

It Takes a Researcher to Know a Researcher: Academic Librarian Perspectives Regarding Skills and Training for Research Data Support in Canada

Alisa B. Rod

Volume 18, numéro 2, 2023

URI : <https://id.erudit.org/iderudit/1101594ar>
DOI : <https://doi.org/10.18438/eblip30297>

[Aller au sommaire du numéro](#)

Éditeur(s)

University of Alberta Library

ISSN

1715-720X (numérique)

[Découvrir la revue](#)

Citer cet article

Rod, A. (2023). It Takes a Researcher to Know a Researcher: Academic Librarian Perspectives Regarding Skills and Training for Research Data Support in Canada. *Evidence Based Library and Information Practice*, 18(2), 44–58. <https://doi.org/10.18438/eblip30297>

Résumé de l'article

Objective – This empirical study aims to contribute qualitative evidence on the perspectives of data-related librarians regarding the necessary skills, education, and training for these roles in the context of Canadian academic libraries. A second aim of this study is to understand the perspectives of data-related librarians regarding the specific role of the MLIS in providing relevant training and education. The definition of a data-related librarian in this study includes any librarian or professional who has a conventional title related to a field of data librarianship (i.e., research data management, data services, GIS, data visualization, data science) or any other librarian or professional whose duties include providing data-related services within an academic institution.

Methods – This study incorporates in-depth qualitative empirical evidence in the form of 12 semi-structured interviews of data-related librarians to investigate first-hand perspectives on the necessary skills required for such positions and the mechanisms for acquiring and maintaining such skills.

Results – The interviews identified four major themes related to the skills required for library-related data services positions, including the perceived importance of experience conducting original research, proficiency in computational coding and quantitative methods, MLIS-related skills such as understanding metadata, and the ability to learn new skills quickly on the job. Overall, the implication of this study regarding the training from MLIS programs concerning data-related librarianship is that although expertise in metadata, documentation, and information management are vital skills for data-related librarians, the MLIS is increasingly less competitive compared with degree programs that offer a greater emphasis on practical experience working with different types of data in a research context and implementing a variety of methodological approaches.

Conclusion – This study demonstrates that an in-depth qualitative portrait of data-related librarians within a national academic ecosystem provides valuable new insights regarding the perceived importance of conducting original empirical research to succeed in these roles.

© Alisa B. Rod, 2023



Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

éru
dit

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

<https://www.erudit.org/fr/>



Research Article

It Takes a Researcher to Know a Researcher: Academic Librarian Perspectives Regarding Skills and Training for Research Data Support in Canada

Alisa B. Rod
Research Data Management Specialist
McGill University
Montreal, Quebec, Canada
Email: alisa.rod@mcgill.ca

Received: 4 Jan. 2023

Accepted: 7 Feb. 2023

© 2023 Rod. 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/eblip30297](https://doi.org/10.18438/eblip30297)

Abstract

Objective – This empirical study aims to contribute qualitative evidence on the perspectives of data-related librarians regarding the necessary skills, education, and training for these roles in the context of Canadian academic libraries. A second aim of this study is to understand the perspectives of data-related librarians regarding the specific role of the MLIS in providing relevant training and education. The definition of a data-related librarian in this study includes any librarian or professional who has a conventional title related to a field of data librarianship (i.e., research data management, data services, GIS, data visualization, data science) or any other librarian or professional whose duties include providing data-related services within an academic institution.

Methods – This study incorporates in-depth qualitative empirical evidence in the form of 12 semi-structured interviews of data-related librarians to investigate first-hand perspectives on the necessary skills required for such positions and the mechanisms for acquiring and maintaining such skills.

Results – The interviews identified four major themes related to the skills required for library-related data services positions, including the perceived importance of experience conducting original research, proficiency in computational coding and quantitative methods, MLIS-related skills such as understanding metadata, and the ability to learn new skills quickly on the job. Overall, the implication of this study regarding the training from MLIS programs concerning data-related librarianship is that although expertise in metadata, documentation, and information management are vital skills for data-related librarians, the MLIS is increasingly less competitive compared with degree programs that offer a greater emphasis on practical experience working with different types of data in a research context and implementing a variety of methodological approaches.

Conclusion – This study demonstrates that an in-depth qualitative portrait of data-related librarians within a national academic ecosystem provides valuable new insights regarding the perceived importance of conducting original empirical research to succeed in these roles.

Introduction

In the context of academic libraries, there has been increasing demand from researchers for data-related services over the past ten years (Barsky, 2019; Barsky et al., 2017; Cox et al., 2017; Cox et al., 2019a; Steeleworthy, 2014). This shift is projected to continue a growth path, which has implications for increased capacity and infrastructure needs at all levels of service provision (Briney et al., 2015; Khair et al., 2020). Academic libraries have responded to increasing demand for research data support by developing capacity in this regard and positioning the library as the appropriate centralized support resource (Ashiq & Warraich, 2022).

