

Implementing Research Data Management Services in a Canadian Context

Abstract: Research data management (RDM) has become an increasingly pressing issue for academic librarians as they strive to assist researchers in addressing new public funding requirements surrounding data dissemination and preservation. Briney, Goben, and Zilinski (2015) reviewed several characteristics of RDM service provision efforts by 206 American research universities. Following a similar methodology, the author of this paper reviewed RDM service development within Canadian research universities and compared the results to the American efforts. This study finds that the main area requiring development in Canada is the provision of RDM services. Therefore, some current best practices for implementing RDM services were gathered through a literature review. The successful approaches highlighted in the literature include awareness of funder and institutional data policies, reaching out to data service providers on campus and beyond, understanding researcher data management needs and finding RDM champions, implementing research data services strategically, planning for growth in RDM services, marketing the RDM services, and creating incentives to create data management plans and utilize RDM services.

About the Author: Tess Grynock is a member of the 2016 graduating class of Dalhousie University's Master of Library and Information Studies program. She has a Bachelor of Science from the University of Alberta and spent a semester studying at the University of Canterbury in Christchurch, New Zealand. Tess' practicum at the Massachusetts Institute of Technology Libraries exposed her to research data management (RDM) services development and she is continuing to build her RDM skills through a project with the Dalhousie University Libraries.

Introduction

Academic libraries have long been expected to be repositories for the research papers and dissertations created by university scholars, and now they are adding local academic research data to their extensive collections as well. Data is a trending topic within academic libraries and has appeared in the form of data curation on the 2012 Association of College and Research Libraries' (ACRL) top trends list, and as an opportunity for researchers and academic librarians to collaborate in the 2014 list (Saunders, 2015). Despite data's effect on academic libraries, only 39.7% of the 63 American academic library strategic plans surveyed by Saunders (2015) included data services. Canadian academic libraries have further incentive to provide data curation services since researchers applying for funding from the tri-agencies—Canadian Institutes of Health Research (CIHR), Natural Sciences and Engineering Research Council of Canada (NSERC), and the Social Sciences and Humanities Research Council of Canada (SSHRC)—will soon be required to submit a data management plan that details how their data will be shared and preserved. Data management plans can also act as useful tools for researchers and ensure that data being deposited into an institutional repository has already gone through the initial steps of the curation process (Varvel & Shen, 2013).

Many American academic libraries developed their research data management (RDM) services more rapidly because they were on a tighter deadline to meet the onrushing requirements from public funders such as the National Science Foundation and National Institutes of Health, who implemented their data sharing policies in 2011 and 2003 respectively. Canadian academic institutions have the luxury of time before the majority of public funding bodies finalize their data management requirements. One exception is the CIHR, which has had an open access policy pertaining to research data since 2007. However, only bioinformatics, atomic, and molecular coordinate data is required to be made available on a public database, while all CIHR funded data has to be stored for a minimum of five years after project completion. Another country Canada can look to in terms of research data management is the United Kingdom (UK), which has become a leader in this field and has developed a national support system through the Digital Curation Centre established in 2004. The national support strategy adopted by the UK has been adopted by other countries, including Australia, and it is likely that Canada will follow a similar model—this has already commenced through the formation of the Portage initiative. With the luxury of time and being able to observe other academic libraries adopt research data services, Canadian academic libraries have the benefit of learning from the complications and accomplishments other libraries experience

Implementing Research Data Management Services in a Canadian Context

when implementing RDM services. Circumstances surrounding RDM, such as the requirements from the different funding bodies, are different in Canada than in the United States and other countries around the world. Therefore, any techniques used by other academic libraries around the globe must be adjusted to the Canadian context. So how are Canadian academic libraries faring with research data management (RDM) today? This paper reviews some successful approaches for developing RDM services gathered through a literature review, and provides an outlook of current research data management services in Canadian research universities.

Literature Review:

Successful Approaches to Implementing Research Data Management Services

Many academic institutions and funders have created policies that either deal primarily with research data or include research data as a larger component of research output. Institutional policies, such as the University of Alberta's (2010) Information Technology Security Policy, mandate a certain level of security on data, and other institutions also have policies that cover areas such as ownership and dissemination. Funders, such as the Bill & Melinda Gates Foundation (who have an open access policy which also covers data) also have their own data policies which may have similar or differing conditions regarding data ownership and dissemination than the institution. Knowledge of institutional and funder data policies not only helps academic libraries establish parameters for their RDM services, but also provides knowledge for librarians helping researchers navigating this landscape. In Canada, the tri-agencies will soon require a data management plan (DMP) to accompany funding requests, and require that the data be stored in a "publically accessible, secure, and curated repository" (Government of Canada, 2015). Canadian researchers also occasionally receive money from government funding agencies in the United States, or are part of a larger research group with funding from other countries. Certain journals such as *Nature*, *Public Library of Science (PLoS)*, and the *Journal of the Medical Library Association* also have data sharing policies (Briney, Goben, & Zilinski, 2015). Once a librarian understands the policies relevant to their institution, they may begin to guide researchers through the intricacies of institutional and funder data policies.

Some institutions do not have data policies. Briney et al. (2015) presents this scenario as an opportunity for librarians to help their institutions create a data policy and promote modification of data policies that are damaging to research. Since researchers must

also respect numerous non-institutional data policies, I would caution institutions from making too many more complications for researchers. A better understanding of what should go into the policy will also arise from practicing RDM.

Reaching out to data service providers on campus and beyond

Discovering other data service providers on campus not only prevents duplication of services, but could also lead to opportunities for collaboration. Occasionally, other data service providers, such as the technology support department and data centres on campus, require reassurance that libraries are not taking over their space (Henderson & Knott, 2015). To help reduce the potential hostility, Stephanie Wright from the University of Washington prefers to take groups that provide data services out for coffee instead of a meeting (Wright, Whitmire, Zilinski, & Minor, 2014). A conversation over coffee not only provides a relaxed informal setting, but is also easier to fit into people's busy schedules.

