Developing new skills for research support librarians

Rebecca Brown, Malcolm Wolski & Joanna Richardson

Division of Information Services, Griffith University, Australia

*Corresponding author

Acknowledgement: None
ARTICLE
Developing new skills for research support librarians

Abstract
In recent years there has been considerable discussion about the key role which university libraries can play by engaging with their research community. As a result libraries are scoping, developing and implementing new roles and service models, especially in the relatively new area of research data. This article explores the specific challenges experienced by a traditional academic librarian at Griffith University as she moved into a new role as a data librarian. It was found that this transition needed to be underpinned by a skills development program, a mentor/coach and a support network of specialists. The authors then outline some strategies to facilitate this type of role transition, which include investing in a range of training and staff development activities, leveraging existing core librarian capabilities, and understanding the researcher perspective. The article concludes with a suggestion that several national organisations will continue to have an important role in supporting librarians as they develop new skills.

Keywords: Data librarian; research data services; research libraries; research data management; library roles

Implications for best practice
- While formal skills training is important as librarians move into new research support roles, there is also a critical need for informal training, mentoring and support networks.
- Library roles which support research need to be scoped to determine the skills and expertise required within a team, faculty and the institution.
- In-depth knowledge of the research process in specific discipline areas may be required to enable librarians to contribute as a full partner in the research activity.
- In Australia national bodies such as CAUL and ANDS will continue to have an important role to play in assisting libraries to provide support networks.

Introduction: Research data is the new gold
There is a new emphasis on re-using and preserving research data. This is a logical consequence of data having become more numerous, more complex, and more important. As the Australian National Data Service (ANDS 2011) observes, ‘…the research process is transforming to become more investigative as it is now possible to assemble significant data collections that enable much broader problems to be addressed. Thus it is critical that research data is managed, discoverable, and connected to enable innovative re-use’.

Funding bodies and national governments are seeking an improved return on investment for funded research. Along with the transformative nature of research, this has been a major driver for initiatives aimed at better access to and the sharing of research data, e.g. UK Data Archive. In addition, governments are providing access to data as part of their ‘open government’ strategy. The Australian Government, for example, in its Declaration of Open
Government is committed to ‘open government based on a culture of engagement, built on better access to and use of government held information, and sustained by the innovative use of technology’ (Tanner 2010). Most Australian State Governments have followed suit with similar strategies.

In Australia re-use of data is typically an existing priority in projects funded by National eResearch Collaboration Tools and Resources (NeCTAR) and ANDS, and is an emerging priority in funding agencies such as the Australian Research Council (ARC) and National Health and Medical Research Council (NHMRC). In addition, ARC applicants are expected to outline their ‘management of data’.

At the same time internationally, digital preservation organisations and services are beginning to evolve in recognition of the importance of preserving not only research data but also digital information more generally as part of its lifecycle. Walters and Skinner (2011) examine in detail some of the types of initiatives (business-driven, community-driven- and library-driven) which underpin these operations.

Role of university libraries in supporting research data

The recent focus on data has resulted in much discussion within universities, often leading to the development of new services. The library is one of the service providers within a university that is seen by many as having a key role in engaging with the research community, particularly in regard to the management of their data (Auckland 2012; Malenfant 2010; Jaguszezewski and Williams 2013). The traditional role of providing information support and training has been expanded to include support in all steps of the research lifecycle (Simon Fraser University Library 2012).

However, many university libraries are grappling with their emerging role in supporting the new area of research data. In a 2009 UK survey (Lewis 2010), the starting point was to question whether the specific activity of managing data was actually a role for university libraries. On the one hand, the answer was ‘no’ because the scale of the challenge in terms of infrastructure, skills and culture change requires concerted action by a range of stakeholders, of which university libraries are just one. On the other hand, the answer was ‘yes’ because ‘data from academic research projects represents an integral part of the global research knowledge base, and so managing it should be a natural extension of the university library’s current role in providing access to the published part of that knowledge base’ (p 146).

How far that current role should be extended has been the topic of much discussion. In 2011 the Association of College and Research Libraries (ACRL) surveyed a cross section of its members in the United States and Canada to provide a baseline assessment of the current state of, and future plans for, research data services (RDS) in academic libraries in these countries. In the resultant report (Tenopir et al, 2012, 3-4), two of the key findings were that:

- Only a small minority of academic libraries in the United States and Canada currently offer research data services (RDS), but a quarter to a third of all academic libraries are planning to offer some services within the next two years
- Libraries on campuses that receive NSF (National Science Foundation) funding are more likely to offer or plan to offer RDS of any type. This suggests that funding agency requirements are driving the need for RDS. As budget decisions move towards even greater accountability, it is likely that more agencies will dictate
responsible data management, so the need for RDS on campus is likely to grow. If the library is not actively involved in providing these services, some other unit is likely to be pressed into service, which can diminish the image of the library as an important partner in the research process.