In this way, academic libraries have spearheaded service models for the provision of data-related research services (Radecki & Springer, 2020). In the academic library context more generally, new positions focused on data have emerged over the past ten years, including positions focusing on research data services (RDS), data curation, and research data management (RDM). RDM, while closely related to RDS, is more focused on organizing and managing research data over a lifecycle rather than providing reference support related to finding sources for data or accessing data held in controlled collections, which is more the purview of RDS. Obscuring the notion of a clearcut definition of these roles is the lack of standardized titles and the conflation of research data support with liaison duties at many institutions. For example, a recent study by Theilen and Neeser (2020) identified 119 discrete job titles for academic library job postings, of which the majority did not include a reference to librarianship, although the most frequently occurring librarian job titles reflect the four most common roles, including data services librarian, data curation librarian, research data management librarian, and data librarian. Thus, for the purposes of this study, the broader term “data-related librarian” is used to represent any librarians or professionals who offer support for research data within an academic context.

There are several additional studies that have investigated the competencies, skills, and educational background required in job advertisements for these and related roles in academic libraries both in North American and globally (ACRL Research Planning and Review Committee, 2020; Cox et al., 2017; Federer, 2018; Fuhr, 2019; Goben & Sapp-Nelson, 2018a; Goben & Sapp-Nelson, 2018b; Thielen & Neeser, 2020; Xia & Wang, 2017). However, it is unclear as to whether job postings are indicative of the profiles of individuals who are performing these roles. In addition, there are also existing studies that have

investigated the self-perceptions of librarians regarding technical readiness and confidence in assuming these roles (Chiware, 2020; Ducas et al., 2020; Joo & Schmidt, 2021; Tang & Hu, 2019; Thomas & Urban, 2018). However, these studies typically rely on survey methodology, which is useful in identifying broad trends but less capable of identifying nuance or providing entirely novel insights.

Aims

This study aims to contribute a qualitative analysis of the perspectives of data-related librarians regarding the necessary skills, experience, and training for these roles. This study is organized around the following two research questions:

RQ1: What are data-related librarians' perspectives regarding the skills required to work effectively as academic data professionals?

RQ2: What are data-related librarians' perspectives regarding the role of the MLIS in providing relevant training and education to prepare academic data professionals?

This study incorporates in-depth qualitative empirical evidence in the form of 12 semi-structured interviews of data-related librarians to investigate first-hand perspectives on the necessary skills and expertise required for such positions and the mechanisms for acquiring and maintaining such skills.

Literature Review

Over the past two decades, technological advances and institutional investments related to digital research infrastructures have rendered the collection of larger and larger quantities of data as ubiquitous to many academic disciplines and research processes (Briney et al., 2015). Thus, research communities and public funding agencies have articulated and socialized best practices for handling data throughout a research projects' lifecycle, culminating with the development and adoption of the FAIR principles as the primary guiding framework for research data (Force11, 2014; Tang & Hu, 2019; Wilkenson et al., 2016). FAIR refers to data that are findable, accessible, interoperable, and re-usable, which may be implemented primarily through infrastructure, such as digital data repositories, designed for enhancing the discoverability, preservation, and sharing of research data. Following a consultation period lasting almost 10 years, the three major funding agencies in Canada, the Tri-Agencies, recently published a policy on RDM highlighting the FAIR principles and focusing on the proper handling of data as a matter of integrity and ethics (Government of Canada, 2021). These large-scale shifts in norms and policies governing accountability and reproducibility regarding research data have necessitated parallel shifts among higher education institutions related to the provision of support and the investment in staffing capacity as a crucial pillar of a modern academic digital research infrastructure program (Baxter et al., 2021; Federer et al., 2020).

In most cases, the library is the campus unit that has been designated with a mandate to provide data-related support services, training, and guidance (Radecki & Springer, 2020). Indeed, the library is the natural home for data-related digital research infrastructure management as this emerging function intersects with other key contemporary academic library roles, including scholarly communications, digital collections more generally, digital preservation, and digital scholarship (ACRL Research Planning and Review Committee, 2020; Steeleworthy, 2014). In addition, the liaison model provides a support system that is embedded within conventional research processes and offers a communication channel by which to facilitate awareness and outreach campaigns (Steeleworthy, 2014). Thus, academic libraries have

spent the past decade investing in capacity-building in terms of technical infrastructure and highly qualified personnel (HQP) in the form of data-related librarian positions or units such as RDS that incorporate several data-related librarian positions.

Several previous studies have mapped both the change in academic library positions regarding data services and the skills required for these and adjacent positions over the past decade (ACRL Research Planning and Review Committee, 2020; Goben & Sapp-Nelson, 2018a; Goben & Sapp-Nelson, 2018b; Tenopir et al., 2015; Tenopir et al., 2019; Thielen & Neeser, 2020). In the past few years, a shift has occurred regarding the educational requirements for data-related librarian positions (Chen & Zhang, 2017; Theilen & Neeser, 2020). A study by Chen and Zhang (2017) found that fewer than 30% of data management academic library job postings between January and April 2015 required the MLIS and that an alternative relevant advanced degree would be accepted. A more recent study by Theilen and Neeser (2020) reviewed postings between 2013 and 2018 for RDS positions within academic libraries. In general, Theilen and Neeser (2020) claim that the requirements and preferred qualifications for data professional job advertisements favors the education and experience of candidates with a non-LIS background. A recent study by Fuhr (2022) also found that hiring for early career RDS librarians is trending toward candidates who do not hold an MLIS. Chen and Zhang (2017) similarly speculated that LIS programs may not be preparing students with expertise in data management or that the relevance of LIS curriculum to data management professions, even in the context of academic libraries, is becoming less apparent. In general, continuing education programs and professional development offerings have emerged to supplement formal LIS education (Davis & Cross, 2015; Read et al., 2019).

