8-27-2019

Whose Responsibility is it Anyway?

Kristine N. Stewart
Zayed University, kristine.stewart@zu.ac.ae

Judith Mavodza
Zayed University

Submitted on: 02.07.2019

Whose Responsibility is it Anyway?

Kristine N. Stewart
Library & Learning Commons, Zayed University, Abu Dhabi, United Arab Emirates.
E-mail address: Kristine.Stewart@zu.ac.ae

Judith Mavodza
Library & Learning Commons, Zayed University, Abu Dhabi, United Arab Emirates.
E-mail address: Judith.Mavodza@zu.ac.ae

Abstract:

The evolution of scholarly research raises questions regarding the role of research libraries in the 21st century. As information and data have taken on new forms, researchers and libraries alike have adapted their skills and services to reflect changes in how information and research are created and conducted, disseminated, and preserved – throughout shifting social and philosophical paradigms as well as in response to emerging technologies. As such, librarianship is an ever-changing field that has advanced to include data management skills as a core competency. Unfortunately, perceptions of the LIS field have not kept up with the pace of its development. Involvement of librarians in data capture and management remains a struggle because those producing data in universities may not necessarily associate their activities with the library – unless there is a system in place that makes it mandatory for them to use a library or other repository in place, e.g. grant funding or promotion requirements.

This calls for information specialists such as data and academic librarians to intervene and provide guidance in numerous areas such as: information management, classification, and basic data literacy skills. The tendency in academe to avoid librarians in the research process is a missed opportunity for many researchers but also requires that librarians step up and make their voices and potentialities be known.

Keywords: Data librarianship, Academic librarianship, Information management, Research data management, Scholarly research.

Introduction

Research libraries are confronted with a requirement to make research data as readily available and useable as scholarly articles. This refers to both in-house data and data from external sources. Data continues to accumulate with increasing scholarly research and has some scholars referring to it as "small data". Pollock (2013) defines small data as "the amount of data you can conveniently store and process on a single machine, and in particular, a high-end laptop or

server." Those tens of thousands of small-data producers in aggregate may well produce as much data (or more, measured in bytes) as big data. As Wang, Xu, Chen, and Chen (2016, p. 2) point out, reality now dictates the library collect a lot of "small research data", created by individual researchers. Kinney (2018) reports that an overabundance of small data sets is a fundamental challenge.

The management of research data has become a challenge for many research and educational institutions due to the quantity of born-digital data produced rapidly in a variety of forms (Pinfield, Cox, & Smith, 2014). In most cases, data management is assumed to be the responsibility of the researcher who has created or collected the data. But then, the reality is that the research process involves many people whose participation impacts the quality of data and safeguarding them. The participation of many players creates the initial stages of facilitating data sharing but must expand and be thoughtfully cultivated to include stakeholders who will have a vital role in an institution's Research Development Program (RDM). Pinfield et al. (2014, p. 8), found in their study of RDMs at UK universities, components of an institutions RDM will vary based on an institution needs and specifications, but will (in general) be composed of the following components:

- Strategies that include a vision with and outline of goals to guide RDM activities;
- Policies that specify procedures and how to cover emergent issues, such as intellectual property rights;
- Guidelines that provide details of how policies will be implemented, specific RDM activities, stakeholders, and roles and responsibilities of stakeholders;
- Specifications of processes to regulate activities with the data life-cycle that assist in activities such as RDM for specific projects, data processing, data systems, data preservation, and the use of standards and procedures;
- Technologies to facilitate the technical components required for data repositories and networking infrastructures for data storage and retrieval;
- Services that support research data life-cycle activities (e.g. creation of data management plans, provision of training, and support services).

An integral part of the aforementioned RDM components is the designation of guidelines to make this process organized. This includes outlining and assigning (rather than presuming) the roles and responsibilities of stakeholders and involving stakeholders from a diverse selection of departments and backgrounds with relevant expertise. This is especially critical due to the collaborative nature and requirements of both research and RDM programs.

The UK Data Service (2019) lists the following as potential stakeholders involved in data management and sharing:

- the project director designing and overseeing the research;
- research staff designing research, collecting, processing and analyzing data;
- laboratory or technical staff generating metadata and documentation;
- a database designer;
- external contractors involved in data collection, data entry, transcribing, processing or analysis;
- support staff managing and administering research and research funding, providing ethical review and assessing Intellectual Property rights;
- institutional IT services staff providing data storage, security and backup services;
- external data centers or web archives that facilitate data sharing.

Training is mentioned as important. But then, the role of librarians is not mentioned in an obvious manner. Librarians may be presumed to be some of the specialists who may coordinate or provide training in various aspects of research data management, but often need to express a need, interest, or jurisdiction to be involved with RDM at an institutional level.

The implication is that librarian training and continuous professional development, and re-training, must move with the demands of the market. This is not just so librarians can be involved in RDM and other institutional initiatives, but so libraries are prepared for rapid changes that occur in the landscape of higher education and may lead the way in developing sustainable approaches to changes.