Once contact is made between academic librarians and other data service providers, this network can be leveraged to create a community of practice. Three examples of communities of practice are presented by Vitale, Marshall, and Nurnberger (2015), who are from Washington University in St. Louis, University of Wisconsin-Madison, and Columbia University respectively. Mannheimer (2014) describes a fourth community of practice at Montana State University called the Data Management Steering Committee. The communities created at these institutions take up various projects from creating educational programs and drafting data management plans to creating best practices, overcoming common challenges, and ensuring that the data services meet campus needs (Mannheimer, 2014; Vitale et al., 2015). The collaboration within communities of practice can also lead to the creation of new tools (Henderson, Raboin, Shorish, & Van Tuyl, 2014). One tool described by Henderson and Knott (2015) that arose from such collaboration is a new DMPTool template with GIS coordinates for the Rice Center for Environmental Life Sciences at Virginia Commonwealth University. The added tool allows the Center to display their different research locations on an interactive map. As with any interdisciplinary team, conflict will likely arise from defining terms for determining desired outcomes (Vitale et al., 2015). To overcome the conflict, communities of practice need a shared vision which includes clear, attainable objectives which are accepted by the individual stakeholders of the group (Vitale et al., 2015).

Relationships built with non-data-service units on campus and data service organizations off campus can also provide guidance for academic librarians when implementing data management services. Wright et al. (2014) emphasize the

importance of the relationship with the Research Office, since this is who provides information about funding sources for researchers, and could also direct researchers to the data management services provided by the library once they are established. Organizations such as the Inter-University Consortium for Political and Social Research (ICPSR) have been preserving data since 1962 (Hayslett, 2015). Libraries have the opportunity to learn from these organizations that have successfully structured research data for the past several decades (Doty, Herndon, Lyle, & Stephenson, 2014). As Hayslett (2015) so aptly states, “I encourage you to explore the landscape of our data world—you may be surprised at the resources available in your own backyard” (p. 4).

Understanding researcher data management needs and finding RDM champions

To better serve researchers, RDM librarians need to be “research-aware” (Newton, Miller, & Bracke, 2011). In other words, they need to understand what research is being performed at the university and what terms are associated with that field of research, particularly in relation to data. Traditional methods of assessing the RDM environment at an institution include sending out surveys and/or performing interviews. A common problem with these methods is that everyone, including librarians, faculty, and staff, are busy, so scheduling interviews or having a representative sample for a survey may be difficult. As an alternative solution, researcher data management needs can be extrapolated from the needs assessments completed at other universities. A recent example of a needs assessment survey was performed by Buys and Shaw (2015) at Northwestern University. Henderson and Knott (2015), who champion this method, highlight that research on information needs and information-seeking behaviour can also inform the RDM needs for different disciplines. Their research also lists the needs assessment sources that they used to inform their own RDM strategy, which range from studies highlighting researcher data management needs to older studies on information-seeking behaviour (Henderson & Knott, 2015). One site that may be of particular use to data librarians is *Purdue’s Data Curation Profiles Directory*, at <http://docs.lib.purdue.edu/dcp/>. Unfortunately, there are currently no Canadian examples, but Purdue provides a toolkit that would enable librarians to produce their own data curation profiles. If each library within the U15 Group of Canadian Research Universities created a data curation profile for a different discipline, they could create five data curation profiles for each of the major Canadian funding bodies—CIHR, NSERC, and SSHRC—thus providing examples of data management strategies to streamline the process.

One advantage of performing surveys and interviews is that they spread the word about upcoming data management services and they can help identify RDM champions, researchers who are eager for RDM services, and willing participants for pilot testing. The RDM champions may also be cultivated from faculty and researchers who have positive existing relationships with the library (Mclure, Level, Cranston, Oehlerts, & Culbertson, 2014; Toups & Hughes, 2013).

Implementing research data services strategically

“Launching new services requires forethought of design, business modeling, marketing, communication, personnel recruitment, and some level of system development within a sustainable framework.”

(Shen & Varvel, 2013)

The Association of European Research Libraries/Ligue des Bibliothèques Européennes de Recherche (LIBER) recommends having an institutional data repository and data services. However, they do not provide direction as to how the implementation should take place within their ten recommendations for libraries getting started with RDM (Christensen-Dalsgaard et al., 2012). RDM services must go through the same planning as any new service as described by Shen and Varvel (2013), but luckily many of these plans have been shared in the literature for libraries to adapt to their own institution. One of the main themes that arose from these plans was the need to start with a small number of RDM services, then look to expand those services (Toups & Hughes, 2013; Wright et al., 2014). Starting with a small number of services may seem problematic if there is a high need at the institution, but, just as with cataloguing gifts and donations, one must perform triage. The plan outlined by Henderson and Knott (2015) for the Virginia Commonwealth University triaged data services on a needs basis. They also caution that curation, such as appraisal and weeding, should perhaps not be the first plan of action (Henderson & Knott, 2015). Another model is presented by Knight (2015) from the London School of Hygiene & Tropical Medicine (LSHTM) who describes the school's three-tiered triage approach: priority is given to researchers performing funded research, questions are directed to existing expertise whenever possible, and each question submitted is recorded to help plan educational support and to reduce the time needed to answer frequently-asked questions. Similar question tracking systems are used by reference services to expedite response time. The objectives listed in a strategic plan can help inform the triage needed when providing initial services, and provide goals to work towards.

