, e.g. libraries are more familiar with publications than with RDM, and legal aspects will be of utmost importance in the next few years. The EU Data Protection Directive will come into effect in 2018, which then has to be realized on the national level.

RDM Training for PhD Students

Monica Allardt, Susanna Nykyri, Katri Larmo and Liisa Siipilehto, University of Helsinki, presented their take on RDM training for PhD students: three approaches at U Helsinki. The courses target all subject fields, and have the following aims:

- To motivate researchers;
- To communicate the benefits of managing research data;
- To share key policies.

In some faculties the training is already very popular. The Medical Faculty, for example, holds 6 courses per year. A course typically starts with a pre-task on a learning platform (Moodle), including two motivational videos, a questionnaire on the student's data, and other material. This is followed by a two-hour hands-on session which covers data management at each stage of the research data life cycle, examples from the medical field and how to select a license. One course assignment asks students to present and discuss their research work and data. A lot of effort goes into encouraging questions.

This approach has been selected, because it fits very well to the limited resources available, and participants can come back with questions any time. As a consequence, the researchers get to know the library's research data team; they know where to find help. Course participants have been very satisfied with this approach. In the Life Sciences, Humanities and Social Sciences the aim is to let them write a DMP for their own work. Pre-exercises based on MANTRA and research guides are provided and they receive feedback about their answers. They write a DMP and get support in using the Tuuli tool. Overall, setting up the trainings involves a lot of work but this way of teaching is very inspiring and mostly positive feedback was received.

RESEARCH DATA MANAGEMENT COURSE GOALS

- to get **motivated** and **see the benefits** of managing research data
- getting to know some of the **key policies** and requirements
- to get an **overview: planning (e.g. DMP Tuuli)**, collecting, managing, citing, sharing and preserving research data
- meeting other PhD students tackling with the same kind of issues; **peer experiences** and solutions
- finding out where to get **help**

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Hybrid RDM Training

Robin Rice, University of Edinburgh, presented on Delivering hybrid RDM training using MANTRA¹⁰ at the University of Edinburgh. Conceptually, the courses are part of a training matrix.

MANTRA is a self-paced course, openly licensed, one can choose different topics. Overall, it is more text-based than the MOOC on research data management which was developed with the University of North Carolina Chapel Hill¹¹.

In addition face-to-face sessions are offered for researchers, announced in staff meetings (with a 5 minutes pitch) and then a follow-up session for an individual. Skills workshops are targeting mostly postgraduates, together with other information skills. Librarians may also find the DIY training kit useful, which is offered together with MANTRA. Other courses cover good practice in RDM and are offered 2-3 times per semester, based on the flipped classroom idea. Each group looks at a different module, comes back with Dos and Don'ts.

¹⁰ <http://datalib.edina.ac.uk/mantra/libtraining.html>

¹¹ <https://www.coursera.org/learn/data-management>

Courses also cover how to create a DMP for grant applications, working with personal and sensitive research data, handling data using the statistic software SPSS, etc. The training links up to policies and research data services: while the responsibility is on the side of researchers the university backs this up with service. All three stages of the process are covered, from planning, working with data to data stewardship. New developments are data carpentry and software carpentry¹² workshops (training fundamental data/software skills) for researchers which will be offered from August 2016.

Group Discussions

Although some of the participants were not yet involved in training, several are seeking collaboration and funding (e.g. Finland). Some have brand new funder expectations to consider (e.g. Switzerland) and learn with researchers in pilot mode (e.g. organising round tables).

One question is how to bring the knowledge to liaison librarians. It has been phrased as 'the new reference service' (Canada) and institutions have to define their levels of services – which echoes results from the LIBER study, as described above. Two group discussions addressed questions related to RDM training:

1. What topics are most valuable in training? What are the benefits for trainers?

The discussion highlighted that researchers need credit and they care about their data.

Training Tips: Try to find active researchers and to match library roles with researchers. Work from a policy. Focus on staff and research. Embed as part of open research. Have academic promotion hindered if they do not comply. Set clear goals for the training. Collect feedback / evaluation (pre and after). Consider covering legal aspects. Explain how to create, review. Use data management plans (also for trainers).

Benefits for trainers are:

- Inspiration to take on a new area and to get the feeling that you are needed (this goes to the whole library)
- Feedback helps to improve training
- Increased knowledge and competencies
- A new career path: if you create new roles you also need to create new positions.

2. How do you link training to RDM services?

Comments and recommendations from the discussion:

- Training should be closely related to researchers. A good understanding of the researcher is needed. Role plays can be applied and there is some analogy with medical teams.
- Build up and learn from projects (like H2020).
- Cooperate with faculty for building up the training. Be there at the right time. But also consider how this scales across the entire university.
- Practical training should be based on existing services. Engage with everyday research activities, reassure the researcher (they keep the control), show the benefits for RDM. Senior researchers are role models; add pressure with funders, journals, etc.
- A different solution is needed for PhD students. It's key to communicate the idea of RDM and related services, and that they can contact the library and come as an individual when they need it. Another solution is direct consultancy for proposals and related DMPs rather than training.

¹² <http://www.datacarpentry.org>, <http://swcarpentry.github.io/website>

- Trainings can be delivered on request vs. scheduled, customised by discipline and school. This requires marketing and to identify needs but the training would only be delivered on request.
- Linking of training and RDS comes automatically in marketing, tell researchers and together with IT and research services. However, there is some concern that we do not know how many will actually request the services and how labor-intense it will be.
- Do silos scale? To what extent do we need to accommodate disciplinary peculiarities?

3. What methods & tools do you consider most useful for delivering institutional training?

Methods-wise, Face-to-face and online training in combination, workshops both general and focused on one theme, mixed methods, mixed target groups (for researchers, research support units, PhD students and supervisors, etc.), all needs to be aligned.

Other ideas and thoughts: Develop a “RDM techniques for dummies”, develop and test training formats and modules, a registry/one-stop-shop for trainers/trainees, re-use teaching material, don’t keep reinventing the wheel. Use of learning platforms, recombine existing materials but also create own materials as this is more motivating. Cooperate to get away from 1:1 solutions based on incoming requests (nationally and internationally).

Tools which are considered useful are the modules from MANTRA, the Coursera RDMS MOOC, YouTube videos and the DMP tool.¹³ In particular, the DMP tool is mutually good for trainers and researchers. In terms of methods for credit courses are considered most valuable, combining online and classroom teaching (flipped classroom), with a focus on discussion in the meetings. In addition, it was recommended to tie the courses in with information literacy, being inclusive in terms of topics. Include 10 minutes of RDM in Masters teaching, bespoke what it means.

4. What train-the-trainer topics would be most beneficial for you?

Streamlining is needed, how to cope with little staff. Training topics include: planning as key topic, finding useful data, keeping control, real-life experience, legal aspects (e.g. privacy, security), reviewing DMPs, metadata.

Further practical tips for organising training:

- Keep terminology as simple as possible (e.g. what does data carpentry mean?). Consider a simple registry of terms.
- When would be a good time to start training? Possibly already at undergraduate level.
- The legal issue is expected to become very relevant. The library should act as a ‘middle man’, not as police!

5. How could LIBER help to strengthen the role of libraries in this activity area?

- More in depth practical solutions and services which researchers use in different disciplines, conversion tools.
- Showcasing of most important services in different fields of science.
- The external pressure is coming from funders and others anyway. Libraries have to anticipate that they have a role and responsibility.
- Collect relevant training materials.
- Continue the advocacy work and facilitating the dialogue between stakeholders.

¹³ See the links above and <https://dmp.cdlib.org>