In addition to educational requirements, several studies have also focused on the skills and/or competencies required for data-related librarian positions. A recent study by Federer (2018) detailed 47 skills and competencies organized according to a taxonomy of nine categories. The overarching categories of skills, knowledge, and competencies required for data librarianship, as identified by Federer (2018), included data management, technology and information technology, evaluation and assessment, teaching and instruction, marketing and outreach, library skills, professional involvement, skills and personal attributes, and education and training. Many of these areas of skills and competencies overlap with other functional or subject-specific librarian roles. The data-specific skills, knowledge, and competencies cover a range of applications of technological and functional abilities and training, including data visualization, scientific programming, geographic information systems (GIS) data and programs, discipline-specific data management skills, and knowledge of best practices regarding data sources, finding data, sharing data, curating data, and developing service models for data support.

In a more recent review, Fuhr (2019) synthesizes the literature in terms of suggested or required skills for RDM-related information professionals, finding that previous research identified skills related to the full research data lifecycle, including planning for the collection and use of data, actively managing data in terms of storage and security, documenting data and incorporating relevant metadata at all phases of a research project, and planning for the dissemination or archiving of data in the long-term. In addition, several studies also identify “soft skills” such as effective communication, team-oriented professionalism, and relationship building to be cited across job postings or identified by practitioners (Chen & Zhang, 2017; Federer, 2018). Other research has also found an increasing emphasis on familiarity or expertise in a programming language or statistical software, GIS, or data visualization (Fuhr, 2022).

In this way, the technological and functionally specific skills required to acquire and perform a data-related role within an academic library context may not map onto the conventional LIS curriculum. Indeed, several recent studies claim that these competencies are not easily acquired through LIS curricula

(Chen & Zhang, 2017; Si et al., 2013; Thielen & Neeser, 2020). As RDS was emerging as a support area within academic libraries, a handful of MLIS-granting institutions developed pathways or certificates in data curation or related topics embedded within the broader LIS curriculum (Corrall, 2012). Currently, courses in RDS, RDM, and related topics are still not widely integrated within LIS programs, although a systematic review of curricula is lacking (Chen & Zhang, 2017). However, a recent study by Wang and Lin (2019) provided an empirical review of the 2018-2019 academic year course offerings of 48 iSchools across the US and found that only 35% of programs offered any courses related to RDS, of which fewer than 10% of the total course offerings were related to RDS.

There are also several recent studies incorporating methodologies to directly investigate the self-perceptions of RDS and RDM librarians regarding the skills and qualifications necessary to perform their roles (Chiwari, 2020; Ducas et al., 2020; Joo & Schmidt, 2021; Tang & Hu, 2019; Thomas & Urban, 2018). Thomas and Urban (2018) surveyed 105 RDS professionals and asked directly for participants to indicate the extent to which their MLIS degree program prepared them for their position. The primary takeaway from the Thomas and Urban (2018) study is that RDS librarians learn about topics that are less conventionally covered in LIS curricula, such as data management, through hands on experience and thus perceive MLIS degree programs to be ill-equipped at preparing professionals for data-intensive librarian roles. A more recent study by Ducas et al. (2020) surveyed 205 librarians in Canada about a range of emerging roles for academic libraries including data management and data curation. Ducas et al. (2020) found that most participants reported needing additional training in several emerging librarian job functions, including data curation, statistical methods and programs, and data management.

Although several recent studies have surveyed librarians in RDS positions, most studies on the background and skills required for data-related librarian positions tend to focus on job postings rather than individuals hired into the positions. Although surveys do provide some insight into the disconnect between educational pathways, job postings, and the actual demands of job duties related to research data, they typically rely on sample sizes that are not adequate for representing larger populations. In general, there has not been an in-depth investigation into the perceptions of information professionals regarding the necessary skills to perform these data-related librarian roles, broadly defined, in the Canadian context, which relies on a nationally integrated infrastructure and support network including federal stakeholder organizations, such as Borealis, The Digital Research Alliance of Canada (the Alliance), and the Canadian Research Knowledge Network (Ducas et al., 2020). Thus, a qualitative in-depth investigation of individuals holding data-related librarian positions in the context of Canadian academic libraries would be a major contribution to this body of literature.

Methodology

The data collection for this study involved conducting 12 semi-structured interviews with data-related librarians at Canadian higher education academic institutions. The interview research described in this article was approved by McGill University's Research Ethics Board (REB File #: 22-01-077). One major methodological challenge of this study was determining an operationalization for the construct of data-related librarian. Previous studies found that data-related librarians have a range of unique titles. Thus, the operationalization of data-related librarian includes any librarian or professional who has a conventional title related to a field of data librarianship (i.e., research data management, data services, GIS, data visualization, data science) or any other librarian or professional whose duties align with librarian job descriptions of data-related services within an academic institution (Springer, 2019). I

randomly selected 13 individuals from a list of 253 data-related librarians and professionals compiled in accordance with this conceptualization (Rod, 2022).