Piloting the RDM services is another technique for starting small and identifying problems with low risk to the entire project. Many different subsets of the target

population have been used to pilot RDM services. The pilot described by Knight (2015) targeted only research projects that had secured funding and did not require a data management plan in the funding process. Although the criterion meant running the risk of researchers not planning for the funding required for RDM, Knight (2015) explains that establishing research data management plans after acquiring funding allowed researchers more time to discuss the plan with their team. Since many universities are multi-disciplinary and deal with different forms of data, the pilot can focus on reflecting the diversity. The pilot described by Johnston (2014) only used five data sets, but from different disciplines, at the University of Minnesota. Piloting with interested users also ensures that the researchers are committed to the project and the pilot will produce useful results (Mclure et al., 2014; Newton et al., 2011). RDM developers can utilize the knowledge of department liaisons to help find potential subjects or leverage existing relationships within the departments (Newton et al., 2011).

Where students are required to submit electronic theses and dissertations, libraries can ask students to submit their data, or the institution may make submission mandatory (Creamer, 2015; Doty, Kowalski, Nash, & O'Riordan, 2015). The student subjects can be used to pilot data management plan services, and their research data can be used to pilot data-specific repositories. Students make excellent pilot subjects; they are more willing to accept new practices, since they are not yet comfortable with a standard practice, and the data management skills that they learn will be brought forward into the rest of their life's work (Choudhury, 2012; Creamer, 2015). Although piloting with student data sets has many benefits, Steve Van Tuyl from the Oregon State University reminds librarians that student data sets may also need the most work to curate (Creamer, 2015). Doty et al. (2015) also note that at their university, where data submission is not mandatory, the number of electronic theses and dissertations being submitted with supplemental data files has doubled in 2014. Although the reason for the jump is unclear, this shows that libraries need to allocate repository space and staff time in order to accommodate growth. If academic libraries are hesitant to pilot data repositories with researcher or student data, they could use the library's own data, such as electronic resource data (Ogier, Hall, Bailey, & Stovall, 2014). As an added benefit, a data audit can be performed at the same time as the pilot (Ogier et al., 2014).

Another method of implementing RDM services strategically is by using pre-existing skill sets among the library staff. The working document from the Joint Task Force on Librarians' Competencies in Support of E-Research and Scholarly Communication (2015) highlights three key competencies needed to support RDM: providing access to data, advocating and supporting data management, and managing data

collections. The survey of the U15 schools revealed a strong contingent of Canadian librarians who already help provide access to data. Other pre-existing skill sets within the library discipline align with these competencies, so even librarians without data management training may already have a basis to draw from (Cox & Corrall, 2013; Mclure et al., 2014). Reference service skills may be applied in the interviews and question development for data management plans (Huwe, 2013). Liaison librarians bring knowledge of research practices within their departments. Cataloguers and collection development staff can apply their knowledge to determine suitable metadata standards and the best use of identifiers. Digital archival skills also apply to electronic data sets, especially in terms of appraisal and curation as the data sets accumulate. Asking for the advice of digital archivists early on will help data repository managers develop strategies for assessing data's value and preserving digital objects. Often skills in reference, liaising, cataloguing, collection development, and digital archiving do not reside within a single person at an academic library, and therefore initial data services may be provided as a team to facilitate sharing of expertise (Davis & Cross, 2015; Varvel & Shen, 2013). For example, having librarians with different skill sets team up to help a researcher construct a data management plan means that they will be able to offer advice on a wider range of areas and provide a more diverse range of observations (Varvel & Shen, 2013).

Libraries also need to outline in their plan how they are going to keep track of progress and frequently asked questions to start building institutional knowledge. Knight (2015) highlights that tracking the number of support requests and staff time devoted to these requests during a pilot can help administrators estimate workload, determine problem areas, and reveal which departments are using the service and which are not. For librarians wishing to prepare themselves for the types of RDM support questions likely to be asked, Knight (2015) provides a good list including questions about data citation and institutional data policies. By the end of the trial phase of the services, the librarians should understand the capabilities and limitations of the system and the user (Newton et al., 2011).

The data management plan not only needs to address short term metrics, but also long term metrics such as data re-use and how much data should be maintained (Shen & Varvel, 2013). These questions are not easy to answer, but Shen and Varvel (2013) suggest some possible benchmarks such as the number of partnerships with external repositories. The institutional data repository is one element of RDM services that can be measured early on through an evaluation tool such as the Trusted Repositories Audit & Certification: Criteria and Checklist (TRAC) (Downs & Chen, 2010). The diversity of data present at many universities requires research data managers to

be hybrid data curators who are “open to and flexible about what to curate while maintaining scientific rigor during the curation process” (Kouper, Akers, & Lavin, 2013, p. 46). The hybridity adds complexity to an already complex task, but libraries and other memory institutions have been curating multi-discipline and multi-media collections for centuries and similar solutions could be drawn from previous obstacles in these areas.

Planning for growth in RDM services

Part of implementing research data services strategically is taking time to consider the possible growth of the RDM services provided by the library. How will the university library meet the data management needs of the different disciplines considering variations in vocabulary and data management knowledge between domains (Shen & Varvel, 2013), curate the data in the future, and create metadata for diverse data sets? Of course, any plans set before RDM services are implemented will change as institutions discover what methods work best in their respective situations.

When developing a data repository, an important element to consider is who should create the metadata (White, 2014). The three main options are professional-created metadata, which is consistent but expensive and takes time, author-created metadata, which may provide better descriptors, or automatic indexing (White, 2014). Researchers prefer descriptive metadata but usually do not want to create it themselves (Greenberg, White, Carrier, & Scherle, 2009). Luckily the three main options can be blended in a hybrid system to best meet the needs of the institution and discipline.