Those university libraries which are adopting a broad perspective are currently beginning to undertake a range of research support activities, including

- Raising awareness of data issues within institutions and the benefits of actively managing research data
- Assisting in developing policies about data management and preservation
- Providing advice to researchers about data management early in the research life cycle; influencing the way researchers will be creating their data, the formats they will use and building a commitment to use a repository to publish/preserve their data
- Working with IT service colleagues to develop appropriate local data storage capacity
- Training and introducing data management and curation concepts to research students
- Exploring methods of moving data from work-in-progress storage spaces to repositories in more seamless ways

Walters and Skinner (2011, 11) have published a report which discusses ‘the emerging practice of digital curation for preservation and how research libraries are fostering curatorial practices in order to ensure that their parent institutions continue to realize their core mission of creating, disseminating, and preserving knowledge. MacColl (2010) reinforces the idea that a vision of a comprehensive and strategic role for libraries includes the curation and preservation of research outputs. Tenopir et al (2012, 41) point to ‘a more active and visible role in the knowledge creation process by placing librarians at all stages in the research planning process and by providing expertise to develop data management plans, identify appropriate data description, and create preservation strategies’.

In July 2012, LIBER’s E-Science working group (Christensen-Dalsgaard et al 2012) released ‘Ten recommendations for libraries to get started with research data management’. Since then, several libraries across Europe either have started to build or have expanded their capacities for research data management, typically combining e-infrastructure and support services. Their progress has been documented in 11 case studies, which describe policies and strategies that pave the way for the creation, institutional integration and the running of support services and underlying infrastructures. In addition, challenges and lessons learned are described, and ways-forward outlined.

Digital scholarship and the challenges associated with research data offer libraries the chance to shed their ‘support service’ label and become research collaborators (Corrall 2013). Sarah Thomas, Vice-President for the Harvard Library, reinforces the likelihood of this new role: ‘I see us moving up the food chain and being co-contributors to the creation of new knowledge’ (Monastersky, 2013, 431).

In a 2014 webinar, three senior information professionals discussed how their respective libraries were offering a growing number of services to support a diverse set of research needs, as both researchers and scholars increasingly move toward data-driven research (Mertens et al 2014). Stuart (2014) has explored the potential role for research libraries in a data-centric age. He highlights, for example, the importance of providing training to researchers in accessing, archiving, publishing and managing data.

In the examples outlined above, each library has had to determine its own approach to
supporting research, based on its respective strategic priorities and those of the parent organisation. For most libraries, defining the scope of the problem and then moving from a theoretical discussion to actually developing and implementing a practical response has not been an easy task since it impacts on all aspects of the library organisation and the staff. In the following sections the authors discuss their own experiences and suggest some strategies to facilitate the transition.

The challenge of putting theory into practice

In 2013 it became evident at the authors’ university that there was a high level mandate for change. This was a result of several key drivers: compliance with funding agency requirements, senior academic recognition of the potential value of data as an asset and the practical need to improve data management and an enterprise need for across the board data classification methods to help manage the risk of data loss. This resulted in changed policies and new guidelines.

At Griffith University the Division of Information Services integrates e-research, library, information and communication technology into a single organisation. A divisional restructure saw the formation of six portfolios, two of which encompassed library operations. The new Information Management portfolio took over governance responsibility of the traditional acquisitions and library systems function as well as corporate records and institutional repositories for scholarship including research outputs. Additionally this portfolio has responsibility for discoverability of content as well as the management role. New teams and staff roles were established to manage these operations. The new Library and Learning Services portfolio assumes the governance of faculty librarians and physical spaces. Faculty librarian roles were also modified to address data management.

At the same time investment in further development of systems and infrastructure to manage data underpinned the establishment of new services. Development of a discovery portal known as the ‘Research Hub’ and development of new repositories to accommodate new demands are ongoing (e.g. addressing the need to mint DOIs). A complementary program of work is underway to implement new services through the faculty librarians. The objective of this program of work is to further engage with researchers and assist researchers in the management of their data.

Similarly to what was noted in the LIBER E-Science working group study above, most of the Griffith staff involved saw a potential for new services in research data but few staff had any practical experience. Carlson has observed: ‘The challenges encountered by librarians seeking to engage in data management and curation issues are found at the individual level (acquiring skills and confidence) and at the organizational level (creating a supportive environment). Both levels will need to be addressed by libraries seeking to develop data services’ (2013, 17).

