Proceedings of the IATUL Conferences

2016 IATUL Proceedings

From Coast to Coast: Canadian Collaboration in a Changing RDM Seascape

Angela Henshilwood
University of Toronto (Toronto, Ontario)

Christina Hwang
University of Alberta (Edmonton, Alberta)

From Coast to Coast:
Canadian Collaboration in a Changing RDM Seascape

Cristina Sewerin, University of Toronto, Canada, Eugene Barsky, University of British Columbia, Canada, Dylanne Dearborn, University of Toronto, Canada, Angela Henshilwood, University of Toronto, Canada, Christina Hwang, University of Alberta, Canada, Sandra Keys, University of Waterloo, Canada, Marjorie Mitchell, University of British Columbia, Canada, Michelle Spence, University of Toronto, Canada, Kathy Szigeti, University of Waterloo, Canada, Tatiana Zaraiskaya, Queen's University, Canada

cris.sewerin@utoronto.ca, eugene.barsky@ubc.ca, dylanne.dearborn@utoronto.ca, angela.henshilwood@utoronto.ca, christina.hwang@ualberta.ca, skeys@uwaterloo.ca, marjorie.mitchell@ubc.ca, michelle.spence@utoronto.ca, kathy.szigeti@uwaterloo.ca, tatiana.zaraiskaya@queensu.ca

Abstract

There is power in collaboration. With academic institutions across Canada preparing for possible changes in funding requirements around sharing and preserving data and the submission of data management plans, librarians from a range of small, medium and large Canadian schools are collaborating to survey their researchers to better understand local needs and to generate together a richer understanding of their respective science and engineering research communities. Some of the questions we hope to answer with this endeavour include: What are the characteristics of the research data produced, and how do researchers in different disciplines manage their data? What attitudes towards research data management (RDM) support services and data sharing can be observed? In this paper, results from schools which have already run the survey will be summarized, including an analysis of comparable data. Similarities and differences observed across disciplines and schools will be highlighted. Participating institutions at time of writing include: University of Toronto, University of British Columbia, University of Waterloo, University of Alberta, Queen's University, University of Ontario Institute of Technology, Dalhousie University, University of Ottawa, and McGill University.

Keywords: research data, faculty practices, faculty attitudes, libraries, Canada, science, engineering

1.0 Introduction

Canadian academic institutions have been expecting changes in the Canadian public funding landscape for research, with possible new requirements around sharing and preserving data and the submission of data management plans. Developments such as the Government of Canada's release of the draft framework for comment on the collective realignment of funding agency policies [Social Sciences and Humanities Research Council, 2013], and the Draft Tri-Agency Statement of Principles on Digital Data Management [Government of Canada, 2015], have prompted librarians at Canadian academic institutions to prepare for expected changes in requirements for managing research data.

The federal government and its funding agencies, the Canadian Institutes of Health Research (CIHR), Natural Sciences and Engineering Research Council of Canada (NSERC), and Social Sciences and Humanities Research Council of Canada (SSHRC), have been consulting with various research communities, including libraries and archives, about the benefits and challenges of research data management (RDM) for some time. As these government entities continue their efforts, and with potential changes to funding requirements looming that could impact researchers, it is important for Canadian institutions to understand researcher practices and attitudes. This can help inform conversations, proactive service and infrastructure development, both at a local and national level.

With institutions across Canada facing similar challenges in understanding RDM practices, workflow and researcher attitudes towards sharing those data, librarians from a number of Canadian universities partnered to seek answers to the following questions:

- What are some of the characteristics of data produced by researchers at these institutions?
- How do researchers in different disciplines manage their data? Are there differences that can be observed between disciplines?
- What attitudes can be observed toward RDM support and services?

Working together, five Canadian universities, to date, have used a common survey instrument to gather information about their respective science and engineering researcher communities and generate a richer understanding of their users' RDM practices and attitudes. The participating universities are: University of Toronto (U of T), University of British Columbia (UBC), University of Waterloo (Waterloo), University of Alberta (U of A), and Queen's University (Queen's). Four other institutions, University of Ontario Institute of Technology, Dalhousie University, University of Ottawa, and McGill University, plan to survey their researcher populations later in the year.