I chose to use random selection to minimize the chances of participant re-identification and to be able to provide participants, if requested, with the precise statistical likelihood of their selection. I intended to sample between 10 and 15 in stages in anticipation that I would need to send multiple batches of invitations. However, 12 of the 13 invited participants from the initial sample agreed to the interview. The interview participants represent a range of roles (see Table 1 for proportions of titles of interview participants and the Appendix for examples of specific job titles for each aggregated category). In addition, the interview participants were located across several provinces, including Ontario, Québec, one Prairie province, and one Atlantic province.

Table 1
Aggregated Titles of Interview Participants

Aggregated Titles	Frequency	Percent of Total Title Categories
Data or GIS Librarian	4	33%
Advisors or Directors	2	16%
Liaison Librarian	3	25%
RDM Librarian or Specialist	3	25%
Total	12	100%

The 30-minute interviews were booked via Microsoft Bookings and took place virtually via Microsoft Teams in the summer of 2022, between June 22 and July 28. The interview questions focused on participants' experience, education, and their perspectives on relevant training and expertise required for these positions. For a full list of interview questions and codebook, see the deposited dataset (Rod, 2023). Interviews were audio recorded and auto-transcribed in Microsoft Word. I reviewed and manually corrected each interview transcription. Overall, the auto-transcriptions were fairly accurate.

The interview transcripts were analyzed using qualitative content analysis, which is an established qualitative methodology for summarizing a series of observations drawn from unstructured data either by human coding or using a text analysis software or program (Bernard et al., 2016). Qualitative content analysis is an iterative process in which one or more coders read through the textual data, usually more than one time to reconcile disagreements and to create and refine a codebook (Creswell, 1994). Once a codebook is developed, the coder(s) re-read the text data and apply the codes to specific words, sentences, or paragraphs. The unit of analysis for this study is the individual participant, meaning that I am comparing across participants. Thus, after collecting all references to each code within a single interview transcript, I aggregated unique codes by each interview.

Following the completion of the interviews, I read a sample of 2 full transcripts to develop an initial codebook. I then read 5 different interview transcripts to adjust and refine the codebook. Finally, I read all interview transcripts, including those that had been previously analyzed, to develop a final version of

the codebook. Once a finalized codebook was established, I re-read and coded all interview transcripts using Taguette, a desktop application for tagging textual data with codes.

These interviews were conducted as part of a larger project on the skills, education, and the responsibilities of data-related librarians in the Canadian context (Rod, 2022). The results reported in the following section reflect an analysis of a portion of the information collected from the interviews that focused on data-related librarians' perspectives of relevant skills and experience for these roles.

Results

To address the first research question, interview participants were asked to reflect on the most essential skills required for roles or positions like theirs at academic libraries. To maintain participants' privacy, pseudonyms are employed to describe results. Interestingly, only two participants indicated that their job duties are clearly defined. Most interview participants indicated that they are, to varying degrees, unclear where the boundaries are between their job and other data-related roles at their library. As summarized by Victoria, an RDM librarian or specialist at a large research university:

What is the difference between the [number] of us? We have no idea. It's never been defined or clarified with us and no matter how many times we ask we never get an answer, so I couldn't tell you what technically the differences are between our roles and what we're supposed to do, and not supposed to do other than what our titles are.

Relatedly, one-third of the interview participants indicated that their data-related duties are done "off the side of their desk" and are split between another position type, such as liaison/subject librarian. In terms of the specific tasks that are associated with library-related data services positions, most interview participants mentioned that they are not an RDM librarian, but that they are responsible for RDM-related tasks as part of their job. The most reported job tasks across all interview participants included data-related reference with individuals, teaching/instruction quantitative or numerical data analysis (e.g., with Excel, R, Python, etc.), facilitating data sharing and deposit for researchers, and consulting or training on writing data management plans (DMPs). Interestingly, many participants are also involved in RDM-related initiatives at the strategic planning level of their institution. This is due to the recent Tri-Agency RDM Policy, which requires institutions administering public research funds in Canada to develop and share an institutional RDM strategy by March 1, 2023 (Government of Canada, 2021). Less commonly, interview participants mentioned that, as part of their job, they are involved as principal investigator (PI) or related staff at the project-level of sponsored research. In addition, fewer than one-third of participants indicated that they perform tasks related to text analysis, outreach, and/or qualitative analysis.

Overall, the findings of the interviews identified four major themes related to the skills required for library-related data services positions, including:

- Experience conducting original research;
- Proficiency in computational coding and quantitative methods;
- MLIS-related skills such as understanding metadata, documentation, preservation, and curation;
- The ability to learn new skills quickly on the job.