Carlson’s observations parallel the experiences of one of the authors of this paper. The author moved from a long-time role as an academic services librarian, working with a Science and Engineering Faculty, to a role as data librarian working on a federally funded climate change adaptation project which aimed to provide advice to Natural Resource Management bodies around Australia. The data management component involved providing data management advice to nine clusters of researchers from the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and academic institutions throughout the project lifecycle, and ensuring that project outputs would be accessible and
discoverable at the project’s completion. The author’s previous role had included collection development, advanced information retrieval training for higher degree students and academics, bibliometrics reporting, and providing high-level advice to research staff on data management planning.

The author considered herself as having a solid, broad-level understanding of data management, including the creation, use, description, storage, and retrieval of data. However the new role proved a substantial challenge. The author did find that having relevant discipline knowledge (in this case a formal background in Environmental Sciences, including experience as a research assistant) was a distinct advantage. Familiarity with scientific terminology and the scientific research process meant that she was comfortable working with climate change researchers from the start, and broadly understood the types of outputs they were producing.

Initially the author was required to provide written, best-practice advice on a suite of data management topics such as in-project data storage, copyright and licensing (including the complex legalities of extensive data re-use), data description and documentation, data identifiers, and long-term deposit for preservation and discoverability. While this proved to be an excellent, on-the-job introduction to many major concepts in data librarianship, becoming a confident ‘expert’ in a short space of time required the rapid uptake of new skills and knowledge.

The author spent considerable time doing background reading and completing a variety of informal training packages (for a list of self-training options and communities of practice for data librarians see Simons and Searle (2014, 12)) as well as participating in an 8 week, assessment-based MOOC on Metadata. Together these tools provided a strong theoretical overview of the field of data management; however the author had difficulty contextualising theory, and felt limited by a lack of both practical data management experience and in-depth understanding of the data on which she was providing advice.

As she wrote in a blog post following the early days of her new role, ‘For many academic librarians moving into the data management space, ‘data’ is just a word. How many academic librarians have seen a dataset recently? ….. If I could have seen a raw dataset or data collection prepared for sharing and reuse, versioned correctly, saved in an appropriate file format, licensed, assigned a DOI, described using an appropriate metadata schema, uploaded into a content management system, and made discoverable for reuse, then I think I could have saved many hours of reading and scratching my head…’ (http://www.samsearle.net/2014/06/reflections-on-path-to-data.html)

The author was also required to provide detailed advice to researchers on creating quality documentation (metadata) for their project outputs, with the aim of storing outputs in the Terra Nova Climate Change Adaptation Information Hub (https://terranova.org.au/). She quickly recognised that her previous career focus on information literacy and information retrieval meant she lacked expertise in cataloguing processes, including the use of authority files and standards, which would have been useful for the role. For example, Terra Nova incorporates a number of geospatial fields and vocabularies from the ANZLIC Metadata Profile (itself based on AS/NZS ISO 19115:2005 Geographic information – Metadata). Rapid upskilling in the use of standards in metadata creation was therefore required.

The author was involved in the ongoing development of the Terra Nova Hub. In spite of her extensive experience in information retrieval from an end-user perspective, her lack of
information technology skills (e.g. lack of technical experience with relational databases and content management systems, web development and design), initially limited the advice she could provide to developers on improving the user experience.

The author was assigned an experienced mentor from within the organisation, which proved critical during the induction period of the role. The mentor was able to point the author to a wealth of existing material to assist in writing best-practice advice, and helped to place new concepts into context by providing real-life examples of their application. The most valuable training material proved to be case studies which attempted to contextualise theory, for example completed data management plans, (such as those found at the Digital Curation Centre website, http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples), actual data curation profiles (such as those at http://datacurationprofiles.org/) and webinars hosted by the Australian National Data Service, which have provided an opportunity to hear speakers describe their involvement in research data service provision in detail (ANDS 2014).

The mentor proofread and edited all content written by the author during the first three months of the role. The mentor also helped the author to understand the scope of her position and provided reassurance about the depth of technical expertise required for the job.

When the author’s mentor moved to a new position within the Division of Information Services and was no longer able to provide mentoring, the author found herself working alone within a team of programmers, web developers, business analysts and project managers. At this point the author successfully turned to both local and online communities of practice external to the organization to seek advice and support, for example, the Australian National Data Service General Discussion Group.

