



Review Article

Shifting Horizons: A Literature Review of Research Data Management Train-the-Trainer Models for Library and Campus-Wide Research Support Staff in Canadian Institutions

Dr. Felicity Tayler
Research Data Management Librarian
University of Ottawa
Ottawa, Ontario, Canada
Email: ftayler@uottawa.ca

Maziar Jafary
PhD Candidate and Part-Time Professor
School of Sociological and Anthropological Studies
University of Ottawa
Ottawa, Ontario, Canada
Email: Mjafary@uottawa.ca

Received: 4 Aug. 2020

Accepted: 16 Nov. 2020

© 2021 Taylor and Jafary. This is an Open Access article distributed under the terms of the Creative Commons-Attribution-Noncommercial-Share Alike License 4.0 International (<http://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly attributed, not used for commercial purposes, and, if transformed, the resulting work is redistributed under the same or similar license to this one.

DOI: 10.18438/eblip29814

Abstract

Objective – In consideration of emerging national Research Data Management (RDM) policy and infrastructure, this literature review seeks answers to the following questions:

- 1) What is the most effective way for a Canadian research university to build capacity among library and campus-wide research support staff, with a view towards providing coordinated RDM support services for our researcher community?
- 2) What international training models and course offerings are available and appropriate for a local context?
- 3) What national guidelines and best practices for pedagogical design and delivery can be

adapted for a local context?

Methods – This literature review synthesizes a total of 13 sources: 9 articles, 2 book chapters, and 2 whitepapers. The whitepapers were selected for a narrative literature review because of their focus on case studies detailing train-the-trainer models. Within the 13 sources we found 14 key case studies. This review serves as a supplement to the 2017 CARL Portage Training Expert Group white paper, “Research Data Management Training Landscape in Canada,” the focus of which was to identify RDM training gaps in order to recommend a coordinated approach to RDM training in a national environment.

Results – The narrative review of case studies revealed three thematic areas. Firstly, pedagogical challenges were identified, including the need to target training to RDM support staff such as librarians and researchers, as they comprise distinct groups of trainees with divergent disciplinary vocabularies and incentives for training. Secondly, the case studies cover a broad range of pedagogical models including single or multiple sessions, self-directed or instructor-led, in-person or online instruction, and a hybrid of the two. Finally, RDM training also emerged as a key factor in community building within library staff units, among service units on campus, and with campus research communities.

Conclusion – RDM training programs at local institutions should be guided by a set of principles aligned with the training methods, modes of assessment, and infrastructure development timeline outlined in a national training strategy. When adapting principles and training strategies to a local context, the following trends in the literature should be considered: librarians and researchers must have meaningful incentives to undertake training in RDM or to join a community of practice; disciplinary-specific instruction is preferable to general instruction; a librarian’s own training opportunities will influence their ability to provide discipline-specific RDM instruction to researchers; in-person training opportunities improve learning retention and produce beneficial secondary effects, whereas online instruction is most effective when paired with an in-person component; generalized third-party RDM training should be adapted to local context to be meaningful. Future directions for RDM training will integrate into open access and digital scholarship training, and into cross-disciplinary, open science communities of practice.

Introduction

This literature review was undertaken to help the Research Services Division of the University of Ottawa Library determine effective training methods for library and campus-wide research support staff, with a view towards providing coordinated RDM support services for the researcher community. For the last four years, University of Ottawa has held an annual, in-person, campus-wide RDM training event, attended by researchers and a wider general audience. The event was also attended by RDM-curious librarians and researchers from other

universities. By 2019 this event had gained national attention as the Shifting Horizons training series. The 2020 edition presented a national training program, developed through Canada’s CARL Portage Network. One of the goals of the program is to provide basic RDM skills training for librarians and for support staff in the Research Office, labs, faculty departments, and Central IT. Despite the event’s success, the library’s Research Services Division needed to evaluate whether a single annual event was the most effective way to achieve the vision of campus-wide RDM awareness and a coordinated service model. As was observed in

the follow up documentation of the event, the [RDM Readiness Report](#), the stakeholders of a coordinated RDM service model at University of Ottawa continue to face the same challenges as those identified by 241 librarians in a study by Tang and Hu (2019). Some of these challenges include issues with staffing and upskilling, promotion of the service, service quality, and shared understanding among campus departments.