	University of Toronto	University of British Columbia	University of Alberta	University of Waterloo	Queen's University
Students	84,556	61,113	38,733	35,900	22,114
Faculty	13,239	5,334	2,046	1,461	1,075

Table 1. Total number of all students and faculty at each institution [Queen's University, 2016a; Queen's University, 2016b; University of Alberta, 2016a; University of Alberta, 2016b; University of British Columbia, 2016; University of Toronto, 2016a; University of Toronto, 2016b; University of Waterloo, 2016.]

In this paper, the combined results of the five surveys currently completed are shared and insights to the questions posed above are explored.

2.0 Methods

The survey instrument was developed at the U of T Libraries and adopted by other libraries to allow direct comparative analysis. The survey instrument consisted of four sections: 1) working with research data; 2) data sharing; 3) funding mandates and RDM services; and 4) demographics and general questions.

Survey participants included all ranks of faculty members, postdoctoral and research fellows, and lecturers from engineering and science departments. Queen's also included graduate students at master's and doctoral levels.

The survey was distributed electronically using the subscription-based Survey Wizard (U of T) or FluidSurveys (Queen's, Waterloo, UBC and U of A) software. Each institution collected data for two or three weeks between April 2015 and December 2015.

Links to the following survey instruments are available: U of T [Sewerin, Dearborn, Henshilwood, Spence, Zahradnik, 2015]; UBC [Barsky, Mitchell, Buhler, 2016]; Waterloo [Szigeti, 2015]; Queen's [Zaraiskaya, Cooper, Moon, Murphy, Saleh, 2016]; U of A [Hwang, 2016].

Assistance with statistical analysis of anonymized survey results was provided by Alexandra Cooper from Queen's University Library and Dr. Monique Herbert from U of T's Ontario Institute for Studies in Education.

3.0 Results and discussion

The survey received a total of 780 responses from researchers at five institutions. Demographics and results are discussed below. Note that not every respondent answered every question and therefore the respondent counts for each question are different.

3.1 Demographics

In this paper, 780 responses from the five universities are included: Queen's (400); U of A (128); U of T (95); UBC (94); and Waterloo (63). The approximate total populations surveyed at each institution were: Queen's (1393; 594 faculty, 799 graduate students); U of A (825); U of T (1116); UBC (950); and Waterloo (786). All institutions included only completed survey responses with the exception of Queen's, which included both the complete and incomplete responses in their data.

A total of 676 respondents self-identified their rank. For all institutions, the highest number of respondents self-identified as professor (44.1%), followed by postdoctoral fellow (26.2%%), graduate student (18.5%), lecturer (3.4%), professor emeritus (2.2%), and adjunct professor (1.5%), with the remainder indicating other (4.1%). See Table 2 for details.

	Professor		Adjunct professor		Professor emeritus		Graduate student		Postdoctoral fellow		Lecturer		Other		Total
	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	
Queen's University	77	22.5%	6	1.8%	0	0.0%	125	36.5%	125	36.5%	9	2.6%	0	0.0%	342
University of Alberta	85	84.2%	4	4.0%	2	2.0%	0	0.0%	0	0.0%	3	3.0%	7	6.9%	101
University of Toronto	56	58.9%	0	0.0%	11	11.6%	0	0.0%	19	20.0%	7	7.4%	2	2.1%	95
University of British Columbia	54	57.4%	0	0.0%	1	1.1%	0	0.0%	21	22.3%	0	0.0%	18	19.1%	94
University of Waterloo	26	59.1%	0	0.0%	1	2.3%	0	0.0%	12	27.3%	4	9.1%	1	2.3%	44
Total by Rank	298		10		15		125		177		23		28		676
	44.1%		1.5%		2.2%		18.5%		26.2%		3.4%		4.1%		

Table 2. Number of respondents by rank for each institution.