Perceptions of the Library's Role

Despite the many changes in higher education, one of the largest challenges faced by libraries remains how they are perceived within higher education. Libraries have evolved to reflect changes in modern research practices that impact how research is conducted, disseminated, and preserved. Unfortunately, the literature indicates that perceptions of LIS have not evolved at the same rate (Benton, 2009; Kroll & Forsman, 2010; Brown & Tucker, 2013; Hollister & Schroeder, 2015; Gabbay & Shonam, 2017). Much of academe appears to maintain the perception expressed by Benton (2009) that:

> Professors and librarians are socialized into different professions with different values that can make us mutually incomprehensible: one emphasizes individual scholarly productivity; the other looks to provide the context in which that work can take place. The two professions are also separated institutionally and, increasingly, there are fewer opportunities for collegial exchange between faculty members and librarians.

This insight was ten odd years ago, but the perception has hardly changed. Although much of the expertise involved in RDM necessitates knowledge about the processes involved in manipulating, storing, retrieving, and classifying recorded information, the library is not always a stakeholder which springs to the minds of researchers and administration. Some academics are unaware or occasionally discouraged from collecting, sharing and preserving their research data, clearly reflecting a disconnect between what the library can do for them and the role of the library in a research environment.

Advances in technology have made vast quantities of information accessible in digital format, enabling researchers to access information remotely. While this has greatly increased access to information, it has led to confusion regarding the source of information and the extent to which researchers *believe* they are using the library. Remote access has reduced how frequently researchers physically visit the library but may be accessing more information from the library using university-provided proxy servers and external sources such as Google Scholar to retrieve information from library databases (Gabbay & Shoham, 2017).

This confusion of access has created a disconnect between those producing and using data in universities and institutions and the services, resources, and expertise provided by libraries. A by-product of this disconnect is that the library's involvement with RDM may struggle because they are no longer seen as instrumental to the research process. In reality, librarians are even more involved in the research process than ever before but work "behind the scenes" developing digital collections and making them accessible through the creation of metadata and cataloguing, system maintenance, and data literacy instruction. Avoidance of librarians in

the research process is a missed opportunity for researchers and a reminder for librarians that they need to be the face of digital scholarship and data management at their institutions.

Active involvement of libraries within RDM is often dependent on where the library is positioned within a university or institution's scholarly communication system. When libraries are part of an open system that encourages the collection, management, and sharing of scholarly works through institutional repositories or other platforms to manage the digital scholarship produced at an institution, this stimulates engagement with the library and provides a built-in leadership role for libraries in RDM (Jain, Bentley, & Oladiran, 2009). This is why it is so crucial that the relationship between research activities (and not only information seeking and retrieval) and the library is made explicit – through marketing of not only what the library can provide researchers in terms of resources, services, and instruction, but how *librarians* contribute to research, scholarship, and research practices.

Some universities and institutions require researchers share their work (e.g. article pre-prints, conference proceedings, data sets, and theses and dissertations) through institutional repositories to gain access to grant funding or promotion. This assists institutions with preservation of their research outputs, while also increasing their visibility and research prestige. In terms of the benefits to libraries, this provides additional opportunities to engage with researchers on broader issues of modern research practices and helps redefine the work of libraries in their minds.

Changing Perceptions or Assuming Responsibility?

Changes in the landscapes of information and higher education have challenged libraries to become more flexible and responsive to remain relevant (White, 2017). Librarians recognize their part in research data management, and LIS curricula reflects that. They are also involved in continuous professional development, and creating strategic alliances within their parent institutions, as well as enhancing researcher support efforts. Assuming responsibility is no longer a choice because it is relevant to all research areas and draws from the larger Big Data picture.

Due to the collaborative requirements of RDM, there are several factors that influence an institutions approach to RDM and the library's involvement. Pinfield et al.'s research (2014) indicated these factors include: acceptance and prioritization of RDM; cultures of various professions and academic disciplines involved; creation of demand or interest in RDM; incentives for participation (e.g. promotion); stakeholder acceptance of assigned roles; governance, institutional power dynamics; funding and staff resources; abilities and gaps in skills; advocacy; and the setting and context of an institution. From this, a complex picture is revealed. These challenges, however, provide numerous opportunities for libraries and librarians to demonstrate leadership and develop or re-develop narratives of the library that place elements of RDM within the jurisdiction of the library.

Assuming responsibility for data management requires a commitment to develop the capabilities of libraries as well as an understanding of how much librarians can feasibly do. While libraries are vested in the management of research data, a DRM program requires ongoing support and funding for staff and additional resources. The nature of RDM demands collaborative work and thus presents opportunities for collaborations with researchers and academic departments, IT departments, administration, research support centers, and other support services (Pinfield, 2014; White, 2017). These collaborations are crucial for many

reasons, one of which is that it increases knowledge of DRM, which will potentially increase buy-in from researchers, creating demand, and opportunities for its development.