Aims

The articles reviewed here supplement the CARL Portage Training Expert Group white paper, “Research Data Management Training Landscape in Canada” (Fry et al., 2017). The purpose of this white paper was to identify “significant issues and gaps in RDM training in Canada” and to recommend a national, coordinated approach to RDM training (Fry et al., 2017, p. 2). This recommendation is driven by the understanding that expertise in data stewardship is unevenly distributed across higher education institutions and is often isolated within disciplinary areas. In contrast to RDM infrastructures elsewhere, which cohere around disciplinary or national service centres, a critical mass of RDM expertise in Canada is organized within the academic library community. To date, this report’s holistic multi-platform vision of a coordinated national training curriculum, to level the “playing field” has been enacted in a modest capacity through best practices, data primers and ad-hoc webinar training, supplemented by single-day, in-person sessions reflecting the individual expertise of members of the Portage Training Expert Group (Fry et al., 2017, p. 7). The day-long training event at University of Ottawa, led by James Doiron, who is both an author of the training landscape white paper and the RDM Services Coordinator at the University of Alberta Libraries, is one example of the in-person sessions currently offered through CARL Portage. Once an institution has participated in

the training, the next steps are unknown. For example, there is no clear direction, recommended strategies, or coordinated curriculum resources, at the national level, to support the long-term development of highly qualified personnel (HQP) providing RDM services at libraries selected to play a leadership role in this area.

Methods

This literature review synthesizes a total of 13 sources, including 9 articles, 2 book chapters, and 2 whitepapers from a larger sample of 35 texts published within the last ten years (2010-2020). The authors cited seven additional supporting sources in the analysis in order to provide the contextual framing for the thematic approach of this narrative review. Keyword searches such as “research data management (and) training” were undertaken in databases including LISTA and Library and Information Science Source. Because RDM training is an emerging field, contingent upon variable jurisdictional challenges, policy, and funding environments, the aim was not to be exhaustive, nor systematic in our searches. Instead, a “snowball” search for key articles, white papers, and reports shared by colleagues on RDM-themed listservs such as CANLIB-DATA, or IASSIST, or referenced at annual RDA Plenaries supplemented these keyword database searches. In addition to the snowball searching, the authors contacted various content experts to review the abstracts collected to ensure that no important sources were missed. Though the number of sources reviewed is minimal, this is an indicator that RDM is an emerging area of librarianship, which is also interdisciplinary in nature. There are simply not that many articles out there yet, and this literature review aims to address this gap while recognizing that there is still work to do in this area.

In the 13 sources selected for synthesis, we found 14 key cases for analysis. Out of the 13

sources selected for synthesis, as outlined above, 9 of the selected sources had a single case study focus (Baker et al., 2016; Grootveld & Verbakel, 2015; Haddow, 2014; Helbig, 2016; Papadopoulou & Miller in Clare et al., 2019; Papadopoulou & Grabauskiene in Clare et al., 2019; Wittenberg et al., 2018; Southall & Scutt, 2017; Read et al., 2019). Out of the original 13, 2 of the selected sources covered multiple case studies (Bryant et al., 2018; Surkis & Read, 2015), while 2 of the sources dealt with the same case study (Tang & Hu, 2019; Shipman & Tang, 2019). In choosing the case studies, the authors prioritized European, North American, and Australian examples as their social and academic contexts are comparable to those of Canada. However, this geographic limitation and focus on English-language sources introduces a bias to this review. This selection bias does not reflect a deliberate exclusion of other regional models, rather it echoes a trend to build Canadian digital research infrastructure on existing models such as the [European Open Science Cloud \(EOSC\)](#), or to look to best practices in RDM established by the [Digital Curation Centre](#) in the UK, and American RDM service models as outlined by [OCLC](#).

Results

The review of the literature is divided into three sections, reflecting themes within the articles and case studies. The first section discusses challenges and opportunities for RDM training in universities. Outreach and pedagogical issues were identified by several authors, including the development of targeted RDM training to two distinct groups of trainees: RDM support staff, including librarians, and researchers. These two groups differ in their incentives for training participation and their use of discipline-specific vocabulary. With these challenges in mind, the evaluation of training models for success and areas of improvement will be discussed. The second section explains different approaches to curriculum and pedagogical design in RDM

training. The case studies cover a range of pedagogical models and whenever possible evaluations of these training methods and formats of pedagogical engagement for RDM training are highlighted. Finally, the third section looks at how RDM training operates as a means of community building within library staff units, between service units on campus, and within campus research communities. This final section also covers internal and external partnerships which are necessary to develop RDM training.