The experience of this author, as outlined above, highlights some of the challenges to a successful transition from traditional academic librarian to data librarian, and points to the types and formats of skills development which could ease the journey.

Discussion
In the particular case outlined above, the leap from theory into practice has been a large one and has needed to be underpinned by a skills development program, a mentor/coach and a support network of specialists.

Training and staff development
When designing training, one needs to consider both the background of the librarian and the details of the expanded role they are expected to play. Auckland outlines the wide variety of data management roles in which an academic librarian may be involved: from a largely advisory role (e.g. providing advice and referral on within-project data management, long-term preservation of research outputs and compliance with policy and funding mandates) to a hands-on role (e.g. applying advanced skills in developing metadata schema specific to disciplinary standards and individual research projects) (2012, 3).

The data management roles of academic librarians, along with the associated required knowledge and skills, can be represented as a continuum of increasing complexity:
<table>
<thead>
<tr>
<th>Advisory role, e.g. academic librarian</th>
<th>eResearch role, e.g. data librarian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of the research process and scholarly communication, including an overview of discipline-based knowledge and outputs</td>
<td>Advanced understanding of discipline-based research process, outputs and scholarly communication, including an understanding of data types and formats typical of specific disciplines</td>
</tr>
<tr>
<td>Knowledge of legal and regulatory frameworks</td>
<td>Advanced understanding of ethics, intellectual property, copyright and licensing</td>
</tr>
<tr>
<td>Overview of good information and data management practices, e.g. safe storage, backup and long-term deposit of data</td>
<td>Advanced understanding of safe storage, backup, and transfer of data, including file formats, version control, file authenticity and security</td>
</tr>
<tr>
<td>Overview of metadata concepts and schemas</td>
<td>Advanced knowledge of discipline-specific metadata schemas and related standards, at both item-level and collection level</td>
</tr>
<tr>
<td>Understanding of mark-up languages such as XML, interoperability and crosswalks.</td>
<td>Understanding of mark-up languages such as XML, interoperability and crosswalks.</td>
</tr>
<tr>
<td>Overview of preservation standards</td>
<td>Knowledge of repository certification schemes and standards</td>
</tr>
<tr>
<td>Overview of semantic web and open data</td>
<td>Knowledge of semantic web standards, open data platforms</td>
</tr>
<tr>
<td>Communication and outreach skills</td>
<td>High level communication and documentation skills, project management skills, systems design skills, business analysis skills.</td>
</tr>
</tbody>
</table>

Table 1. Increasing complexity of librarian roles supporting research data

There also needs to be a continual program of in-service training to provide ongoing skills development. Some specific skills needed include negotiation, advocacy and communication. Simons and Searle (2014, p 10) have identified three broad training pathways for librarians moving into the data management space: formal (tertiary) education, training courses (in-house or externally provided), and informal learning, either self-directed or supervisor/peer-assisted. Searle (2014) has specifically outlined the benefits of introducing a component of scenario-based learning into introductory research data management workshops for librarians. In addition she provides practical advice as to how to develop scenarios and integrate them within an institutional staff development program.

This has highlighted the need for a good support network of specialists to form a virtual team. Such a network enables the data librarian to acquire specific knowledge from domain experts on the job. It also allows the data librarian to call in a specialist to work on complex problems beyond their level of expertise. Some recent Australian examples include:

1. Secondment of subject librarians to work in specialist teams for a short period to build skills (La Trobe University) (Huggard et al 2014)
2. Creation of virtual teams including library staff to develop services for researchers (University of South Australia) (Healey et al 2014)
3. Secondment of librarians to work on projects to address data management activities (Griffith University) (Wolski and Richardson 2014)

The common thread is that librarians have been placed in teams with other specialists. A key finding from these examples has been the importance of librarians being able to tap into other librarians’ experiences. As a result national organizations such as ANDS and CAUL have facilitated such activities through workshops and webinars. However there is still an unmet need to further develop support network librarians working in domain specific areas e.g. bioscience, ecology.

Alex Ball (2013, 10) from the Digital Curation Centre (DCC)/UKOLN Informatics has recently stated: ‘I am confident we will see more specialists with scientific, technical, engineering and medical backgrounds being brought into the library profession to deal with specialist data issues’.

While many data management practices are cross-disciplinary, there are discipline areas in which domain knowledge would be a distinct advantage. Complex datasets resulting from scientific instruments, modelling results, and geographic information system (GIS) layers, for example, need accurate and extensive description to be reusable. This level of description would require a librarian to have, or to develop, some expert knowledge.