The final metadata created will not only ensure that the data is searchable but also accessible (Walker & Keenan, 2015). In the rapidly growing field of RDM, accessibility is easier to incorporate earlier rather than later. One example that Walker and Keenan (2015) provide is the offering of assistive technology for visually impaired students to navigate a table of data if they cannot view a pie chart. Although this area of RDM is still evolving, Walker and Keenan (2015) list a number of useful resources such as WebAIM (<http://webaim.org/>), Center for Universal Design (<http://www.ncsu.edu/ncsu/design/cud/>), Diagram Center (<http://diagramcenter.org/>), and the California State University Accessible Technology Initiative (<http://www.calstate.edu/accessibility/tutorials/math.shtml>).

Metadata creation and selection are the main components of traditional data curation, which happens after data is deposited. Both components are time consuming and complicated by the context-dependent nature of data. Alternative

forms of curation include community-based data curation and active curation (Giarlo, 2013; Gordon, Millman, Steiger, Adolph, & Gilmore, 2015). Community-based curation, as explained by Giarlo (2013), leverages crowd-sourcing to multiply traditional curation and has been used in projects such as Galaxy Zoo and Zooniverse. In active curation, librarians encourage researchers to use institutional data repositories to store their data throughout their research, which allows researchers to add the appropriate context and ensure the most current version is available (Gordon et al., 2015).

Marketing RDM services

Marketing RDM services was not addressed in the LIBER Working Group's recommendations, but plays an essential role in RDM service implementation (Christensen-Dalsgaard et al., 2012). Both Fox (2013) and Knight (2015) acknowledge that researchers often lack knowledge of the RDM services provided by their own institutions. A communication plan helps maintain a consistent message, and determines target audiences for marketing RDM services effectively (Henderson & Knott, 2015). Through marketing RDM services, Toups and Hughes (2013) have found that interpersonal engagement is more successful than impersonal marketing. Interpersonal engagement is already used in service-oriented libraries and can be seen in reference exchanges and encounters between liaison librarians and their departments. Due to the feedback Toups and Hughes (2013) received on their marketing, they changed the title of their services to "data support," which was more understandable and less ambiguous than data or digital curation. Researchers also prefer targeted information, and faculty liaisons can provide RDM developers with such relevant information (Mannheimer, 2014). Mannheimer (2014) also highlights the need for simplified information. One method of advertising that is suggested by Toups & Hughes (2013), and is already being used in Canada, is the creation of a RDM research guide. Examples of RDM research guides can be seen on the Dalhousie University Libraries website (<http://dal.ca.libguides.com/rdm>) and the Western Libraries website (<http://guides.lib.uwo.ca/rdm>).

As research data managers, libraries need to become open data promoters and spokespeople for advancing data management standards, skills, and education (Newton et al., 2011). Since data management plans are required for funding, librarians may find it easy to convince faculty and researchers to complete data management plans but may also find that the DMPs are viewed as an added burden to the application process as opposed to an asset to the faculty's and researchers' projects (Shen & Varvel, 2013). Read et al. (2015) addresses some of the biggest problems different disciplines have with data management that can be solved with

the creation of a data management plan. The biggest obstacle for basic science researchers is a perceived lack of data standards and procedures for uniform collection (Read et al., 2015). Temporary researchers, both post-docs and graduate students, pose another big problem for basic science researchers because they often take data with them and are difficult to reach when they move on to new research opportunities (Read et al., 2015). Clinical researchers are more concerned with the quality of data, caused by factors such as inconsistent data collection methods with multiple individuals collecting data (Read et al., 2015). Librarians can position data management plans as solutions to these concerns. Herold (2015) also suggests that data management plans that incorporate appropriate public data sharing measures have been proven to increase citation rates. In short, enlightening researchers and faculty to the benefits of data management plans will provide greater incentive to create and follow data management plans.

Although the LIBER Working Group does not include marketing in its recommendations, it does include education of data management (Christensen-Dalsgaard et al., 2012). Educating faculty, students, and library staff in RDM plays an important role in marketing RDM services. Bishoff and Johnston's (2015) study of National Science Foundation data management plans illustrates why education is needed. One troubling finding was the inclusion of the general public as an audience 60% of the time, but with only 40% actually having publically accessible data sharing methods (Bishoff & Johnston, 2015). Researchers often assumed journals and conferences were adequate means of public sharing when many of these dissemination forums are "pay-to-view" (Bishoff & Johnston, 2015). By educating faculty, students, and library staff, such assumptions may be corrected. Educating researchers on the discipline-specific repositories has also proven beneficial in reducing the initial inflow of data into an institutional data repository (Knuth, Johnson, & Hauser, 2015). Researchers with no previous knowledge of discipline-specific repositories may be more inclined to use them once they know about them and therefore reduce load stress on institutional repositories that are just starting out. Lists of discipline-specific repositories have already been created on sites such as Databib, DataCite, and the Registry of Research Data Repositories (re3data) (Hayslett, 2015). RDM education would not only help market the library's services but also correct data management assumptions and help moderate data deposition into the institutional repository.

Workshops on data transfer and storage are not the only possibilities for workshops. Knuth et al. (2015) lists the workshops provided at the University of Colorado that range from best practices in data management to learning the differences between specific data formats. Workshops can also extend beyond faculty and principal investigators.

Regina Raboin (2015) remarked that the best people to reach out to are the lab managers since they spend a lot of time alongside the students in the lab, and additionally, the students trust their expertise. Whether or not students are used for piloting an institution's data repository, they are still a valuable audience for which to provide data management instruction. Doty et al. (2015) notes that even though faculty want their students to gain data management skills, they do not want to spend time on data management when it delays the research process and especially when faculty do not feel knowledgeable enough to instruct students on the process. Not all students want or need to learn proper RDM techniques; therefore, targeting workshops towards graduate students and students enrolled in research programs or courses ensures an appropriate audience for RDM instruction (Toups & Hughes, 2013; Whitmire, 2015). Although Whitmire (2015) notes areas for improvement in her RDM course for graduate students, such as more real world examples, she provides one successful example of a term length course. At the writing of her article, she had heard mainly positive feedback and that the students were able to implement the skills they had learned (Whitmire, 2015). Librarians do not need to start their RDM workshops from scratch since there are a number of curriculums already available, such as the DataONE Data Management Education Modules, MANTRA from the University of Edinburgh, and the New England Collaborative Data Management Curriculum (NECDMC) (Eaker, 2014). Librarians should be forewarned that statistics collected by Knight (2015) indicate a spike in RDM support requests after information seminars, which indicates the success of education as a marketing strategy.