Discussion

Challenges and Opportunities for RDM Training in Universities

While many of the texts that were retrieved in the searches addressed developing RDM services around best practices, or outlined approaches for broader data literacy training strategies, this literature review focuses on train-the-trainer models as a unique subset of the RDM training landscape. Because the literature in this area is emerging, this review presents a combination of conclusions drawn from train-the-trainer models alongside approaches to training researchers. In a train-the-trainer model, the targeted audience of trainees are librarians and other research support staff. In the researcher trainer model, the targeted audience members are typically faculty, student research assistants, and other affiliates of disciplinary research projects. However, in practice the line between these roles is blurry, as trainers often become a secondary audience of the training for researchers, and researchers can also benefit from train-the-trainer sessions as they can perform a trainer role as part of their own research team. Furthermore, as this review demonstrates, there is a correlation between the pedagogical model applied to train-the-trainer sessions and the effectiveness of these trainers to then shape learning experiences for researchers. By outlining the challenges to providing RDM

training to researchers in this section, the recommended best-practices can inform approaches to train-the-trainer models. We begin with the principle that RDM is not generic. Instead, librarians and other research support staff need a fundamental understanding of how data flows and data management differ between disciplinary research methods, and how to recommend relevant engagement with local, national, and international infrastructure contexts.

RDM training for librarians and other research support staff will have an impact on the success of RDM services delivered. Both Tang and Hu (2019) and Surkis and Read (2015) identify significant barriers and pedagogical challenges of RDM training for librarians and other research support staff, beyond the administrative concerns of budget and capacity. For example, librarian language and vocabulary does not translate well to the disciplinary environment of researchers and other stakeholders. Such specialized RDM vocabulary might not be well received or even understood by researchers. Another challenge could be a lack of training for librarians and research support staff on different approaches to research data management within the field of study, as defined by the researchers' peers and funding bodies. Tang and Hu's (2019) needs assessment highlighted the need for key training in strategic communication of RDM service models to library and university administration, while Surkis and Read (2015) instead stress that when the goal is the improvement of training offerings for researchers, instructors from the library sector (and related fields), as part of their own training, should engage in interviews with researchers in different fields. This exercise would help librarians better understand researchers' needs and expectations from RDM services. A later study by Read et al. (2019) further explored this lack of disciplinary knowledge as a high barrier to librarian engagement with RDM services in biomedical

fields, due to a "lack of comfort engaging with researchers" (p. 2). Read et al. (2019) also noted a double gap in the training landscape, identifying that a "lack of satisfactory curricula" (p. 2) to train *both* librarians and researchers in RDM further contributed to the lack of RDM service offerings in biomedical fields.

Engaging Researchers with Data Management: The Cookbook (Clare et al., 2019), includes several case studies of RDM engagement and collaborations among researchers. The case studies demonstrate how librarians and other research support staff with disciplinary awareness can encourage researchers to consider research data management practices and services as an extension of their disciplinary peer communities. In one of the chapters focusing particularly on RDM training, Papadopoulou and Miller evaluate the format of training "mini-events" for their impact on building a community of RDM supports and data management best practices at the Vilnius University Library in Lithuania. Each of these mini-events (delivered either as half-day or full-day workshops) consisted of three incremental phases: familiarity of the participants with RDM support services; learning how to use various available tools; sharing research data in practice (Papadopoulou & Grabauskiene, 2019). Papadopoulou and Grabauskiene specify that one of the challenges faced by these RDM training sessions is reaching out to, and persuading, the uninterested researchers to attend. One proposed strategy is to do peer outreach rather than through a generic unit, such as Information Services. Secondly, based on their study of a conference at the University of Edinburgh, Papadopoulou and Miller propose that the events should include presentations by researchers from multiple university faculties. Such presentations might discuss RDM best practices and their impact on researchers' work, thereby encouraging their disciplinary peers to participate. Thus, the presentations can also be interactive sessions

among the researcher peers themselves (Papadopoulou & Miller, 2019).

Approaches to Research Data Management Training

The previous section outlined challenges of RDM training such as the gap in terminology shared by research support staff and the researcher community that they support, and the researchers' lack of interest in RDM if it is perceived to be beyond the scope of methodologies shared by their disciplinary community. These challenges support the *Research data management training landscape in Canada: A white paper* finding that pedagogical design needs to be mapped to trainee needs and is a necessary learning objective for librarians and other research support service providers (Fry et al., 2017). This review has revealed multiple approaches to RDM training, specific to the trainee contexts. Although we focus here on librarians and other research support staff as "trainees", it is with an understanding that their training opportunities have an impact on the quality of RDM training and service provision available to researchers. Further, this review notes several approaches to pedagogical design for RDM training, which can be broadly categorized as: generalized instruction or discipline-specific, single or multiple sessions, self-directed or instructor-led, in-person or online instruction (and most often, a hybrid of the two).