First, the interviews highlighted the perceived importance of conducting original research. Ten of the twelve interview participants explicitly discussed the importance of conducting original research as a qualification for a data-related librarian position. Across these ten interview participants, several

participants noted that conducting research helps to bridge the gap between researchers/faculty and librarians/library-related data services staff. As Anna, a data or GIS Librarian at a large research university summarized “if you’ve done some research in the past, I think that’s extremely helpful. If you’ve analyzed data in some way, [if] you’ve written a paper, then you’ve done something that’s going to help you in this job.”

Relatedly, the second theme identified across the interviews is proficiency in computational methods (e.g., coding), quantitative methods, and data literacy. Most interview participants discussed the importance of working with various types of data, various quantitative methodologies, and having an advanced understanding of the research data lifecycle. More than half of the interview participants mentioned the importance of technical skills related to different types of data and familiarity with a wide range of data types and tools.

The third theme that emerged across most interview participants relates to skills that are developed through training in MLIS programs and coursework. Participants consistently indicated that skills and expertise acquired through the standard MLIS curriculum, such as understanding metadata, documentation, preservation, and curation, are the bedrock of succeeding at data-related librarian positions. For example, Lily, a data or GIS Librarian, discussed the value of having expertise in thinking critically about empirical information divorced of disciplinary biases and argued that acquiring this expertise is unique to MLIS programs.

The fourth theme that emerged across interview participants relates to the importance of the meta-skill of acquiring new skills and expertise quickly on the job. Interestingly, interview participants weighed the soft skill of gaining new expertise quickly as relatively equally important as hard skills such as technical proficiency in data analysis programs. For example, Claudia, an RDM librarian or specialist, discussed a project where they were asked to provide advice on data sharing for a medical research project. Although they were not familiar with medical research, they were able to learn quickly and apply research principles that are cross-cutting. Other participants defined learning quickly on the job in terms of the importance of understanding the high-level principles of coding and data analysis so that when trends regarding specific tools (e.g., SPSS, R, Python, etc.) change, it is easier to adapt and evolve. Jake, a liaison librarian with data responsibilities, provided an analogy to explain that data-related librarian positions are not inherently or necessarily different from other types of librarian roles:

I’m going to pick on [law librarians] specifically 'cause I know they tend to have this idea of: “No, you should not be a law librarian unless you are a lawyer, and this is very much a world that nobody else can ever understand unless you have this training.” I go in the exact opposite direction. You could come into data librarianship from anywhere and really, it’s more about having the fluency of knowing how to understand different ways of talking about things.

Indeed, five interview participants expressed the view that having specific skills and expertise, including the ability to learn new skills quickly, is more important than any specific disciplinary background, work experience, or formal academic training.

To address the second research question regarding the importance of the MLIS in providing relevant training and education for these positions, interview participants were asked about their educational background and previous work experience. Seven of the interview participants mentioned that their work experience as a librarian translated directly into their acquisition of relevant skills and expertise. These seven participants indicated that they had worked as data-related librarians in a previous

position(s) or have already been working in the same data-related position for a significant number of years. In addition, half of the participants explicitly mentioned that obtaining an MLIS degree facilitated the acquisition of many relevant skills and expertise.

When asked to reflect on the ideal educational background for information professionals working in data-related librarian positions, nine of the interview participants discussed the importance of the MLIS degree. Overall, the interview participants were highly supportive of the MLIS, while acknowledging that this perspective is self-serving. However, another consistent theme that emerged across the participants was that the MLIS does not incorporate enough data-specific courses to facilitate gaining advanced proficiency in the relevant skills required for data-related librarians. As articulated by Lily, one issue identified is that:

People don't understand what librarians do and there's also just a huge variety of things that librarians do, so I think people sometimes wonder...why couldn't someone with a PhD in that field do that better than you? My response to that is... Just because someone has, for example, like in my field in [STEM], someone who has a [STEM] PhD isn't necessarily going to know how databases work within the sciences...or how to manage that data in ways that preserve that data for the future.

On the other hand, three participants argued that a master's degree in a quantitative field, or a PhD, should substitute or augment an MLIS specifically for data-related librarian positions. As Claudia argued, "in my opinion it's not possible to respond to faculty needs without an advanced degree at the level of a PhD." In this way, the interviews offer views of how information professionals working in data-related librarian positions in Canada have acquired the relevant skills that are mostly reflective of the interviewee's own experiences. However, the interview participants were generally still divided in terms of their perspectives regarding the relevance of specific educational backgrounds or prior work experience.

Discussion

The findings of this qualitative study reinforce several emerging trends identified within recent literature. For example, recent research has echoed the finding that there is increasingly less of an emphasis on the MLIS and that data-related librarians offer higher levels of technical assistance relative to the state of the field ten years ago (Fuhr, 2022; Plassche, 2022). Overall, the themes that emerged from the interviews provides evidence that there are at least three key implications for data-related librarianship training and skills acquisition.

First, interview participants identified the importance of data-related librarians to have first-hand experience conducting empirical research. Several participants argued that research experience should be a requirement for data-related librarian positions compared with liaison/subject positions. The overall rationale for this perspective is that academic researchers (e.g., non-librarian faculty members, professors, etc.), whether warranted or not, do not respect the advice of colleagues who do not "understand" the full research project lifecycle and all that it entails (e.g., writing grants and data management plans, collecting data in compliance with institutional and national privacy policies, storing data, sharing data among collaborators, analyzing empirical data, and writing up findings based on their analysis). According to this perspective, the only way to gain this understanding is to go through the process of conducting research. This perspective in many ways is antithetical to the discipline and conventional professional

practice of librarianship, where meta-knowledge of the organization of information can be applied across domains without necessarily having experience in each domain or in-depth disciplinary expertise.