Faculty, students, and researchers are significant groups to provide RDM workshops for, but the most important group is the library staff. Providing library staff with enough information to understand research data management means they will know how to field questions from researchers and will be able to direct more complicated questions to the RDM Services team (Toups & Hughes, 2013). Being knowledgeable about RDM also means that librarians outside of the RDM Services team are able to promote RDM services in their regular interactions with faculty and students. Methods of instructing staff range from providing a whole day of instruction, such as *DataDay!* delivered by Johnson and Bresnahan (2015), to a time-sensitive deposit-a-thon described by Amy Hodge from Stanford University, where a number of librarians gathered to eat pizza and deposit data sets into the institutional repository (Briney, 2014). In another lightning talk covered by Briney (2014), Kristen Burgess stresses that if libraries are to encourage others to practice good data management, libraries also need to practice it themselves.

User education can also be provided when librarians/researchers are navigating the institutional repository. One method of doing instruction is through a live help chat widget on the data repository page, but the repository manager needs to negotiate with online reference staff before implementing this service (Witt, 2012). Online reference staff would also need training for how to deal with questions coming from the site. Witt (2012) asked online reference staff to post a script that asks if the user's question can be "referred to a data librarian who will get back to them in 1-2 business days" (p. 184). Helping users to navigate the data repository can also provide valuable feedback for improvements and ways to target different disciplines (Gordon et al., 2015). While implementing RDM practices into curricula and facilitating cultural change in academia are the ultimate goals, libraries can provide educational support until these changes take place (Eaker, 2014; Shen & Varvel, 2013).

Creating incentives

Some libraries found that the ease of accessing and storing well-organized data has not provided enough incentive to create good data management plans. The solution at the University of Colorado Boulder was to provide monetary incentives (Knuth et al., 2015). The school created a \$2000 award for the "Best Digital Data Management Plans and Practices Competition" with winners in five disciplines: arts and humanities, social sciences, life sciences, physical sciences, and engineering (Knuth et al., 2015). Of all the marketing performed by Knuth et al. (2015), the financial incentives were the most successful outreach effort. While financial incentives of such magnitude may not be possible in the current climate of budget constraints, other incentives such as a discount on other library data services may be given to researchers with data management plans who meet certain criteria for excellence.

Methods

The methods used for this study are based off a similar study performed by Briney, Goben, and Zilinski (2015) that reviewed library websites and institutional data policies of 206 research-intensive American universities. The search strategies and definitions from Briney et al.'s (2015) study were applied to the U15 Group of Canadian Research Universities: a group of 15 research-intensive universities in Canada. For each U15 university, the author of this study searched the university website for publically available university data policies, either standalone or as part of an intellectual property policy. The search used similar terms to Briney et al. (2015) including "data policy", "data retention", and "data management". The words "data" and "policy" were used if the initial search returned no results. For French websites, the search terms "données" and "politiques" were used in place of "data" and "policies" respectively.

The author's knowledge of the French language and Google page translations were used to help navigate and search the French websites. Following Briney et al.'s (2015) example, data policies for a specific department, a whole university system, and institutional data were excluded.

Information on how the U15 university libraries supported research data management was collected through the respective library websites. The information included the presence or absence of a data librarian, RDM services, and a repository specifically designed for data or an institutional repository that accepts data submissions. The author used Briney et al.'s (2015) definition of a data librarian: a librarian with data as part of their title or primary component of their job description. When accessible, the author searched the university's staff listing for data librarians. The author also used Briney et al.'s (2015) limitation that training or consultation were necessary components to qualify as data services, therefore RDM support had to go beyond a research guide. The author maintained that for the institutional repository to be considered a data-specific repository, the institutional repository documentation "must specifically mention data as an acceptable type of content for deposit" (Briney et al., 2015, p. 7). The information retrieval was completed in November 2015 and changes may have occurred at the universities since the date of collection.

Limitations

Universities may be falsely coded since this study lacks the inter-coder reliability present in Briney et al.'s (2015) multi-author study. French is not the first language of the author, which caused some difficulty on the French language sites, and therefore limitations in the author's French translation skills were compensated through Google translation software. Translation software is not always accurate, particularly in areas of semantics. This study only looked at information pertaining to RDM services that is publically available. RDM services may be taking place at some university libraries that do not advertise these services publically because they are still in the trial phase of implementation or for other reasons. The assumption being made is that if research data services are being provided, the library would want to advertise these services and their website is a convenient location to do so.

Results

Table 1 illustrates which universities have a data policy (either standalone or as part of an intellectual property policy), a data librarian, RDM services, and a repository specifically for data or an institutional repository that accepts data submissions. While the majority of U15 universities have university data policies (53%), data librarians (60%),

and repositories for data (53%), many universities do not provide research data management services (33%) (Figure 1).

Table 1. The presence of data policies, data librarians, research data services, and a repository for data storage (either specifically for data (SD) or accepted into the institutional repository (IR)) as of November 8, 2015 within the Canadian U15 universities.