The literature shows that there are significant advantages to delivering discipline-specific or targeted RDM training. However, a generalized approach to RDM training may be favoured due to perceived scalability. As mentioned, Read et al. (2019) note that available online training for librarians is inadequate to build RDM service capacity in biomedical fields, as none have the necessary disciplinary focus. This focus on general RDM training for librarians further contributes to a gap in disciplinary-specific

training curricula for researchers. After reviewing Humboldt University of Berlin's RDM initiative, launched as a joint venture between Computer and Media Service, the Research Service Centre, the University Library, and the Vice President for Research, Helbig (2016) similarly concludes, "Although general workshops on research data management are more scalable in comparison to discipline-specific workshops, the advantages of a tailored approach outweighed this concern" (p. 2). Humboldt University's RDM training initiative consisted of one-day workshops aimed at helping PhD students and researchers in the Geography Department. Groups of six to eight trainees were formed in order to facilitate the learning process. RDM specialists at the University felt that a targeted approach would be advantageous. Through a priori surveys and interviews with researchers and graduate students, the workshops were designed for the specific needs of that department. By understanding the nature of RDM in geography, specialists were able to provide an interactive session encouraging the full participation of the trainees. Other universities such as Monash University in Australia, University of Edinburgh in the United Kingdom, and University of Illinois, in the United States, offer courses to targeted campus groups based on their needs. Such needs are identified through consultation with strategic research management services at these universities, as well as in-person discussions with individual researchers around the campus. Bryant et al. (2018) explain that the integrated instruction model in a semester-long course is a preferable method because it is sustainable, as they observe, "the most resource-intensive approach to supporting RDM education is through in-person, instructor-led workshops" (p. 10). However, if a workshop approach is taken over a course integration approach, Bryant et al. (2018), argue that RDM educational services should strategically align their workshops with course content and with broader institutional policies of the respective

university (such as conforming to the requirements of Data Management Plans).

Within the literature, the choice between disciplinary focus or generalized curriculum models, is paralleled by the choice of delivery mode through online modules, in-person sessions, or a hybrid of the two. Online training modules are among the most popular among RDM professionals because they are thought to allow flexibility for accommodating work schedules (Tang & Hu, 2019). Read et al. (2019) note that the required time commitment is a strain on working librarians and there is a significant rate of non-completion of online training. Read et al. (2019) also showed that while online modules improve the “understanding of and comfort level with RDM” in-person instruction resulted in “improved RDM practices” (p. 1). The differing experiences between online and in-person learning led Read et al. (2019) to develop a hybrid, or “two-tier” coordinated approach to RDM training for health sciences librarians, and for biomedical researchers that the librarians will, in turn, train and support. There were seven self-paced, multi-media, online modules produced to train librarians. The modules covered general RDM topics and applications of RDM in health science methodologies and discipline-specific data standards. An evaluation form embedded at the end of each module was included for self-assessment. Once a librarian indicated comfort with the content, they received a Teaching Toolkit which included a lesson plan and related materials to teach RDM to biomedical researchers via a 60-90 minute in-person session. This hybrid, coordinated model improved the librarian’s ability to deliver an RDM session for researchers; as Read et al. (2017) observe, “the online modules were concise and directly tied to the Teaching Toolkit, a curriculum specifically created for use by the librarians to teach RDM locally, thus addressing the time constraints of working professionals...” (p. 8).

The learning objectives of online training options are improved when paired with in-person instruction. Bryant et al. (2018) explain that the MANTRA Research Data Management Training modules, promoted on the website as “a free online course for those who manage digital data as part of their research project” (p.10), is a series of eight generic self-paced modules and tutorials that are supplemented by in-person training courses by RDM professionals, at the University of Edinburgh. The online modules, initially built for researchers and graduate students, have influenced pedagogical design of RDM training for librarians and research support staff, not only at the host institution, but also for researchers and staff at other institutions. In 2013, MANTRA launched a DIY Training Kit for Librarians to facilitate the remote training modules. Built for the UK research and funding environment, the course can be adapted locally to include online and in-person instruction, covering data management planning, organizing and documenting data, data storage, data sharing and ethics, and questions around data management. In a blog post, Haddow (2014) writes of the experience of adapting and delivering the MANTRA DIY Training Kit for Librarians at the Sterling University of Edinburgh. According to Haddow (2014), the subject librarian members of a dedicated local RDM Task Force, “found it beneficial to set time aside as a team to look at this issue;” (para. 2) however, they noted challenges and significant time investment for the local facilitator to adapt the course content. As Haddow (2014) explains: “the instructions were sometimes not clear but by the end I figured out that I just needed to look at the manual.” (para. 4)