Table 3 below shows the breakdown by discipline of the 546 respondents who identified both rank and home faculty, institute, or department. It should be noted that not every institution surveyed a department, school, or research centre in every discipline in the sciences and engineering. The areas surveyed for engineering include: aerospace, biomaterials and biomedical engineering, chemical, civil and mineral, engineering science, environmental and energy systems, electrical and computer, mechanical and industrial, and materials science. The areas surveyed for science include: agriculture; astronomy and astrophysics; biology; chemistry; computer science; earth and environment science; life and environmental sciences; marine science; mathematics and statistics; oceanography; optometry; pharmacy; physics; and psychology and neuroscience. These areas have been collapsed to disciplinary categories to preserve anonymity and to account for organizational differences between institutions. The disciplines, by rank, collated from the five institutions, and number of responses from each, are:

Engineering: civil/mineral/mining/environmental (73); biological/chemical/materials/mechanical (82); electrical/computer engineering (12).

Science: chemistry (48); computer science (49); earth science (66); mathematics (24); physics/astronomy (71); biology (67); other (54).

	Professor		Adjunct professor		Professor emeritus		Graduate Student		Postdoctoral Fellow		Lecturer		Other		Total
	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	
Civil/mineral/mining/ environmental engineering	23	31.5%	1	1.4%	3	4.1%	29	39.7%	8	11.0%	3	4.1%	6	8.2%	73
Biological/chemical/materials/ mechanical engineering	41	50.0%	1	1.2%	5	6.1%	21	25.6%	10	12.2%	2	2.4%	2	2.4%	82
Electrical/computer engineering	6	50.0%	0	0.0%	1	8.3%	0	0.0%	5	41.7%	0	0.0%	0	0.0%	12
Physics/astronomy	40	56.3%	1	1.4%	7	9.9%	9	12.7%	11	15.5%	1	1.4%	2	2.8%	71
Biology	45	67.2%	1	1.5%	1	1.5%	7	10.4%	9	13.4%	1	1.5%	3	4.5%	67
Earth science	33	50.0%	2	3.0%	4	6.1%	14	21.2%	4	6.1%	1	1.5%	8	12.1%	66
Computer science	14	28.6%	2	4.1%	0	0.0%	28	57.1%	1	2.0%	1	2.0%	3	6.1%	49
Chemistry	17	35.4%	0	0.0%	4	8.3%	14	29.2%	9	18.8%	1	2.1%	3	6.3%	48
Mathematics	13	54.2%	0	0.0%	2	8.3%	3	12.5%	1	4.2%	2	8.3%	3	12.5%	24
Other	42	77.8%	2	3.7%	0	0.0%	0	0.0%	2	3.7%	2	3.7%	6	11.1%	54
Total	274		10		27		125		60		14		36		546
	50.2%		1.8%		4.9%		22.9%		11.0%		2.6%		6.6%		

Table 3. Number of respondents for discipline as per rank. Note this table only includes responses from those that identified both rank and discipline.

As shown in Table 3, the majority of respondents (n=546) from most disciplines self-identified as professors (biology, 67.2%; physics/astronomy, 56.3%; mathematics, 54.2%; biological/chemical/materials/mechanical engineering, 50.0%; electrical/computer engineering, 50.0%; earth science, 50.0%; chemistry, 35.4%; and other, 77.8%), with the exception of computer science and civil/mineral/mining/environmental engineering, both of which had the highest number of respondents self-identify as graduate students (57.1% and 39.7%, respectively).