Of course, several participants argued that it is possible to gain this experience through an MLIS program, either through conducting a short original research project within the context of a course or through a multi-term thesis or project. However, other participants argued that professors and researchers who have gone through the process of acquiring doctoral degrees are biased regarding the importance of conducting original research at that level. In addition, participants also argued that the methodological and research training in MLIS programs are insufficient and are not often required as part of a core curriculum. As Roberta, an advisor or director who previously worked as a librarian supporting RDM and RDS services, reflected,

If they said, “you need a PhD or you need a background exactly in data science or something like that,” then I would never have gotten this position. Maybe that's okay too. Maybe they would have gotten someone who has all of that and it would be even better for the community.

A second key finding of this study is that although there is empirical evidence, in addition to the findings from other recent studies, which depicts a general slight decline in the requirement of an MLIS for data-related librarian positions, data-related librarians in Canada generally still view the MLIS as valuable. This is consistent with a recent study by Plassche (2022) finding that map and geospatial library positions are increasingly requiring domain-specific degrees (e.g., in geography, GIS, or a related field) and technical backgrounds, but still strongly favor an MLIS. Interview participants were highly supportive of the MLIS even while acknowledging that the ability to pursue data-related courses within MLIS programs greatly varies and curricula are generally currently misaligned to the actual demands of data-related librarian positions. Overall, the implication of this study regarding the training from MLIS programs for data-related librarianship is that although expertise in metadata, documentation, and information management are vital skills for data-related librarians, the MLIS is increasingly less competitive compared with degree programs that offer a greater emphasis on practical experience working with different types of data in a research context and implementing a variety of methodological approaches (e.g., domain-specific empirical master's degrees or doctoral degrees).

A final key finding from this study aligns with previous research identifying the particular importance of soft skills (e.g., communication, learning quickly, networking, leadership, etc.) for data-related librarian positions (Chen & Zhang, 2017; Harp & Ogborn, 2019; Federer, 2018). Interview participants in this study generally reported that learning quickly on the job is a top skill for data-related librarians, given that educational pathways to these positions may be diverse and the landscape defining vital technical skills changes rapidly.

In addition, as reinforced by Cox et al. (2019b), the rise of funder mandates requiring better data management across disciplines has carved a role for data-related librarians regarding strategic planning at the institutional level. For example, in 2021, the three major public funding agencies in Canada, or the Tri-Agency, released a harmonized policy on RDM for which one requirement is that higher education institutions that administer Tri-Agency funds develop and publicly post an institutional RDM strategy by March 1, 2023. The development of these institutional RDM strategies has facilitated cross-functional collaborations and discussions among research administration, the library, IT departments, and other campus units relevant to the management of research data. In this way, data-related librarians in Canada

have increased opportunities to employ networking skills in developing and formalizing relationships with various campus stakeholders (see also Pinfield et al., 2014; Harp & Ogborn, 2019).

A key limitation of this study is that it is focused on Canadian data-related librarians, and thus may not reflect the perspectives of data-related librarians in other geographic communities. Future research could expand this methodology to conduct qualitative research across a variety of geographic regions and contexts. A second limitation involves reliance on a single coder, thus potentially introducing measurement bias. However, this is mitigated by the subject and methodological expertise and experience of the coder, in addition to the number of iterations of coding, and the qualitative nature of the study, which is not inherently designed to maximize replicability, but rather to uncover new information that would not be easily surfaced using other methods or approaches.

Conclusion

This study demonstrates that an in-depth qualitative portrait of data-related librarians within a national academic ecosystem provides valuable new insights regarding the perceived importance of the specific training and skills required to succeed in these roles. The interviews identified four major themes related to the skills required for library-related data services positions in Canada, including the perceived importance of experience conducting original research, proficiency in computational coding and quantitative methods, MLIS-related skills such as understanding metadata, and the ability to learn new skills quickly on the job. Overall, the implication of this study regarding the training from MLIS programs regarding data-related librarianship is that although expertise in metadata, documentation, and information management are vital skills for data-related librarians, the MLIS is increasingly less competitive compared with more technical (e.g., programs in STEM fields) or research-oriented degrees (e.g., empirical PhD programs). A potential future line of inquiry could involve investigating the opportunities that exist within MLIS programs for conducting original research and what those research projects or course components entail.

In general, this study contributes to the literature on data-related librarianship in terms of providing new qualitative evidence on perspectives regarding the skills, education, and training required for these roles. In addition, this study provides a valuable methodological approach for conceptualizing and operationalizing data-related librarians according to inclusive parameters (e.g., operationalizing data librarian to include any librarian with responsibilities related to RDM, RDS, data visualization, data science, GIS, etc. in a single analysis) and can be applied to other contexts or geographical regions.