The “Data Intelligence 4 Librarians course” was released in 2011 by 3TU.Datacentrum, a partnership among three universities in The Netherlands (the partnership was later called 4TU.ResearchData (2020)). This course provides another example of a learning platform targeted

to digital preservation professionals and included two in-person sessions at the beginning and the end of the training period. During the in-person sessions, coaches would teach the trainees, while during the online sessions, trainees were expected to be prepared for each unit and complete assignments by themselves or in pairs. Throughout the online portion, trainees could reach out to their respective coaches through an established online platform. Later, the course was transformed into “Essentials 4 Data Support” whose target group was a more widely-defined group of professionals identified as “data supporters,” further defined as “people who support researchers in storing, managing, archiving and sharing their research data” (p. 244). Trainees from multiple institutions attended and worked mostly in pairs, learning how to write research data plans for fictional scenarios. Participant surveys and networking through online forums following the training were completed (Grootveld & Verbakel, 2015). Feedback indicated that homework assignments were the most valuable element of the course, as the pairing of trainees led to enjoyable discussions. Participants also appreciated learning from researchers, including how they deal with data management issues and about differences between disciplines (Grootveld & Verbakel, 2015). Trainees admitted that the use of audio-visual elements was helpful for their learning experience. Current versions of *4TU.ResearchData* consist of three variants: a combination of in-person sessions and online training platforms, supervised by coaches and open to online discussion forums; a self-directed, online course, open to online discussion forums; a self-directed, online course with no access to coaches or discussion forums.

A recent example of generalized, online RDM training includes the Research Data Management Librarian Academy (RDMLA), for librarians from multiple institutions around the globe (Shipman & Tang, 2019). The curriculum

was based on needs gathered from interviews and a survey conducted by Tang and Hu (2019), as previously discussed, and its intent was to fill gaps training for librarians in higher education, through online training. Although its success cannot be confirmed at this time, the online-only format of RDMLA should be assessed in terms of its ability for librarian trainees to translate their knowledge into researcher training, in consideration of completion rates and the findings of studies on hybrid or in-person models. It is important to note that the RDMLA training is underwritten by the publisher Elsevier, with modules promoting tools in which Elsevier has a vested interest, while the other trainings reviewed are developed through public or local institutional funding streams.

Despite available online solutions to local training gaps, in-person instruction remains a popular approach, as it catalyzes communities of practice around complex skillsets. Wittenberg et al. (2018) discussed workshops launched by a research data management team at the University of California in Berkeley, and show that in-person, ongoing, and discipline-based consultations on RDM by specialized liaison librarians are among the most successful methods of RDM support by university libraries. As they mention, “participants, on average, were more satisfied with domain-based RDM training than they were with general RDM training” (p.328). At the same time, Wittenberg et al. (2018) admit that the success of discipline-based training depends on a scientific community built around RDM, which is mainly based on continuous connections between liaison librarians and researchers.

The *Library Carpentry* workshops with RDM-focused content, as discussed by Baker et al. (2016), are a worthwhile comparison to the online or hybrid teaching models available to librarians, due to the strong emphasis placed on in-person skill sharing and long-term community building. The multi-session

workshop took place in the fall of 2015 over four, three-hour weekly evening sessions at the City University London Centre for Information Science. The workshops had three aims: to blend non-library specific software skills training with existing library specific programs; to collect data on software skills in university libraries; and to build the foundations of a distributed community model for embracing and sustaining software skills in the library. Prior to the sessions, attendees were asked to make a name badge, also identifying their level of knowledge of RDM and related software, for presenters to better guide the attendees. Participants were also encouraged to note the level of knowledge of others to better assist them during the workshop, if needed. In this way, peer-to-peer collaborations were built into the workshop design. Participants shaped workshop content. Session one began with an introduction to basic programming concepts and attendees were asked to reflect on words and phrases associated with programming, code, and software from which they could benefit. Baker et al. (2016) note that many universities around the world use “Data Carpentry Workshops” formats and materials adapted to their local needs, which demonstrates the success of the project. However, they still recognize the need to develop a set of resources to enable workshop attendees to share software skills in their home libraries. It is anticipated that these resources would be predicated on the idea that the best way to reinforce one’s own software skills is through teaching others.