3.2 Working with Research Data

To gain a better understanding of current practices and to try to anticipate future needs, survey respondents were asked how they work with, manage, and store their data. When asked "How many research projects did you lead in the past year, for example, as a principal investigator or project lead?" out of 643 responses, 39.3% indicated they led 1-2 projects, 28.5% indicated they led 3-5 projects, and 15.9% indicated they led more than 5 projects. When asked "How much data storage do you estimate you use in an average research project?" 46.0% of the 643 respondents indicated they needed <50 GB per project. However, 3.9% of respondents indicated that they used between 4 and 500TB, and 1.9% used over 500TB per project, demonstrating that if institutions were interested in providing data repositories, they must be able to accommodate projects that produce a large amount of data. Even more interesting, there were respondents (0.3%) that typically required more than 500TB of storage that also led more than 5 research projects. While that is a small percentage of all respondents, these high-use projects must be taken into account when designing institutional data repositories.
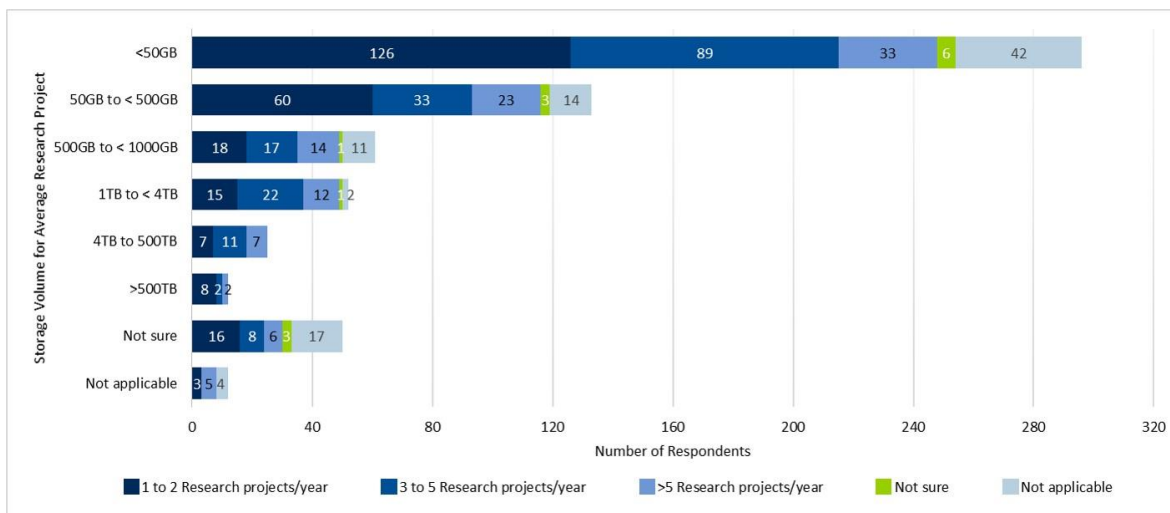
Figure 1. Responses to the question "How many research projects did you lead in the past year, for example, as a Principal Investigator or project lead?" in relation to responses to the question "How much data storage do you estimate you use in an average research project?" * Note: relationship statistically significant, $c^2(28) = 71.08$, $p < .01$

With the potential need for institutional research data repositories on the horizon, knowing the kinds of files that may need to be stored is important. To that end, researchers were asked "which of the following best describes the type of research data you generate or use in a typical research project" and were given a number of choices. Out of 639 respondents, 63.8% indicated that they used or generated numerical data (e.g. CSV, MAT, XLS, SPSS), while 60.9% indicated they used or generated text (e.g. TXT, DOC, PDF, RTF, HTML, XML). Other responses included: multimedia (e.g. JPEG, TIFF, MPEG, Quicktime, Bitmap) (39.7%), models (e.g. 3D, statistical, similitude, macroeconomic, causal) (36.8%), software (e.g. Java, C, Perl, Python, Ruby, PHP) (33.3%), instrument specific (e.g. Olympus Confocal Microscope Data Format, FLIR Infrared Camera (SEQ)) (31.1%), geospatial (e.g. raster, vector, grid) (15.5%), and other (e.g. discipline specific such as CIF, FITS, DICOM) (12.2%). Respondents could choose more than one response.

When asked where they store their research data from their current projects, 60.9% of the 752 respondents stated they used a computer hard drive (i.e. local hard drive). Encouragingly, 59.6% indicated that they used an external data repository (e.g. Protein Data Bank, Cambridge Structural Database, GitHub, Dryad, Figshare), indicating that at least some researchers are already using data repositories. Other responses included: external hard drive (52.5%), laptop hard drive (50.8%), flash drive/USB (36.7%), cloud/web based solution (e.g. Dropbox, Google Drive, Amazon Cloud, Microsoft Cloud) (33.1%), shared drive/university or departmental server (28.3%), hard drive of the instrument/sensor which generates the data (16.0%), physical copy retained (in boxes, cabinets, etc.) (12.2%), CD/DVD (9.3%), and grid/high performance computing (HPC) centre (6.6%). Respondents could indicate more than one storage option. A few respondents (0.8%) were not sure where they stored their data, and 2.5% used something other than the options listed to store their data. These results indicate that although some researchers are using data repositories to store data and practices may already be in place to preserve and share data, data security may be an issue for others, as noted by the percentage of respondents indicating they use flash drive/USBs and laptop computers to store data.