Acknowledgement

This manuscript draws on data collected as part of a Master's degree research project for the McGill University Master of Information Studies program. The author wishes to thank her project supervisor, Dr. Rebekah Willson, for providing valuable feedback on previous versions of this manuscript.

References

- ACRL Research Planning and Review Committee. (2020). 2020 top trends in academic libraries: A review of the trends and issues affecting academic libraries in higher education. *College & Research Libraries News*, 81(6), 270-278. <https://doi.org/10.5860/crln.81.6.270>
- Ashiq, M., & Warraich, N. F. (2022). A systematized review on data librarianship literature: Current services, challenges, skills, and motivational factors. *Journal of Librarianship and Information Science*. Advance online publication. <https://doi.org/10.1177/09610006221083675>
- Barsky, E. (2019, June 20). *Research data management: A view from Canada* [presentation]. <http://dx.doi.org/10.14288/1.0379514>
- Barsky, E., Laliberté L., Leahey, A., and Trimble, L. (2017). Collaborative research data curation services: A view from Canada. In L. R. Johnston (Ed.), *Curating research data, volume one: Practical strategies for your digital repository* (pp. 79-101). Association of College and Research Libraries.
- Baxter, G., Beard, L., Beattie, G., Blake, M., Greenhall, M., Lingstadt, K., Nixon, W., & Reimer, T. (2021). Covid-19 and the future of the digital shift amongst research libraries: an RLUK perspective in context. *New Review of Academic Librarianship*. Advance online publication. <https://doi.org/10.1080/13614533.2021.1976232>
- Bernard, H. R., Wutich, A., & Ryan, G. W. (2016). *Analyzing qualitative data: Systematic approaches*. Sage Publications.
- 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 and Scholarly Communication* 3(2), eP1232. <https://doi.org/10.7710/2162-3309.1232>
- Chen, H. L., & Zhang, Y. (2017). Educating data management professionals: A content analysis of job descriptions. *The Journal of Academic Librarianship*, 43(1), 18-24. <https://doi.org/10.1016/j.acalib.2016.11.002>
- Chiware, E. R. T. (2020). Data librarianship in South African academic and research libraries: a survey. *Library Management*, 41(6/7), 401-416. <https://doi.org/10.1108/LM-03-2020-0045>
- Corrall, S. (2012). Roles and responsibilities: Libraries, librarians, and data. In G. Pryor (Ed.), *Managing Research Data*. (pp. 105-133). Facet.
- Cox, A. M., Kennan, M. A., Lyon, L., & Pinfield, S. (2017). Developments in research data management in academic libraries: Towards an understanding of research data service maturity. *Journal of the Association for Information Science and Technology*, 68(9), 2182-2200. <https://doi.org/10.1002/asi.23781>
- Cox, A. M., Kennan, M. A., Lyon, L., Pinfield, S., & Sbaffi, L. (2019a). Maturing research data services and the transformation of academic libraries. *Journal of Documentation*, 75(6), 1432-1462. <https://doi.org/10.1108/JD-12-2018-0211>

- Cox, A. M., Kennan, M. A., Lyon, L., Pinfield, S., & Sbaffi, L. (2019b). Progress in research data services: An international survey of university libraries. *International Journal of Digital Curation*, 14(1), 126-135. <https://doi.org/10.2218/ijdc.v14i1.595>
- Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Sage.
- Davis, H. M., & Cross, W. M. (2015). Using a data management plan review service as a training ground for librarians. *Journal of Librarianship and Scholarly Communication*, 3(2), eP1243. <https://doi.org/10.7710/2162-3309.1243>
- Ducas, A., Michaud-Oystryk, N., & Speare, M. (2020). Reinventing ourselves: New and emerging roles of academic librarians in Canadian research-intensive universities. *College & Research Libraries*, 81(1), 43. <https://doi.org/10.5860/crl.81.1.43>
- Federer, L. (2018). Defining data librarianship: a survey of competencies, skills, and training. *Journal of the Medical Library Association*, 106(3), 294-303. <https://doi.org/10.5195/jmla.2018.306>
- Federer, L., Clarke, S. C., & Zaringhalam, M. (2020, January 16). *Developing the librarian workforce for data science and open science*. OSF Preprints. <https://doi.org/10.31219/osf.io/uycax>
- Force11. (2014). *Guiding principles for findable, accessible, interoperable and re-usable data publishing version b1.0*. <https://www.force11.org/fairprinciples>
- Fuhr, J. (2019). "How do I do that?" A literature review of research data management skill gaps of Canadian health sciences information professionals. *Journal of the Canadian Health Libraries Association/Journal de l'Association des bibliothèques de la santé du Canada*, 40(2), 51-60. <https://doi.org/10.29173/jchla29371>
- Fuhr, J. (2022). Developing data services skills in academic libraries. *College & Research Libraries*, 83(3), 474-502. <https://doi.org/10.5860/crl.83.3.474>
- Goben, A., & Sapp Nelson, M. (2018a). The data engagement opportunities scaffold: Development and implementation. *Journal of eScience Librarianship*, 7(2), e1128. <https://doi.org/10.7191/jeslib.2018.1128>.
- Goben, A., & Sapp Nelson, M. (2018b). Teaching librarians about data: The ACRL research data management roadshow. *College & Research Libraries News*, 79(7), 354-357. <https://doi.org/10.5860/crln.79.7.354>
- Government of Canada. (2021). *Tri-Agency research data management policy*. https://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html
- Harp, M. R. & Ogborn, M. (2019). Collaborating externally and training internally to support research data services. *Journal of eScience Librarianship*, 8(2), e1165. <https://doi.org/10.7191/jeslib.2019.1165>
- Joo, S., & Schmidt, G. M. (2021). Research data services from the perspective of academic librarians. *Digital Library Perspectives*, 37(3), 242-256. <https://doi.org/10.1108/DLP-10-2020-0106>