RDM Training as a Means of Community Building

Research data management training landscape in Canada: A white paper (2017) outlined eight principles for developing a coordinated national training curriculum. Several of these principles foreground the community of practice approach adopted by the librarian-led Portage Network RDM Expert Groups. The notion of RDM as a

set of skills and practices shared by a community, whether disciplinary, institutional, professional, or otherwise, is consistent with several of the articles reviewed in this paper, as well as the “data communities” model of researcher behaviour in data sharing, described by Danielle Cooper and Rebecca Springer (2019). However, while communities of practice may be wrapped in a myth of informal organizing, in reality they require leadership and intentional cultivation, particularly as Etienne and Beverly Wenger-Trayner (2015) observe, if they are used for developing the “strategic capability” of an organization or its personnel. Indeed, the strategy of nurturing national RDM infrastructure, training, and support by “building partnerships in the face of complexity” has been carefully crafted by Portage since its early stages (Humphrey, 2020, p. 2). From this perspective, RDM Librarians and other research support staff have a key role in training, as universities develop capacity to comply with RDM requirements of national and international funding agencies. For this reason, this literature review will conclude with the seven principles of RDM training developed at TU Delft (2019), as well as new approaches to librarian RDM training that build upon the intersections of research data management with the workflows, best practices, and scholarly communities of open science.

The TU Delft (2019) principles provide a framework whereby RDM training becomes the mechanism for cultivating a community of practice that is both campus-wide and disciplinary-focused, while reaching beyond the campus into the information circuits of the scholarly community. Significantly, these principles encourage a researcher-focused RDM vocabulary; they foster collaboration between faculty and research support staff across multiple university departments and service providers; and furthermore, there is recognition that the university must provide meaningful incentives that motivate trainees, whether they

are administrators, librarians, research support staff, researchers, or students, to join the community of practice. The TU Delft “Open Working” website (2019) outlines some principles including: “whenever possible, data and software management training should be built upon the existing faculty-specific courses”; “building and delivering such training must be a collaborative effort between faculties, the library, graduate school and other university services”; and “library and graduate schools should continuously engage in consultation processes with PhD students and researchers.” (para. 2, 4, 6) At the same time, the principles recommend engagement with organizations outside universities as vital in making training resources sustainable. In order to successfully implement this vision, the TU Delft principles recognize that researchers must receive the proper incentives to participate and contribute to the training. The library should also solicit feedback from researchers to iteratively improve and update the training content. Finally, the principles reinforce that courses should be accompanied by clear learning objectives, a lesson plan, and a description of the methods selected for the training (TU Delft, 2019).

Looking forward, one can imagine integrated training for librarians and researchers that establishes RDM as the foundation for data-sharing workflows and other best practices of open science scholarly communications. The international [principles of FAIR data](#), findability, accessibility, interoperability and reuse, can be a shared method between cross-disciplinary open scholarship practices due to a common engagement with digital assets. As Higman et al. (2019) argue, “Researchers often want to be FAIR, and sometimes open; they are noble aspirations... By using the language of FAIR and open, we can engage people in data management too” (p. 2). The Bodleian Libraries at the University of Oxford offers a model of how the integration of RDM training with other areas of open scholarship might be achieved for

librarians. Library RDM services are led by one specialist who has developed an RDM training series for researchers addressing key issues, such as working with confidential data, secondary use of data, and data deposit and preservation. This training series is often team-taught with IT representatives or library staff with complementary expertise, highlighting the need for researchers to first contact their subject librarians with queries. RDM platforms are also supported by multiple members of library staff, not only the RDM specialist. The collaborative approach to RDM training for researchers, and a distributed technical RDM service “serves to reinforce the message of the training aimed at library staff, namely that RDM is an area that library staff across the board can support to some extent” (Southall & Scutt, 2017, p. 307). RDM training for librarians and library staff mirrors the content of training for researchers. There are two workshops that cover basic principles of RDM, trends in scholarly communications, and concrete examples of data management, with an emphasis placed on an “increased understanding of digital scholarship, RDM issues and where these sit in relation to the work of the academic library and new areas of scholarly activity such as Open Access (Southall & Scutt, 2017, p. 308).