The next series of questions focused on data documentation. When asked if there was sufficient documentation and description (for example, file naming, cells & values, defined parameters, scripts to run) for another person outside their lab to understand and use the research data, only 38.5% of the 633 respondents said "yes", while 61.5% responded "no" or "not sure," indicating that this may be an area that researchers require guidance or assistance. Interestingly, when asked if there was sufficient documentation and description retained in the same file, folder or document for another

person outside their lab to replicate the methodologies that produced the data, 42.0% of respondents (n=633) indicated that there was sufficient documentation.

Respondents were asked about their practices regarding how long they kept their data. When asked how long they keep source material or raw data, 41.4% of respondents (n=628) indicated that they keep it until the data become inaccessible or lost, while 5.4% indicated they keep it only for the length of the project. When it comes to intermediate/working data, however, the percentage drops down to 35.2% for the respondents (n=623) that keep the data until it becomes inaccessible or lost, with 12.4% keeping the data for the length of the project. For processed data ready for publication, 45.3% of respondents (n=623) reported that they kept it until it became inaccessible or lost and 4.2% kept it for the length of the project. These results may indicate that researchers are not making active decisions about the retention or destruction of their data.

3.3 Sharing Research Data

Respondents were also asked to reflect on their practices and attitudes around sharing data. When asked about current sharing methods, 57.6% of the 597 respondents indicated that the most common method they used to share data was "by personal request only". Disciplines with the highest number of respondents that indicated they were not currently sharing data were computer science (44.9%, n=22) and the engineering areas (civil/mineral/environmental, 34.2%, n=25; biological/chemical/materials/mechanical, 24.1%, n=20; electrical/computer, 25.0%, n=3). Respondents from biology had the highest response rates in current sharing practices in both supplementary materials (52.9%, n=36) and repositories (45.6%, n=31). Respondents were also asked about their willingness to share in the future (n=244). Overall, respondents noted that they would be willing to consider sharing more openly, with 30.3% stating that they would consider using a general or discipline specific repository, and only 11.5% stating that they would not be planning to share their research data.

Interestingly, when asked with whom they would be willing to share data with, only 3.5% of the 594 respondents responded "nobody". Conversely, 34.5% of respondents responded that they would be willing to share with anybody (including the general public), though most indicated that they may be restrictive with who they were willing share, with the highest number showing openness to sharing with researchers in their field (49.2%).

When examining restrictions or embargoes that may limit the ability to share data, 32.8% of respondents (n=595) noted that their data would not be subject to any. Overall, the restriction options with the highest number of responses included: the need to publish data before sharing them (48.1%), that sharing data may jeopardize intellectual property rights (23.0%), a contractual obligation with a third party (19.7%), and data that are subject to privacy or confidentiality restrictions (17.0%). The discipline with respondents that appeared to be subject to the lowest number of restrictions was physics/astronomy (n=70), with responses below average to each type of restriction and 55.7% stating that no restrictions or embargoes prevented them from sharing. Electrical/computer engineering (n=12) responded above the average response to each type of restriction (with the exception of needing to publish data before sharing, and data being a matter of public safety or of a sensitive nature) and only two of these respondents stated that there were no restrictions or embargoes preventing them from sharing. Understanding differences in restrictions noted between the disciplines can help to develop training and infrastructure or mediate concerns to ensure that these restrictions or embargoes are respected.

Three hundred and ninety respondents expressed a level of interest in a service that provided assistance with issues associated with data preservation and/or sharing (n=550). To gain a better understanding of the issues these particular respondents may be interested in assistance with, their responses to the question about restrictions preventing them from sharing were examined (387 respondents answered both questions). The most noted restrictions preventing their data from being shared (outside of waiting to publish) were intellectual property rights (26.4%), contractual obligations

(22.2%), and privacy or confidentiality restrictions (20.2%). This list sheds light on the specific issues with which respondents interested in a service providing assistance with data preservation and/or sharing may like assistance.