Therefore it is important within the library that any librarian roles which support research need to be scoped to determine the specific skills and expertise required within the team, faculty and institution.

**Leveraging existing core capabilities**

An aspirational role for librarians was affirmed at the May 2014 ARL membership meeting (Association of Research Libraries 2014) which described the research library and university of 2033 as ‘a rich and diverse learning/research ecosystem’ (p 16), with the research library shifting from ‘its role as knowledge service provider within the university to become a collaborative partner that catalyzes evolution’ (p 17).

As discussed previously, research support librarians generally have to acquire new skills to function effectively in their newly expanded roles. In terms of partnering with researchers, however, they can also draw on those existing core capabilities they already have by virtue of having been trained as a librarian. O’Brien and Richardson (2015) have discussed the core capabilities that position librarians well to be partners in the process of research. These include structured thinking, knowledge of information management theory, ability to communicate, understanding of knowledge dissemination, awareness of trends, etc. It is their ability to utilise both existing capabilities and newly-acquired skills which helps to establish the librarian as a core member of a research support team.

**Researcher perspective**

Whenever you develop a service you need to understand the perspective of the intended audience. This is a fundamental to a good communication strategy. In the case of specifically supporting data management, librarians may encounter resistance. For example, a challenge is a lack of buy-in from the data owners / creators themselves. Many of the steps in data management are labour intensive and require time and money, which have often not been factored into grant-based research.

One of the questions often asked by researchers is ‘why should I spend my precious time and money doing this?’ ‘So others can build on your research’ may ring alarm bells for researchers working in a highly competitive and financially constrained research environment. ‘So others can cite your data and your citation rate will improve’ is likely to elicit a request for statistically significant proof that data sharing is increasing citation rates.

More recently the compliance flag is having some success. ‘Because your funding
body/publisher says you have to’ is more likely to bring the researcher to the table. But at this point the research librarian must present a complete package of tools and procedures to make data management as seamless and pain-free as possible for the researcher. That is, he or she must ‘know their stuff’, which is a challenge in such a new profession or for those libraries developing a new service.

In Australia and internationally there has been much discussion about how to improve engagement and support within the institutions. For example, a recent survey by Research Data Alliance Europe (2014) has made recommendations about how to improve research practice. It should not be assumed that all researchers do not exercise good practice in managing their data; some may do so by using a variety of readily available tools and technologies. The challenge is how to bring about a wider change in behaviour so that all researchers understand the importance of continually reassessing their current practices and, if necessary, adopting new practices. The issue of behavioural change within universities is emerging as a critical factor in responding to the rapidly changing research practice (Yanosky 2009; O’Reilly et al 2012; Andreoli-Versbach and Mueller-Langer 2014; Wolski and Richardson 2015). Therefore the success of assisting researchers in managing data for example depends on both the upskilling of library support staff and the willingness of researchers to engage in the process.

Conclusion
University libraries are now seen as having a key role in engaging with their research community. As a result traditional roles providing information support and training have been expanded to include support in all aspects of the research lifecycle. Libraries are having to determine their approach to supporting research in their respective institutions in response to their own strategic priorities as well as those of the parent organisation. The challenge is to not only scope the changes required but also to develop and implement an effective support model.

In this paper the authors have discussed their own experiences specifically in transitioning a traditional academic librarian to a new role as a data librarian. Several key findings have emerged from addressing training needs and working directly with researchers. Firstly, while formal skills training is important as librarians move into new research support roles, there is also a critical need for informal training, mentoring and support networks. Secondly, library roles which support research need to be scoped to determine the skills and expertise required within a team, faculty and the institution. This is because not only will all support librarians not have the same roles but also there is a need to have expertise in some areas, e.g. information technology, standards, project management. Thirdly, there may be a need to have in-depth knowledge of the research process in specific discipline areas to be able to contribute as a full partner in the research activity. For example, how an ecologist, in contrast with a bioclinician, finds and collects data, and then processes and analyses it through to publication.

Whatever approach a library takes, there will be opportunities for libraries to respond to a rapidly changing environment through collaboration, e.g. especially in providing support networks. In Australia national bodies such as CAUL and ANDS will continue to have an important role to play in this respect.

Note
1. An earlier version of this paper was presented at the VALA Conference, Melbourne, 7 February 2014. The substantially revised paper published in this issue of the Australian Library Journal has been double-blind peer reviewed to meet the Department of Education’s HERDC requirements.
References


Simon Fraser University Library. 2012 “Research Commons: Research Lifecycle for Graduate Researchers”. Last modified August 24 http://www.lib.sfu.ca/research-commons/research/research-lifecycle