Conclusion

The aim of this literature review of 13 sources, containing 14 case studies, was to survey a range of RDM training and capacity-building approaches, in order to determine the next steps for our own local context at the University of Ottawa. A method of looking at international training models was used in order to identify and supplement a gap in the emerging national RDM policy, infrastructure, and training environment. For instance, a notable challenge in the Canadian RDM training space is that many institutions have not yet developed the RDM Institutional Policies that are anticipated by the [Draft Tri-Agency Research Data](#)

[Management Policy](#) (2018). This layer of institutional strategy will enable the building of RDM into graduate-level curricula for both researchers and librarians. In the meantime, the next steps for building a training program at our local institution will begin with establishing a set of principles based on the findings of this literature review. The plan is to align these principles with the training methods, modes of assessment, and infrastructure development timeline outlined in a national training strategy anticipated for release in fall 2020 by the Portage Training Expert Group as a follow-up to the 2017 white paper, tentatively titled, *Building a Portage Network Training Strategy: A Canadian Approach to Research Data Management*.

The following trends emerged through this literature review, which have informed the national training strategy, and will be taken into consideration when building local training options for librarians, research support staff, and researchers at the University of Ottawa. Librarians and researchers must have enough incentive to undertake training in RDM or to join a community of practice. Training requires a significant investment of time, whether online or in-person, and librarians are unlikely to take on additional training, or to complete the training once enrolled, without a perceived benefit or reinforcement through regular RDM service provision. Disciplinary-specific instruction is preferable over general instruction for both librarians and researchers, however, a librarian's own training opportunities will influence their ability to provide discipline-specific RDM instruction to researchers. There is a double gap in the training landscape, as the lack of disciplinary-specific training opportunities for librarians further contributes to a lack of training options and service offerings for distinct research areas.

The range of pedagogical designs reflected in the case studies make it difficult to draw conclusions as to whether intensive events, or a

series of shorter time-commitments over a longer time period, is preferable for learning outcomes. In-person training opportunities emerged as the preferred option for learning retention and secondary effects of building a community of practice. For the same reasons, online instruction was found to be most effective when paired with an in-person component. The sources in this literature review predate the global COVID-19 pandemic, which has shifted higher-education into online delivery in historically unprecedented ways. This context may present an opportunity to apply the best practices of online learning design to close the gap between the benefits of in-person training and the low retention in online learning environments. Initiatives such as the University of British Columbia's [RDM Fall Series 2020](#) are early responses to virtual RDM instruction during the pandemic, demonstrating the importance of the adaptation to local contexts, for example. The literature review highlights the recommendation that in order to be meaningful, generalized RDM training offered by third parties must be adapted to local contexts. Discipline-specific training, in-person training, and adaptation to local contexts are all resource intensive activities but they are worth the investment. Librarians and other research support staff with disciplinary awareness will be more successful as they engage with researchers and help them to adopt research data management practices as an extension of their disciplinary peer communities. Finally, future directions for RDM training will be integrated into open access and digital scholarship awareness training, as well as cross-disciplinary, open science communities of practice that reach beyond local campuses.

Author Contributions Statement

Felicity Tayler: Conceptualization (lead), Methodology (lead), Writing – original draft (lead) review & editing (equal) **Maziar Jafary:** Methodology (supporting), Writing – original

draft (supporting), Writing - review & editing (equal)

content/uploads/2019/05/SR-Issue-Brief-Data-Communities-05132019.pdf

Acknowledgements

The authors would like to thank Chantal Ripp, Jane Fry, James Doiron, Lindsey Sikora and Kim Powroz for valuable feed-back on drafts of this review.

Fry, J., Doiron, J., Létourneau, D., Perrier, L., Perry, C., & Watkins, W. (2017). *Research data management training landscape in Canada: A white paper*, Prepared by the Portage Training Expert Group on behalf of the Canadian Association of Research Libraries (CARL). <http://dx.doi.org/10.14288/1.0372048>

References

4TU Research Data. (2020).

<https://researchdata.4tu.nl/en/>

Baker, J., Moore, C., Priego, E., Alegre, R., Cope, J., Price, L., Stephens, O., Strien, D. van, & Wilson, G. (2016). Library Carpentry: Software skills training for library professionals. *LIBER Quarterly*, 26(3), 141–162.

<https://doi.org/10.18352/lq.10176>

Grootveld, M. J., & Verbakel, E. (2015).

Essentials for data support: Training the front office. *International Journal of Digital Curation*, 10(1), 240-248.