Outside of restrictions or embargoes, respondents were also asked to identify the reasons they would not be willing to share their data. For the 586 respondents, the top reasons noted included: the data being incomplete or not finished (46.6%), still wishing to derive value from them (36.5%), insufficient time (27.8%), and lack of standards to make them usable by others (24.6%). Examining this question from a disciplinary perspective can offer some understanding of particular pain-points unique to the type of research.

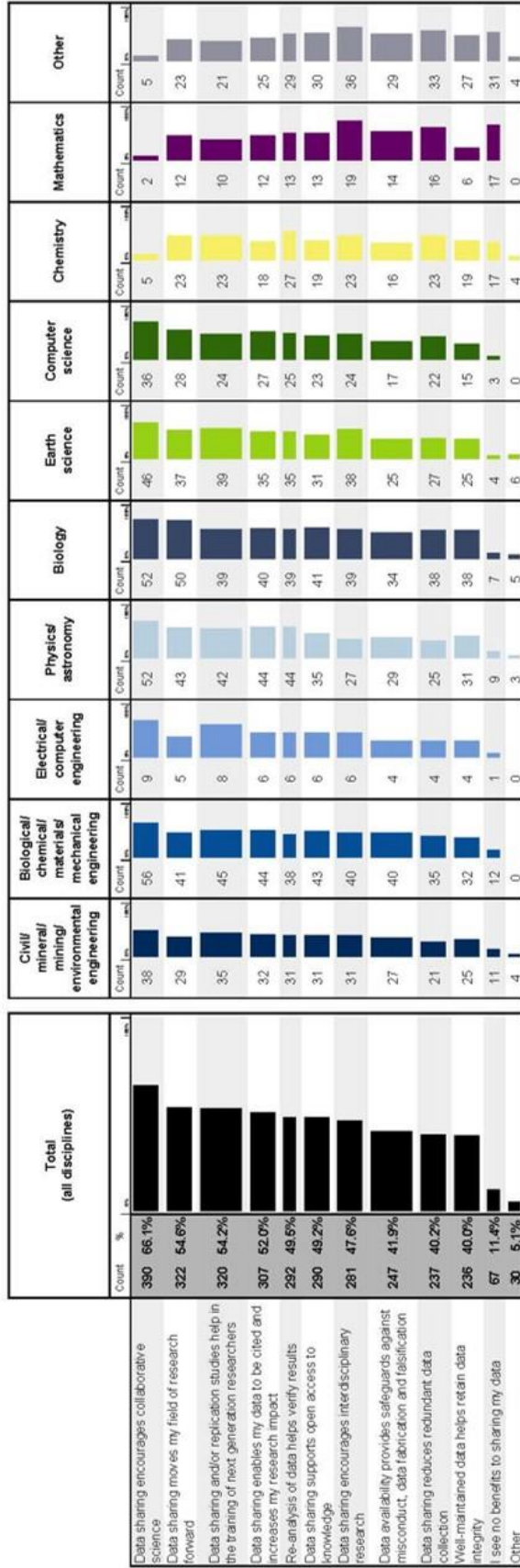| Benefit	Total (all disciplines)		Civil/ mineral/ mining/ environmental engineering	Biological/ chemical/ materials/ mechanical engineering	Electrical/ computer engineering	Physics/ astronomy	Biology	Earth science	Computer science	Chemistry	Mathematics	Other
	Count	%	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count
Data sharing encourages collaborative science	390	66.1%	38	56	9	52	52	46	36	5	2	5
Data sharing moves my field of research forward	322	54.6%	29	41	5	43	50	37	28	23	12	23
Data sharing and/or replication studies help in the training of next generation researchers	320	54.2%	35	45	8	42	39	39	24	23	10	21
Data sharing enables my data to be cited and increases my research impact	307	52.0%	32	44	6	44	40	35	27	18	12	25
Re-analysis of data helps verify results	292	49.5%	31	38	6	44	39	35	25	27	13	29
Data sharing supports open access to knowledge	290	49.2%	31	43	6	35	41	31	23	19	13	30
Data sharing encourages interdisciplinary research	281	47.6%	31	40	6	27	39	38	24	23	19	36
Data availability provides safeguards against misconduct, data fabrication and falsification	247	41.9%	27	40	4	29	34	25	17	16	14	29
Data sharing reduces redundant data collection	237	40.2%	21	35	4	25	38	27	22	23	16	33
Well-maintained data helps retain data integrity	236	40.0%	25	32	4	31	38	25	15	19	6	27
I see no benefits to sharing my data	67	11.4%	11	12	1	9	7	4	3	17	17	31
Other	30	5.1%	4	0	0	3	5	6	0	4	0	4