As noted by Cox and Verbaan (2018), there are compelling reasons for researchers to share their data openly, but many have hesitations stemming from concerns that data analyses will be duplicated or that standards or quality of work will be degraded if research is free and openly available to access and use. While the latter may be a misguided concern, based on a lack of understanding of the publishing process and changes that have occurred that now frequently allow for researchers to upload pre-publication copies of articles to repositories.

As a central activity for many universities and institutions, research is a billion-dollar business. Sharing data and research allows for more research and different approaches to data sets, additional citations for publications and data sets increases the visibility of data and research, opens opportunities for collaboration, and assists in the ability to reproduce studies, thus ensuring the integrity of research findings. Data sharing also has great benefits in terms of providing access to research that has been publicly funded, compliance with funding and publication regulations and requirements (Cox & Verbaan, 2018).

Research only has value if it is something that can be accessed for further use. As such, the importance and value of RDM lies in the ability of potential users to find and access research and research data (Cox & Verbaan, 2018). In terms of finding research and research data, metadata and documentation that describes data is vital to any institution that aims to preserve and share research and data. Depending on the research, source of funding, and its discipline will affect its description and sometimes even the metadata schema used. Although there are many reasons why libraries and librarians have an integral role in DRM, metadata and description are some of the most significant as they ensure the "discoverability" of research and data, crucial tools in the discovery of research and contributors to an institution's visibility and research profile.

Another area of vital importance to accessing research and research data is the cultivation of abilities and skills needed to retrieve this information. Information literacy instruction has long been the work of librarians and information professionals. The concept of data literacy has evolved from information literacy to foster an understanding of the nature of research data, legal and ethical aspects of data collection and management, as well as knowledge of data practices, such as analysis, visualization, documentation, sharing, preservation, and citation (Cox & Verbaan, 2018). Delivery of data literacy instruction and training of researchers, students, and other stakeholders will vary based on institutional structures, but are often done by a combination of librarians, IT departments, or researchers themselves. Again, demonstrating the need for collaboration in RDM programs.

Although libraries and librarians should and have assumed responsibility for RDM, it should not and cannot rest on their shoulders alone. RDM programs are organically collaborative, requiring interdisciplinary skills and knowledge for their management (White, 2017). RDM projects are labor intensive, costly, and require ongoing maintenance. The benefits of RDM are vast and range from assisting in the preservation of research outputs to increasing the visibility and prestige of institutions. However, the benefits of RDM are only viable if researchers participate in these programs, share their research, provide feedback based on their use of platforms such as institutional repositories, and acknowledge the benefits of their participation.

References

Association of College & Research Libraries. (2019). *Scholarly communication toolkit: Repositories.* Retrieved from https://acrl.libguides.com/scholcomm/toolkit/repositories

Benton, T.H. (2009). A laboratory of collaborative learning. *Chronicle of Higher Education, 55*(43), A31-A33.

Brown, J., & Tucker, C. (2013). Expanding library support of faculty research: Exploring readiness. *Portal: Libraries and the Academy, 13*(3), 283-299.

Cox, A., & Verbaan, E. (2018). *Exploring research data management*. London: Facet Publishing.

Gabbay, L.K., & Shoham, S. (2017). The role of academic libraries in research and teaching. *Journal of Librarianship and Information Science, 00*(0), 1-16.

Hollister, C.V., & Schroeder, R. (2015). The impact of library support on education faculty research productivity: An exploratory study. *Behavioral & Social Sciences Librarian, 34*, 97-115.

Jain, P., Bentley, G., & Oladiran, T. (2009). *The role of institutional repository in digital scholarship communication.* Retrieved from https://www.researchgate.net/publication/237322259_The_Role_of_Institutional_Repository_in_Digital_Scholarly_Communications

Kinney, J. (2018). *The big problem of small data: A new approach*. Retrieved from https://www.cshl.edu/the-big-problem-of-small-data-a-new-approach/

Kroll, S., & Forsman, R. (2010). *A slice of research life: Information support for research in the United States.* Retrieved from https://www.oclc.org/content/dam/research/publications/library/2010/2010-15.pdf?urlm=162948

Pinfield, S., Cox, A.M., & Smith, J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLoS One, 9*(12), 1-28.

Pollock, R. (2013). *What do we mean by small data.* Retrieved from https://blog.okfn.org/2013/04/26/what-do-we-mean-by-small-data/

UK Data Service. (2019). *Roles and responsibilities.* Retrieved from https://www.ukdataservice.ac.uk/manage-data/plan/roles-and-responsibilities

Wang, C., Xu, S., Chen, L., & Chen, X. (2016). Exposing library data with big data technology: A review. In *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)* (pp. 1-6). IEEE.

White, W. (2017). Libraries and research: Five key themes for sustainable innovation in strategy and services. *New Review of Academic Librarianship, 23*(2), 85-88.