<https://doi.org/10.2218/ijdc.v10i1.364>

Haddow, L. (2014, January 9). Training subject librarians in RDM. *Edinburgh Research Data Blog*.

<http://datablog.is.ed.ac.uk/2014/01/09/training-subject-librarians-in-rdm/>

Bryant, R., Lavoie, B., & Malpas, C. (2018). *Sourcing and Scaling University RDM Services* (The Realities of Research Data Management, Part 4). Dublin, OH: OCLC Research.

<https://www.oclc.org/research/publications/2017/oclcresearch-rdm-part-four-sourcing-scaling.html>

Helbig, K. (2016). Research data management training for geographers: First impressions. *ISPRS International Journal of Geo-Information*, 5(4), 1-9.

<https://doi.org/10.3390/ijgi5040040>

Clare, C., Cruz, M., Papadopoulou, E., Savage, J., Teperek, M., Wang, Y., Witkowska, I., & Yeomans, J. (2019). *Engaging researchers with data management: The cookbook*. Open Book Publishers.

<https://openresearchlibrary.org/content/03b1aed1-e57e-4305-ae7c-27c9d88b02fd/details>

Higman, R., Bangert, D., & Jones, S. (2019). Three camps, one destination: The intersections of research data management, FAIR and Open. *Insights*, 32(1), 1-9.

<https://doi.org/10.1629/uksg.468>

Cooper, D., & Springer, R. (2019, May 13). *Data communities: A new model for supporting STEM data sharing* (Issue Brief). Ithaca S+R. <https://sr.ithaka.org/wp->

Humphrey, C. (2020). The CARL Portage partnership story. *Partnership: The Canadian Journal of Library and Information Practice and Research*, 15(1), 1-7.

<https://doi.org/10.21083/partnership.v15i1.5825>

- Papadopoulou, E. & Miller, K. (2019). 'Dealing with Data' Conference at University of Edinburgh. In Clare, C., Cruz, M., Papadopoulou, E., Savage, J., Teperek, M., Wang, Y., Witkowska, I., & Yeomans, J., *Engaging researchers with data management: The cookbook* (pp. 70-72). Cambridge, UK: Open Book Publishers.
- Papadopoulou, E. & Grabauskiene, R. (2019). DuoDi: The 'Days of Data' at Vilnius University. In Clare, C., Cruz, M., Papadopoulou, E., Savage, J., Teperek, M., Wang, Y., Witkowska, I., & Yeomans, J., *Engaging researchers with data management: The cookbook* (pp. 74-76). Cambridge, UK: Open Book Publishers.
- Read, K. B., Larson, C., Gillespie, C., Oh, S. Y., & Surkis, A. (2019). A two-tiered curriculum to improve data management practices for researchers. *PLoS ONE*, 14(5), 1-14. <https://doi.org/10.1371/journal.pone.0215509>
- Shipman, J. P., & Tang, R. (2019). The collaborative creation of a Research Data Management Librarian Academy (RDMLA). *Information Services & Use*, 39(3), 243-247. <https://doi.org/10.3233/ISU-190050>
- Southall, J., & Scutt, C. (2017). Training for research data management at the Bodleian Libraries: National contexts and local implementation for researchers and librarians. *New Review of Academic Librarianship*, 23(2-3), 303-322. <https://doi.org/10.1080/13614533.2017.1318766>
- Surkis, A., & Read, K. (2015). Research data management. *Journal of the Medical Library Association*, 103(3), 154-156. <https://doi.org/10.3163/1536-5050.103.3.011>
- Tang, R., & Hu, Z. (2019). Providing research data management (RDM) services in libraries: Preparedness, roles, challenges, and training for RDM practice. *Data and Information Management*, 3(2), 84-101. <https://doi.org/10.2478/dim-2019-0009>
- TU Delft. (2019, October 25). Vision for Research Data Management Training at TU Delft. *Open Working*. <https://openworking.wordpress.com/2019/10/25/vision-for-research-data-management-training-at-tu-delft/>
- Wenger-Trayner, E., & Wenger-Trayner, B. (2015). *Communities of practice: A brief introduction*. Wenger-Trayner. <https://wenger-trayner.com/wp-content/uploads/2015/04/07-Brief-introduction-to-communities-of-practice.pdf>
- Wittenberg, J., Sackmann, A., & Jaffe, R. (2018). Situating expertise in practice: Domain-based data management training for liaison librarians. *The Journal of Academic Librarianship*, 44(3), 323-329. <https://doi.org/10.1016/j.acalib.2018.04.004>