

ALIA CONFERENCE: FOR MENTOR REVIEW

ABSTRACT TITLE: Research Data: key to the transformation to the next generation of Academic Libraries and Librarians

Author/s & Organisation/s

1. Cynthia Love, CSIRO
2. Karen Visser, Australian National Data Service

Presenter/s:

1. Name: Cynthia Love, Executive Manager, CSIRO Information Services
Postal address: CSIRO, Gate 5, Normanby Rd, Clayton, Victoria, 3168 Australia
Email address: cynthia.love@csiro.au
Phone: +61 3 9545 2908
2. Name: Karen Visser, Program Leader, Skills Policy and Resources, ANDS
Postal address: WK Hancock Building, ANU, Acton, ACT 0200 Australia
Email address: karen.visser@ands.org.au
Phone: +61 2 6125 7121

CONTENT STREAM: Academic Libraries, Future Roles for Academic Libraries

INTRODUCTION

Q: What is almost always open access, an institutional asset and a necessity in the collections and service offerings for every academic library? A: Research data. Research data products enable academic Libraries to continue to offer their staff and students relevant scholarly products for the new worlds of academic endeavour. There are also critical new opportunities afforded to those Libraries who engage their staff in the discovery and publication of their institution research data assets, as well as augmenting their collections with research data collections sourced elsewhere. Academic Libraries and Librarians are adept at not only recognising the value of emergent product opportunities but also adapting well honed practices to embrace new scholarly products.

Changes in research imperatives have impacted on this area. Research data outputs provide academic Libraries with the ability to value-add existing connections between journal and monograph products and their publishers with this new information format. Data has moved from being a by-product of research to being a first class output and there is a growing culture of data citation and providing data to support publication claims. Reproducibility, which is increasingly critical for the publication and accountability of research, requires connections between publications, data, software and other research artefacts. This connectivity also supports citation likelihood and assists research impact. Research has become increasingly collaborative, which means that researchers' demand for discovery of both publications and data has also increased. Researchers are distinguishing less between publications and data as research outputs. To them they are all outputs of their research and potential inputs into other research. Funding organisations are increasingly requiring the publication of data as a condition of the grant and this means that it needs to be properly managed and widely available.

Data is an institutional asset –and in a competitive environment publication of data is becoming a further demonstration of research impact and evidence of an organisation's contribution to addressing national and international research challenges. The inclusion of

data in collections with descriptions that facilitate syndication to promote access and subsequent re-use, is critical to the curation and effective presentation of the organisation's data assets. Combine this with the fact that the library services and collections model needs to respond to the new "DIY resource discovery" world and much more sophisticated and complex publication options – and this new academic world now includes data.

SO, WHY LIBRARIANS AND RESEARCH DATA?

Including research data products within a Library's publishing, catalogues, training and reference services is an increasingly necessary extension of existing Library paradigms. Research data aligns well with current publication and citation models and is increasingly integral to article publication processes. High school students are coming to university 'digital and data savvy' and expect that resources which support their tertiary education experiences will include products far beyond digital books and online journal articles. Resources for tertiary teaching and learning are a matrix of scholarly and generalist publications, software, visualisations, raw and processed data, interactive graphs, diagrams, embedded videos and more. Well crafted research data collections will include all of these products. The burgeoning open access movement has created a vast new world of free resources which are essential to the business model for academic libraries.

Libraries and Librarians now have critical opportunities to partner with researchers in ways never before possible and to change the ways in which, and with whom, Librarians work. Successful partnerships for research data publication include representatives from the academic community, IT, Research Office and the Library. Data Librarians are at the forefront of turning research data into institutional assets. Their ingest work includes the discovery, description and publishing of research data collections in tandem with the research team who created this product. At the other end of this product development is Librarians working with staff and students to correctly find, cite and reuse research data.

DATA VALUE-ADD FUNCTIONS FOR LIBRARIES

Specifically there are areas of activity where libraries can play a high value-add role:

- Support discovery – recognising that researchers are distinguishing less between publications, data, software and other research outputs, means there is a need to increase the breadth of information discovery to include the information inputs into research. Similarly, companies like Ex Libris are now syndicating the data collection records in Research Data Australia into their federated search tools. Additionally institutions are now making the link between publications in their publications repository and records in their data repositories and vice versa. CSIRO is also accommodating the publication of software in its Data Access Portal and librarians have been working with researchers, IT and legal staff developing a schema for software description, DOIs and appropriate licensing.
- Support data management – with expertise in description, information structures and retrieval, Librarians are perfectly placed to work with IT and researchers to develop supporting infrastructure and provide assistance with description and data management planning. Researchers recognise this and are calling for greater involvement from librarians in the management of their research data. That call has been heard by a number of organisations. Recognising the value of public sector data, there are now imperatives to make it freely available and data.gov.au is forging ahead. This represents an opportunity for government libraries to get involved and expand their service and influence.