Figure 2. Responses to the question "What benefits do you see to sharing your research data?" in relation to total responses, and by discipline.

As shown in Figure 2, when asked about the benefits they saw to sharing their data, only 11.4% of the 590 respondents indicated they saw no benefit. Overall, respondents identified with the beliefs that data sharing encourages collaborative science (66.1%), data sharing moves their field of research forward (54.6%), data sharing and/or replication studies help in the training of next generation researchers (54.2%), and that data sharing enables data to be cited and increases their research impact (52.0%). Looking further into perceived benefits of sharing can offer insights for discussion within specific disciplines. For example, physics/astronomy respondents (n=71) noted they perceived that data sharing encourages collaborative science (73.2%), that re-analysis of data helps verify results (62.0%), and that data sharing enables data to be cited and increases research impact (62.0%). In contrast, chemistry respondents (n=48) noted that data sharing and/or replication studies help in the training of next generation researchers (56.3%), that well-maintained data helps retain data integrity (47.9%), and that data sharing moves their field of research forward (47.9%). In general, civil/mineral/mining/environmental engineering respondents answered below average to each perceived benefit to sharing, while biology respondents answered above average to each, showing opportunity for discussion with each, but perhaps targeted with different approaches.

Communication and information about funding and journal requirements regarding research data may be one service that could facilitate more open sharing. Respondents that indicated a level of interest in this type of service also expressed an increased interest in sharing their data in the future in external repositories or with publishers as supplementary files. This increase is demonstrated by examining how those interested in this type of service also answered the two questions about current and future sharing practices. Of the 438 respondents interested in this type of service, 437 answered the question about current sharing practices and 184 answered the question about future sharing practices (note that UBC and Queen's did not include the question about future sharing). Current practices indicated that only 25.4% (n=437) currently submit supplementary data files to a journal publisher, with 52.7% (n=184) that would consider it in the future, and that 15.6% (n=437) currently deposit in a repository, rising to 34.8% (n=184) for those considering this option for the future.

3.4 Funding mandates and RDM services

In this section, respondents were asked to identify the funding sources they used in the last five years or planned to use in the next five years. Of 379 respondents, 86.3% selected an NSERC grant as a source of funding, 33.2% selected a grant from the Canadian Foundation for Innovation (CFI), and industry was selected as a funding source by 25.3%. Of the 358 respondents who identified at least one of the Tri-Agencies (CIHR, SSHRC, NSERC) as a funding source, 82.9% said they would need or want assistance with drafting a data management plan as part of a grant application.