University	Data policy?	Data librarian?	RDM services ?	Repository specifically for data (SD) or accepts data into institutional repository (IR)?
University of Alberta	Yes	No	Yes	SD
University of British Columbia	Yes	Yes	Yes	SD
University of Calgary	Yes	Yes	No	No
Dalhousie University	No	Yes	Yes	IR, SD currently in testing
Université Laval	No	No	No	No
University of Manitoba	No	No	No	No ¹
McGill University	No	Yes	No	No
Université de Montréal	Yes	No	No	IR
University of Ottawa	Yes	Yes	No	IR
Queen's University	No	Yes	Yes	SD
University of Saskatchewan	Yes	Yes	No	No
University of Toronto	Yes	Yes	Yes	SD
University of Waterloo	Yes	No	No	No ²

Western University	No	Yes	No	SD
--------------------	----	-----	----	----

¹ Due to the lack of mention of data within their institutional repository's documentation, the University of Manitoba is listed as not accepting data into their institutional repository, even though the repository contains data.

² The University of Waterloo accepts video, audio, and image files, but does not accept all forms of data, therefore it was classified as an institution that does not have a repository for data.

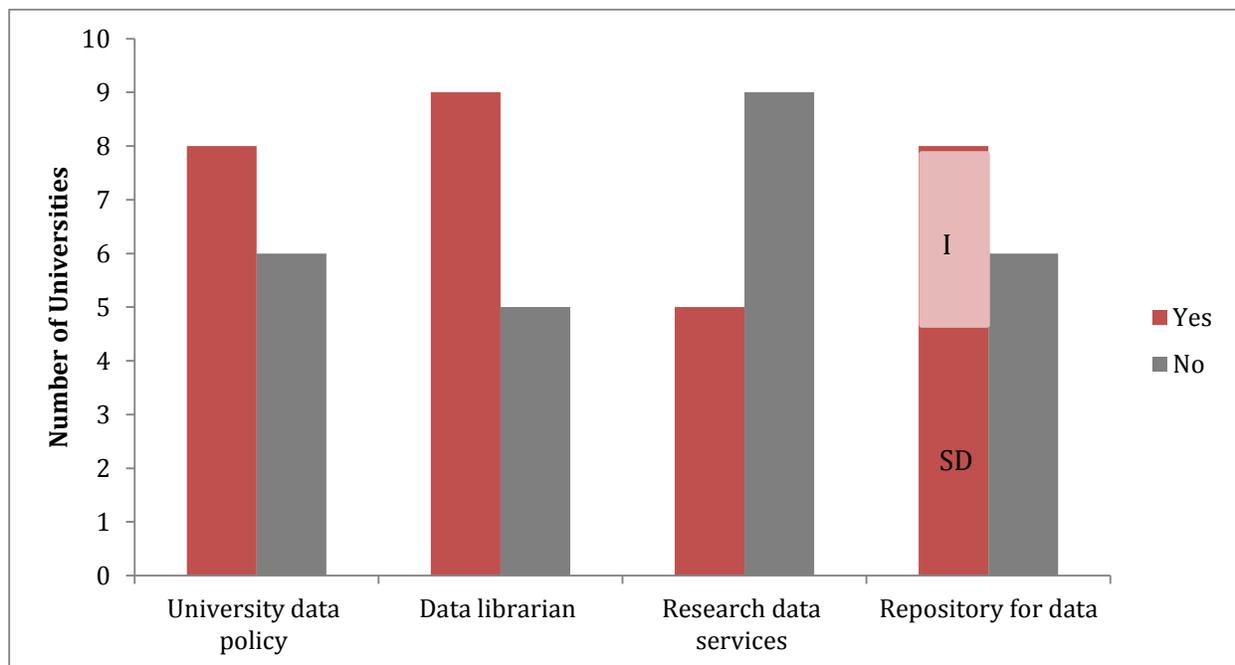


Figure 1. The presence of data policies, data librarians, research data services, and a repository for data storage (either specifically for data (SD) or accepted into the institutional repository (IR)) as of November 8, 2015 within the Canadian U15 universities.

Discussion

The methods used by Briney et al. (2015) attempt to illustrate what a researcher at the institution or an external collaborator would find if they were looking for the institution's data policy or for help with RDM at the institution's library. In Briney et al.'s (2015) study, only 44% of universities studied had some type of research data policy, whereas more than half of the Canadian universities in my sample had the same type of policies (Figure 1). The U15 universities also have a higher ratio of data-specific repositories to institutional repositories than their American counterparts, of which 11% had data-specific repositories and 58% had institutional repositories that accepted data (Briney et al., 2015). The U15 universities had more data librarians than data services, which is

the opposite of what Briney et al. (2015) found in American universities. In libraries that do not provide research data services, the data librarians were often specialists who could help users find data and sometimes use data analysis tools, but not help curate their own data. The University of Alberta does provide data management services, but does not currently have a data librarian, although this may change in the near future as they are looking to hire a RDM services coordinator. At institutions that did not have a “data librarian”, it may be that a librarian manages RDM requests, but does not have data in their title or clearly indicated as part of their job description on the library’s website, which may be the case for the University of Alberta. The importance of having an identifiable data librarian is illustrated by Knight’s (2015) finding that researchers prefer to contact a person as opposed to contacting a generic department email.

The findings display a positive start to opening Canadian research data to the world, but the next step is developing RDM services to help researchers develop better data management strategies.

Conclusion and further studies

Looking back at the compiled approaches, a single full time librarian—let alone a librarian with non-RDM commitments—would be hard-pressed to be aware of funder and institutional data policies, reach out to data service providers on campus and beyond, understand researcher data management needs and find RDM champions, implement research data services strategically, plan for growth in RDM services, market the RDM services, and create incentives (Creamer, 2015). All of these recommendations may be taken on sequentially, but the policies, researcher needs, and software are constantly changing. Even in regards to getting RDM services started, Knight (2015) comments that the initiation of RDM services at the LSHTM would have been slower if the “existing library and archives staff had been asked to perform the work as a portion of their existing roles” (p. 436). Of course, the reality of decreasing budgets means a reduced ability to hire new staff, and more tasks being taken on by fewer people. Therefore, a team approach to data services may be the best option for implementing data management services in an efficient and cost-effective manner.