- Support diversified Collection management - there is a rapid growth in the volume of data being generated and the capacity to store it, but it needs to be managed. There is a potential for the environment to be overwhelmed by floods of unmanaged data with little to no discoverability. Describing data is key to managing it in order to “weed” the data collection effectively and identify valuable data for retention and publication. Managing data is a matter of curation – it is part of the evolution in collection development. Describing and showcasing major collections and working to support feeds of metadata to improve context and increase access points is another area of expertise that the librarian can bring to this endeavour. Organisations are working with subject discipline areas to ensure that data flows are supporting national and international initiatives. The Australian National Data Service (ANDS) has worked with over 90 Australian research institutions and their librarians and data managers to bring together in excess of 100,000 Australian research data collection records. In CSIRO librarians are working with researchers and IT on the Australian Square Kilometre Array Pathfinder project defining collections from the data that will be streaming from instruments. This will allow manageable groups of data to be used and allow meaningful citation to be assigned. However, the challenge is to retain the context in the overall major collection.
- Support institutional publication – “Metadata has value for data users, data developers, and organizations. No dataset should be considered complete without accompanying metadata. Data without metadata is useless.” (1) Librarians play a key role in understanding the links between metadata and discovery, reuse and citation. Often, Libraries are responsible for maintaining institutional (data and print) repositories and work closely with IT to link to storage and compute. Other roles can include: assistance with strengthening link between publications and data, and assigning DOIs to publications (and increasingly to data) to ensuring that all citations are part of bibliometric reporting.

So where do I start?

- Search for resources across the new data horizon. Librarians are more than adept at developing literature reviews and finding information sources. There is a global network of discipline and institutional data repositories, packed with data freely available. The ANDS website contains hundreds of links, articles and infographics to help build a knowledge of tools and collection building options for Librarians.
- Tap into Professional development and networks. ANDS is partnering with CAUL and ALIA, to develop professional courses in the immediate term. In the longer term, data management should be incorporated into the training of librarians. ANDS also runs a number of series of webinars and workshops as well as providing information to support the development of data librarians.
- Talk to your colleagues. There is now a core of librarians who are actively working in this area and who have developed expertise. Communities of practice are being established and there are active conversations on social media such as Twitter #DataLibrarian. There is a real opportunity here for ALIA to support this growing community of practice.
- Audit your environment for data opportunities. What and where are the data outputs of your organisation? What opportunities are there for the Library to be involved in the publication and dissemination of data your organisation? Can data be infused into existing services and business models, or are new services required? What opportunities can develop by partnering with IT, researchers and data owners? How do your services and potential services map against the research lifecycle?

The last decade has seen Libraries take full advantage of the opportunities afforded through the digital revolution – online has meant new and expanded business models for libraries with Librarians being at the forefront of digital literacies. Adding Research data products is a logical and necessary extension to the evolving set of skills necessary for effective information services.

CONCLUSION

The role of libraries in research has always been critical in the publication and discovery of research outputs. With the rise of data as a first class output of research and its link to literature, it is natural that research libraries now have a role in the management and publication of data. Libraries touch on many points of the research lifecycle and the role of the library professional is becoming increasingly sophisticated as the information landscape is widening. Our clients are demanding ways to connect their information schemas – they don't just want literature but they need software, data, workflows and more. 'Would you like data with that?' is the new Library reality and there are huge opportunities for the brave to embrace the future and add data management to professional development.

Footnotes

1. US Geological Survey (2014) *The Value of Metadata* U.S. Department of the Interior. Page accessed 8 July 2014.
http://www.usgs.gov/core_science_systems/csas/metadata/value_of_metadata.html

References

The Australian National Data Service. Page accessed 1 July 2014. <http://ands.org.au/>

ANDS (2013) *Information specialists and data librarians*. Page accessed 1 July 2014
<http://ands.org.au/guides/dmframework/dmskills-information.html>

Schmidt, Heinz (2014) *Architecting extreme computing and data services for global research platforms* CSIRO CSS TCP & eResearch Annual Conference, Melbourne, 25-28 March 2014

Stodden, Victoria (2014) *Reproducibility in Computational Science: Opportunities and Challenges* CSIRO CSS TCP & eResearch Annual Conference, Melbourne, 25-28 March 2014

US Geological Survey (2014) *The Value of Metadata* U.S. Department of the Interior. Page accessed 8 July 2014.

http://www.usgs.gov/core_science_systems/csas/metadata/value_of_metadata.html

eResearch South Australia (2014) *Infographic: Understanding Metadata*. Page accessed 3 July 2014 <http://www.ersa.edu.au/enewsjune2014/understanding-metadata>