- Khair, S., Dara, R., Haigh, S., Leggott, M., Milligan, I., Moon, J., Payne, K., Portales-Casamar, E., Roquet, G., & Wilson, L. (2020). *The current state of research data management in Canada: An update to the LCDRI data management position paper*. New Digital Research Infrastructure Organization (NDRIO). https://engagedri.ca/assets/documents/NDRIO_RDM_PositionPaper.pdf
- Pinfield, S., Cox, A. M., & Smith, J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLoS ONE*, 9(12), e114734. <https://doi.org/10.1371/journal.pone.0114734>
- Plassche, K. A. (2022). Evaluating map and geospatial academic library position descriptions. *College & Research Libraries*, 83(4), 631-645. <https://doi.org/10.5860/crl.83.4.631>
- Radecki, J. & Springer, R. (2020, November 18). *Research data services in US higher education* [research report]. Ithaka S+R. <https://sr.ithaka.org/publications/research-data-services-in-us-higher-education/>
- Rod, A.B. (2022). *A Mixed-Methods Study of Data Professionals in the Canadian Academic Library Ecosystem* [Unpublished manuscript]. School of Information Studies, McGill University.
- Rod, A.B. (2023). *A Qualitative Study of Academic Data-Related Librarians in Canada* (Version 1) [Data set]. Borealis Dataverse. <https://doi.org/10.5683/SP3/WEOGMV>
- Read, K., Larson, C., Gillespie, C., Oh, S., & Surkis, A. (2019). A two-tiered curriculum to improve data management practices for researchers. *PLoS ONE*, 14(5), e0215509. <https://doi.org/10.1371/journal.pone.0215509>
- Si, L., Zhuang, X., Xing, W., & Guo, W. (2013). The cultivation of scientific data specialists: Development of LIS education oriented to e-science service requirements. *Library Hi Tech*, 31(4), 700-724. <https://doi.org/10.1108/LHT-06-2013-0070>
- Steeleworthy, M. (2014). Research data management and the Canadian academic library: an organizational consideration of data management and data stewardship. *Partnership: The Canadian Journal of Library and Information Practice and Research*, 9(1). <https://journal.lib.uoguelph.ca/index.php/perj/article/view/2990>
- 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>
- Tenopir, C., Hughes, D., Allard, S., Frame, M., Birch, B., Baird, L., Sandusky, R., Langseth, M., & Lundeen, A. (2015). Research data services in academic libraries: Data intensive roles for the future? *Journal of EScience Librarianship*, 4(2), e1085. <https://doi.org/10.7191/jeslib.2015.1085>
- Tenopir, C., Kaufman, J., Sandusky, R., & Pollock, D. (2019). *Research data services in academic libraries: Where are we today?* [White paper]. ACRL/Choice 360. <https://www.choice360.org/research/research-data-services-in-academic-libraries-where-are-we-today/>

- Thielen, J., & Neeser, A. (2020). Making job postings more equitable: Evidence based recommendations from an analysis of data professionals job postings between 2013-2018. *Evidence Based Library and Information Practice*, 15(3), 103-156. <https://doi.org/10.18438/ebliip29674>
- Thomas, C. V., & Urban, R. J. (2017). What do data librarians think of the MLIS? Professionals' perceptions of knowledge transfer, trends and challenges. *College and Research Libraries*. 79(3), 401-423. <https://doi.org/10.5860/crl.79.3.401>
- Wang, Y.-Y., & Lin, C.-S. (2019). A survey of data science programs and courses in the iSchools. *Proceedings of the Association for Information Science and Technology*, 56(1), 801-802. <https://doi.org/10.1002/pr2.184>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J., Bonino da Silva Santos, L.O., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>
- Xia, J., & Wang, M. (2017). Competencies and responsibilities of social science data librarians: An analysis of job descriptions. *College & Research Libraries*, 75(3), 362-388 <https://doi.org/10.5860/crl13-435>

Appendix

Examples of Aggregated Title Categories

1. Examples of the Advisors or Directors category include high-level administrators, research officers, unit heads, associate university librarians, directors, or other managers.
2. Examples of the Data or GIS librarian category include titles containing the terms data and data services, GIS or geospatial, data visualization, data curation, or data steward.
3. Examples of the Other Functional Librarian category include systems, metadata, scholarly communications, digital scholarship, digital curation, or digital preservation.
4. Examples of the Liaison Librarian category include STEM librarian, academic librarian, business librarian, or reference librarian.