Canadian academic libraries are well positioned to develop research data management services. The majority of U15 universities have data librarians, data repositories, and data policies and many of the libraries already have RDM research guides. Libraries across Canada are also collaborating to create data management planning tools, like Portage, to help researchers get the most out of data

management plans and enable sharing and team comments on the plan (Moon, 2015). A Canadian curriculum for research data management workshops has yet to be published, which indicates another project that could emerge from collaborations between Canadian academic libraries. A review of recent literature revealed a number of successes and complications involved with implementation of RDM services in universities around the world. The highlighted approaches provide examples of successful techniques that Canadian university libraries may find useful as they progress through the implementation process. Alternatively, the final report of the LIBER working group on E-Science/Research Data Management provides ten concise recommendations for libraries getting started with RDM (Christensen-Dalsgaard et al., 2012).

Many gaps in the current literature of research data management have already been addressed by previous writers including Herold (2015), who advocates for a closer examination of journal preservation practices and guidelines to ensure that depositing data in journals complies with funder requirements. The deep web concept, which addresses the large amount of information that cannot be found by search engines such as Google and Yahoo and only found on local sites due to lack of indexing, also prompts the question of the discoverability of data sets by web search engines. The information gathered on the U15 universities for this research only corresponds to a portion of the information collected in the Briney et al. (2015) study. Further investigation could be made into the content of data policies in Canadian research universities, similar to Briney et al.'s (2015) work. Interesting times lie ahead for Canadian academic libraries and the implementation of research data management services.

References

- Bishoff, C., & Johnston, L. (2015). Approaches to data sharing: An analysis of NSF data management plans from a large research university. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–27. doi: 10.7710/2162-3309.1231
- Briney, K. (2014). Lightning talk summary. *Bulletin of the Association for Information Science & Technology*, 40(6), 35–37. Retrieved from [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2373-9223](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2373-9223)
- Briney, K., Goben, A., & Zilinski, L. (2015). Do you have an institutional data policy? A review of the current landscape of library data services and institutional data policies. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–25. doi: 10.7710/2162-3309.1232
- Buys, C. M., & Shaw, P. L. (2015). Data management practices across an institution: Survey and report. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–24. doi: 10.7710/2162-3309.1225
- Choudhury, S. (2012). ASERL webinar: Data conservancy & data management services at Johns Hopkins University Libraries. Retrieved from <https://vimeo.com/36050328>
- Christensen-Dalsgaard, B., van den Berg, M., Grim, R., Horstmann, W., Jansen, D., Pollard, T., & Roos, A. (2012). Ten recommendations for libraries to get started with research data management: Final report of the LIBER working group on E-Science / Research Data Management. Retrieved November 13, 2015, from http://libereurope.eu/wp-content/uploads/The_research_data_group_2012v7_final.pdf
- Cox, A. M., & Corral, S. (2013). Evolving academic library specialties. *Journal of the American Society for Information Science & Technology*, 64(8), 1526–1542. doi: 10.1002/asi.22847
- Creamer, A. (2015). Current issues and approaches to curating student research data. *Bulletin of the Association for Information Science & Technology*, 41(6), 22–25. doi: 10.1002/bult.2015.1720410610

- Davis, H. M., & Cross, W. M. (2015). Using a data management plan review service as a training ground for librarians. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–20. doi: 10.7710/21623309.1243
- Doty, J., Herndon, J., Lyle, J., & Stephenson, L. (2014). Learning to curate. *Bulletin of the Association for Information Science & Technology*, 40(6), 31–34. Retrieved from [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2373-9223](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2373-9223)
- Doty, J., Kowalski, M. T., Nash, B. C., & O’Riordan, S. F. (2015). Making student research data discoverable: A pilot program using Dataverse. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–25. doi: 10.7710/2162-3309.1234
- Eaker, C. (2014). Educating researchers for effective data management. *Bulletin of the Association for Information Science & Technology*, 40(3), 45–46. Retrieved from [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2373-9223](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2373-9223)
- Fox, R. (2013). The art and science of data curation. *OCLC Systems & Services*, 29(4), 195–199. doi: 10.1108/OCLC-07-2013-0021
- Giarlo, M. J. (2013). Academic libraries as data quality hubs. *Journal of Librarianship & Scholarly Communication*, 1(3), 1–10. doi: 10.7710/2162-3309.1059
- Gordon, A. S., Millman, D. S., Steiger, L., Adolph, K. E., & Gilmore, R. O. (2015). Researcher-library collaborations: Data repositories as a service for researchers. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–17. doi: 10.7710/2162-3309.1238
- Government of Canada. (2015). Draft Tri-Agency Statement of Principles on Digital Data Management - Science.gc.ca. Retrieved from <http://www.science.gc.ca/default.asp?lang=En&n=83F7624E-1>
- Greenberg, J., White, H. C., Carrier, S., & Scherle, R. (2009). A metadata best practice for a scientific data repository. *Journal of Library Metadata*, 9(3/4), 194–212. doi: 10.1080/19386380903405090
- Hayslett, M. (2015). Data world does not lack standards. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–5. doi: 10.7710/2162-3309.1245

