Ten recommendations for libraries to get started with research data management

Final report of the LIBER working group on E-Science / Research Data Management

Birte Christensen-Dalsgaard et al
4-7-2012

Introduction

LIBER installed the 'E-Science working group' in 2010 to investigate the role libraries can and should play in the field of E-Science. The group decided to focus on research data as it was felt to be the most urgent element of e-science that is of relevance to the community of (research) libraries. The group has held three workshops, the first during the LIBER-conference 2011 in Barcelona, the second during the IDCC 2011 conference in Bristol and the third and last one during the LIBER-conference 2012 in Tartu. The results of the first two workshops were used as a basis for compiling recommendations to the LIBER-community. The "10 recommendations for libraries to support research data management" (see side bar) were finalized and prioritized during the final workshop at the LIBER-conference in Tartu.

Summary of the 1st workshop

At the workshop in Barcelona, the focus was on collaboration and partnership. Two essential observations were made:

- most participants saw a potential for new services in research data
- few participants had practical experience

The workshop featured four contributions on current projects and infrastructures, which offered different but complementary views on the role libraries can play when 'dealing with data'.

Following the presentations a number of issues were addressed:

- It is important to engage with researchers, and libraries may be the institutional liaison.
- Incentives for researchers need to be created through proper attribution and citation.
- The issue of ownership to the data needs to be addressed.

Ten recommendations for libraries to get started with research data management

1. Offer research data management support, including data management plans for grant applications, intellectual property rights advice and information materials. Assist faculty with data management plans and the integration of data management into the curriculum.
2. Engage in the development of metadata and data standards and provide metadata services for research data.
3. Create Data Librarian posts and develop professional staff skills for data librarianship.
4. Actively participate in institutional research data policy development, including resource plans. Encourage and adopt open data policies where appropriate in the research data life cycle.
5. Liaise and partner with researchers, research groups, data archives and data centers to foster an interoperable infrastructure for data access, discovery and data sharing.
6. Support the lifecycle for research data by providing services for storage, discovery and permanent access.
7. Promote research data citation by applying persistent identifiers to research data.
8. Provide an institutional Data Catalogue or Data Repository, depending on available infrastructure.
9. Get involved in subject specific data management practice.
10. Offer or mediate secure storage for dynamic and static research data in co-operation with institutional IT units and/or seek exploitation of appropriate cloud services.

- The funding issue for data management and data planning is a real concern.
- It is not going to be 'one size fits all'; different libraries will take different roles. No library probably needs to provide all data services and certainly not at once.
- Libraries need to ensure, that they remain relevant. Services around research data are a new area where the expertise of the library can be put to good use.

A general comment was made: when exploring new services, make room for experiments/pilots, keep what works and learn from your mistakes!

Summary of the 2nd workshop

The workshop in Bristol had a different, more technical, scope. The theme was 'Prerequisites for re-using data'. This workshop also started with four presentations but provided some more time for discussions. At the end of the workshop a plenary wrap up was held and the most important and agreeable issues were identified:

- Introduce a point/star system for the quality of datasets based on the associated metadata (content, context, etc.), openness and adherence to standards.
- Libraries should start supporting researchers in using standardized methods – through advice and actual support with data management plans.
- Libraries can play an important role in the social infrastructure for research data. Libraries can contribute to the use and re-use of research data, teaching facilities for research data and promoting open data access.

Summary of the 3d and final workshop

The working group drafted recommendations at an internal meeting during the Bielefeld Conference 2012 and presented these at the final workshop during the LIBER-conference 2012 in Tartu. Their validity was discussed with the audience in a session that started off with an 'Oxford style debate'. An overwhelming majority of the participants held the opinion that research libraries *can* and *should* play a role in supporting researchers with data management and data planning. Few participants thought that this was better left to research groups and data centers and that this was no place for libraries to step into, if only for the reason that they were not up to it. The participants agreed on a few changes in the wording and the emphasis of the recommendations:

- Libraries should be careful not to place too much emphasis on 'open' (as in 'open data policies') because especially in the early stages of a research project most researchers do not want their data to be 'open', certainly not without their consent and outside their control.
- Libraries should not only liaise with data archives and data centers, but should focus specifically on local collaborations with research groups and even individual researchers where appropriate.
- The recommendation to co-operate with IT units was found to be a bit too restrictive. This recommendation should be expanded to allow for instance the usage of cloud services.

The working group next asked the participants to prioritize the list of recommendations through a simple voting session. Together with the comments mentioned above, this 'wisdom of a small crowd' resulted in the *list of prioritized recommendations* to the LIBER community presented in the side bar on page 1 of this report.

Finally the working group asked the participants of the Tartu workshop what is was that they would need to be able to begin, or make further progress, with supporting research data management. From the lively discussion these common points were abstracted:

- There is no need for research libraries to start with *all* recommendations or to try to deliver a full spectrum of data services at once. Small steps will do. Learn from others (successes as well as mistakes), copy and adopt best practices. Already there is a lot of experience within research libraries worldwide.
- The *reskilling* issue is very important since very few libraries nowadays have the luxury of being able to hire new, specialized staff. Data librarianship probably is a profession in itself but a lot of work regarding data services can also be done by e.g. (reskilled) information specialists.
- Scepticism with researchers or governing boards towards the role libraries can or should play regarding research data management can be overcome by making very clear how good data management practices directly benefit researchers (and hence the institution) and also by pointing to recent fraud scandals that would not have taken place or would have been much less severe if good data management practices had been installed.

The Working Group

The LIBER working group on E-science was led by Birte Christensen-Dalsgaard of the Danish Royal Library and has had the following 'core members' (in alphabetical order):

Marc van den Berg	Director of Library and IT Services, Tilburg University
Birte Christensen-Dalsgaard (chair)	Deputy Director General and Head of Information Technology Services, Danish National Library
Rob Grim	Head of Research Support and Data Librarian, Tilburg University
Wolfram Horstmann	Associate Director, Digital Library Programmes and IT, Bodleian Libraries, University of Oxford
Dafne Jansen	Project Manager R&D, Utrecht University Library
Tom Pollard	(former) Datasets Outreach Officer, The British Library
Annikki Roos	Library Director, University of Helsinki

The working group would like to especially thank Stefan Kramer, Abteilungsleiter Forschungsinfrastruktur Deutsches Institut für Wirtschaftsforschung, Berlin, for his contributions to the final draft of the recommendations.