

Learned Publishing, 27: S33–S34
doi:10.1087/20140506

POINT OF VIEW

Rounding up the data: libraries pushing new frontiers

Susan K. REILLY

LIBER, the Association of European Research Libraries

© Susan K. Reilly 2014

In March 2103 *Nature* magazine published a piece hailing libraries as ‘the new data wranglers’,¹ racing to reinvent themselves in order to keep up with changes in scholarly communications. It is true that in the last few years new opportunities have arisen for libraries to reinvent or ‘re-engineer’² themselves in light of the ‘data deluge’ and the increasing level of recognition by funders and policymakers of the value of open data. Libraries have reacted quickly to these opportunities, developing research data management support services, repurposing institutional repositories to accommodate heterogeneous long-tail datasets, collaborating to foster interoperability, contributing to the development of standards for metadata and citation, working with data centres to encourage researchers to deposit their data, as well as supporting the discovery and reuse of data. At a point when so much change is occurring it is worthwhile looking back at the drivers for libraries to increase their role in research data management, assessing the landscape, and looking forward to identify new frontiers for libraries in supporting the publication of data.

Towards the end of 2011, LIBER, conducted a survey³ for the Opportunities for Data Exchange project, with the aim of exploring the evolving role of libraries in support of research data sharing. The results of the survey presented a compelling case for libraries to step up their activities in the provision of research data management services: over 80% of the libraries reported demand for such services, and massive gaps between demand and supply were to

be found in the provision of support services for data management plans and archiving. At around the same time the LIBER e-Science Working Group published *Ten Recommendations for Libraries to Get Started in Research Data Management*.⁴ Almost two years on, how have libraries gotten started in research data management?

It is clear that funder mandates have been a major driver for the establishment of data management support services in libraries. In the US the National Science Foundation and several other federal funding agencies now require the submission of data management plans. In Europe a major driver for the establishment of data management support services is the Horizon 2020 (H2020) Open Data Pilot, which will require the submission of data management plans that outline how data will be made available. However, something that may have more impact on libraries over the long term is recognition by institutions of the value of making data open, for example as outlined by the Royal Society’s ‘Science as an Open Enterprise’.⁵ The League of European Research Universities, in

their Roadmap for Research Data,⁶ adopt a position of ‘open by default’ and advise universities not just to develop a policy for research data management, but also to create an institutional roadmap for research data to set out objectives, actions, and tasks. It is these developments at institutional level that create the space for libraries to develop and expand their data management support services.

Libraries cannot provide data management solutions on their own. Many libraries have collaborated with relevant institutional departments such as research support and IT services to establish research data management services and institutional data centres. University College London, the University of Oxford, Humbolt University, and the University of Helsinki are all examples of libraries that have taken this approach. Their role in these relationships is often to provide guidance and training, as well as support for curation of the data throughout its lifecycle.

Libraries are also building relationships and collaborating with data cen-

tres, for example in the Netherlands a partnership has emerged between the Data Archiving and Networked Service (DANS) and university libraries to roll out what they refer to as a front-office/back-office model in which the libraries help researchers deposit their data in the data centres. The benefit to the data centre is that the libraries are better positioned to work closely with researchers to prepare and deposit their data in the archive, whilst the library need not offer huge amounts of data storage.

Also, by collaborating and working together on issues such as interoperability and minimum metadata standards, and connecting their institutional repositories to shared infrastructure such as OpenAire in Europe and SHARE (SHared Access Research Ecosystem) in the US, libraries are helping to surface long-tail data, link data to publications, and build shared tools for bibliometrics and the exploitation of data.

Now that libraries are establishing research data management services and partnerships, what are the frontiers that still remain to be pushed?

It would seem that a lot of work remains to be done around advocacy for open data and research data management. A European project, Policy RECommendations for Open Access to Research Data in Europe (RECODE⁷), investigating policy recommendations for open access to research data from the perspectives of different disciplines, has found huge disparities across the disciplines in attitudes towards the value of making data open.⁷ Some libraries may feel that the push towards making data open is best left to funders and institutional leadership. However, advocating the benefits and helping to embed best practice for making data open into research workflows for researchers on the ground may well be a more successful approach in the long term than funder mandates.

The current model or paradigm,

which rewards researchers on the basis of their final publication, is also inhibiting data sharing. Rather than discussing and implementing embargo periods for data, a more constructive approach to incentivizing data sharing is to devise means for researchers to be recognized for their data. It does not make sense that a researcher receives more recognition for producing a single paper than for producing a dataset that becomes the foundation for numerous papers. Supporting data as publications, helping to make them findable, accessible, reusable, and citable is one way that libraries can help to shift this paradigm.

Just as there is scope for libraries to become partners in the process of making data reusable, through support for data management plans, helping researchers to provide adequate metadata and documentation for data peer review, and ensuring that data is available over the long term, there is scope for libraries actively to promote and facilitate the reuse of data. Data registries and discovery services will help to make the data easier to find but libraries can also become more active partners in data reuse.

Facilitating access, manipulation, visualization, and reuse of data is an exciting new frontier for libraries. Already several libraries in the US have repurposed part of their library to become e-scholarship laboratories, e.g. Collaborative University Research & Visualization Environment (CURVE) in Georgia State University Library,⁸ providing access to software for the manipulation of data and immersive displays for data visualization and analysis. Such activity works on two levels: (1) to showcase the rich data being produced by institutions and (2) to bring together researchers, libraries, and technologists to share expertise and take a collaborative approach to accessing and visualizing data as

well as to data creation and, the final frontier, data reuse.

References

1. Monastersky, R. 2013. Publishing frontiers: the library reboot. *Nature News*, 27 March 2013. <http://www.nature.com/news/publishing-frontiers-the-library-reboot-1.12664>
2. Lyon, L. 2012. The informatics transform: re-engineering libraries for the data decade. *International Journal of Digital Curation* 7: 126–138, <http://ucsclibrary.pbworks.com/w/file/fetch/67752593/220-923-1-PB.pdf>
3. Kotarski, R., Reilly, S., Schrimpf, S., Smit, E., and Walshe, K. 2012. Report on best practices for citability of data and on evolving roles in scholarly communication. <http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/08/ODE-ReportBestPracticesCitabilityDataEvolvingRolesScholarlyCommunication.pdf>
4. Liber 2012. *Ten Recommendations for Libraries to Get Started with Research Data Management*. <http://libereurope.eu/news/ten-recommendations-for-libraries-to-get-started-with-research-data-management/>
5. Royal Society. 2012. *Science as an Open Enterprise*. Royal Society, London, <https://royalsociety.org/policy/projects/science-public-enterprise/Report/>
6. LERU Research Data Working Group. 2013. *LERU roadmap for Research Data*. http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf
7. Recode, Stakeholder Values and Ecosystems: http://recodeproject.eu/wp-content/uploads/2013/10/RECODE_D1-Stakeholder-values-and-ecosystems_Sept2013.pdf
8. CURVE Design & Planning: <http://research.library.gsu.edu/curve>

Susan K. Reilly

LIBER, the Association of European Research Libraries

Email: susan.reilly@kb.nl



Susan K. Reilly