- Henderson, M. E., & Knott, T. L. (2015). Starting a research data management program based in a university library. *Medical Reference Services Quarterly*, 34(1), 47–59. doi: 10.1080/02763869.2015.986783
- Henderson, M., Raboin, R., Shorish, Y., & Van Tuyl, S. (2014). Research data management on a shoestring budget. *Bulletin of the Association for Information Science & Technology*, 40(6), 14–17. Retrieved from [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2373-9223](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2373-9223)
- Herold, P. (2015). Data sharing among ecology, evolution, and natural resources scientists: An analysis of selected publications. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–23. doi: 10.7710/2162-3309.1244
- Huwe, T. K. (2013). Data discovery and data curation going hand in hand. *Computers in Libraries*, 33(3), 17–19. Retrieved from <http://www.infotoday.com/cilmag/>
- Johnston, L. R. (2014). Developing a data curation service: Step #1: Work with what you've got. *Bulletin of the Association for Information Science & Technology*, 40(4), 45–47. Retrieved from [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2373-9223](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2373-9223)
- Joint Task Force on Librarian's Competencies in Support of E-Research and Scholarly Communication. (2015). Librarians' competencies profile for research data management. Retrieved from <https://docs.google.com/document/d/15Kqg6w6AFsdbbaxcRCx7DZ3OERnQk6E0UhQ4UeRu0Cc/edit>
- Knight, G. (2015). Building a research data management service for the London School of Hygiene & Tropical Medicine. *Program: Electronic Library & Information Systems*, 49(4), 424–439. doi: 10.1108/PROG-012015-0011
- Knuth, S. L., Johnson, A. M., & Hauser, T. (2015). Research data services at the University of Colorado Boulder. *Bulletin of the Association for Information Science & Technology*, 41(6), 35–38. doi: 10.1002/bult.2015.1720410614
- Kouper, I., Akers, K., & Lavin, M. (2013). Data curators at work: Focus on projects and experiences. *Bulletin of the Association for Information Science &*

Technology, 40(1), 45–46. Retrieved from
[http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2373-9223](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2373-9223)

Mannheimer, S. (2014). Ready, engage! Outreach for library data services. *Bulletin of the Association for Information Science & Technology*, 41(1), 42–44. Retrieved from [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2373-9223](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2373-9223)

Mclure, M., Level, A. V., Cranston, C. L., Oehlerts, B., & Culbertson, M. (2014). Data curation: A study of researcher practices and needs. *Portal: Libraries and the Academy*, 14(2), 139–164.

Moon, J. (2015). Portage: Supporting Canadian innovation through shared expertise and stewardship of research data. Retrieved from <http://www.slideshare.net/CASRAI/casrai-2015-carl-portagedmpassistant-jeffmoon>

Newton, M. P., Miller, C. C., & Bracke, M. S. (2011). Librarian roles in institutional repository data set collecting: Outcomes of a research library task force. *Collection Management*, 36(1), 53–67. doi: 10.1080/01462679.2011.530546

Ogier, A., Hall, M., Bailey, A., & Stovall, C. (2014). Data management inside the library: Assessing electronic resources data using the data asset framework methodology. *Journal of Electronic Resources Librarianship*, 26(2), 101–113. doi: 10.1080/1941126X.2014.910406

Raboin, R. (2015, June 13). Testimony from “Teaching research data management with the New England Collaborative Data Management Curriculum”, SLA 2015 Annual Conference, Boston, Massachusetts.

Read, K. B., Surkis, A., Larson, C., McCrillis, A., Graff, A., Nicholson, J., & Xu, J. (2015). Starting the data conversation: Informing data services at an academic health sciences library. *Journal of the Medical Library Association*, 103(3), 131–135. doi: 10.3163/1536-5050.103.3.005

Saunders, L. (2015). Academic libraries' strategic plans: Top trends and under-recognized areas. *Journal of Academic Librarianship*, 41(3), 285–291. doi: 10.1016/j.acalib.2015.03.011

- Shen, Y., & Varvel, V. E. (2013). Developing data management services at the Johns Hopkins University. *Journal of Academic Librarianship*, 39(6), 552-557. doi: 10.1016/j.acalib.2013.06.002
- Toups, M., & Hughes, M. (2013). When data curation isn't: A redefinition for liberal arts universities. *Journal of Library Administration*, 53(4), 223–233. doi: 10.1080/01930826.2013.865386
- University of Alberta. (2010). Information technology security policy. Retrieved from <https://policiesonline.ualberta.ca/PoliciesProcedures/Policies/Information-Technology-Security-Policy.pdf>
- Varvel, V. E., & Shen, Y. (2013). Data management consulting at the Johns Hopkins University. *New Review of Academic Librarianship*, 19(3), 224–245. doi: 10.1080/13614533.2013.768277
- Vitale, C. R. H., Marshall, B., & Nurnberger, A. (2015). You're in good company: Unifying campus research data services. *Bulletin of the Association for Information Science & Technology*, 41(6), 26-28. doi: 10.1002/bult.2015.1720410611
- Walker, W., & Keenan, T. (2015). Going beyond availability: Truly accessible research data. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–8. doi: 10.7710/2162-3309.1223
- White, H. C. (2014). Descriptive metadata for scientific data repositories: A comparison of information scientist and scientist organizing behaviors. *Journal of Library Metadata*, 14(1), 24–51. doi: 10.1080/19386389.2014.891896
- Whitmire, A. L. (2015). Implementing a graduate-level research data management course: Approach, outcomes, and lessons learned. *Journal of Librarianship & Scholarly Communication*, 3(2), 1–22. doi: 10.7710/2162-3309.1246
- Witt, M. (2012). Co-designing, co-developing, and co-implementing an institutional data repository service. *Journal of Library Administration*, 52(2), 172–188. doi: 10.1080/01930826.2012.655607
- Wright, S., Whitmire, A., Zilinski, L., & Minor, D. (2014). Collaboration and tension between institutions and units providing data management support. *Bulletin*

of the Association for Information Science & Technology, 40(6), 18–21.
Retrieved from [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2373-9223](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2373-9223)