Respondents were also asked to consider possible services related to research data and asked to indicate their level of interest in each. The services that had the highest response rates included: 72.1% showed a level of interest in personalized consultation on data management practices for specific research groups or projects (n=549), 79.5% showed a level of interest in communication and information about funding requirements and journal requirements regarding research data (n=551), and 73.5% showed a level of interest in data storage and backup during active research projects (n=550).
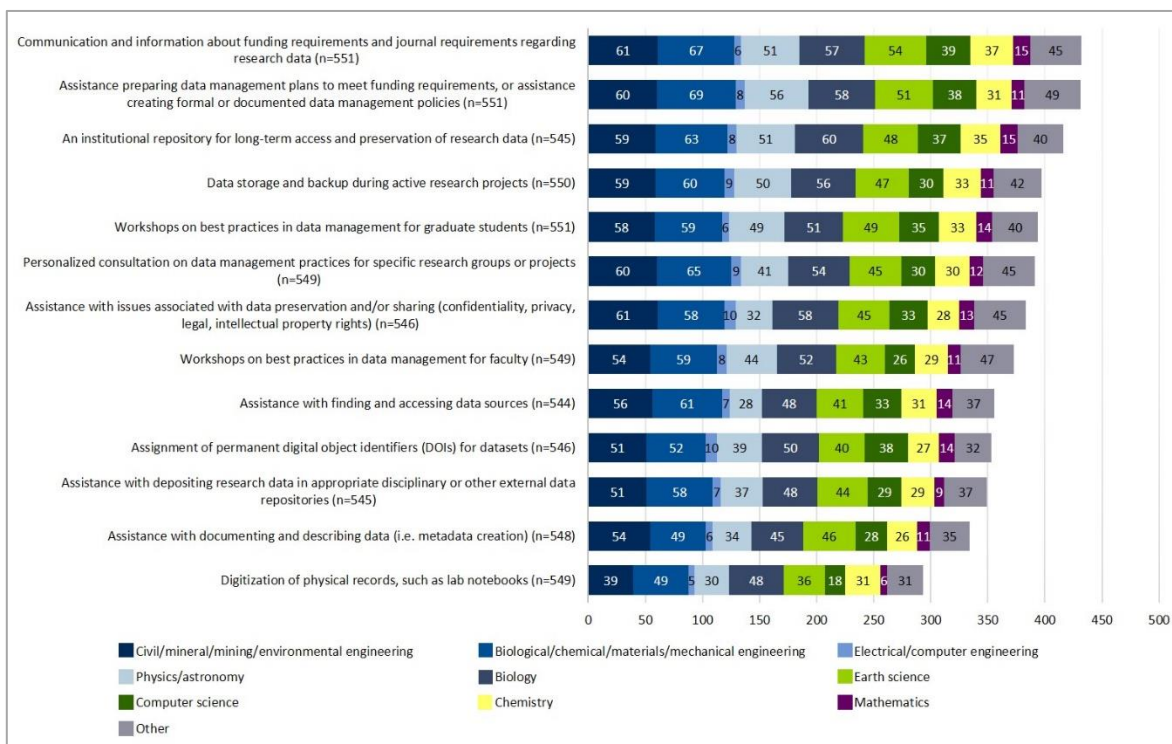
Figure 3. A level of interest shown in services, as broken down by discipline. Note that 'not applicable' and 'not interested' answers are not included.

At least half of all respondents showed some interest in every service proposed. Looking at the level of interest in services by discipline can help liaison librarians in conversations with their respective faculty or it may help prioritize library service development. For example, 72 respondents from the civil/mineral/mining/environmental engineering disciplines showed a level of interest in services such as workshops on best practices in data management for graduate students, which is slightly higher than other disciplines surveyed. Also, 71 respondents showed a level of interest in a service providing assistance with issues associated with data preservation and/or sharing (confidentiality, privacy, legal, intellectual property rights), which is also slightly higher than other disciplines surveyed. Liaison librarians could frame their discussions with researchers in this area around the development of these services.

4.0 Conclusions

The survey received 780 responses from five institutions, with an approximate total response rate of 15.4%. Some key findings for these responses include:

- The most common types of research data generated are numerical and textual, though multimedia, models, software, and instrument specific data were also identified by numerous researchers, pointing to a number of potential challenges in services such as storage and researcher support.
- The majority of respondents are currently depositing research data in external data repositories.
- Researchers may require guidance or assistance in documenting and describing their data, as evidenced by the clear majority of respondents who did not believe, or were unsure if there is sufficient documentation and description for another person outside their lab to understand and use their research data.
- Data security may be an issue for some researchers, as noted by the percentage of researchers indicating they use flash or USB drives and laptop computers to store data.

- Results may indicate that researchers are not making active decisions about the retention or destruction of their data, as evidenced by the percentage of respondents who indicated they keep raw, intermediate and processed data until they are lost or inaccessible.
- The most common method of sharing that respondents currently use is to share by personal request only. In terms of disciplinary differences, the highest number of responses from those willing to share in both supplementary material and repositories were from biology, while the highest number of respondents indicating that they were not currently sharing data were in the computer science and engineering areas.
- Most respondents saw some benefit to sharing data, though disciplines identified differently with perceived benefits, which may reveal incentives, as well as opportunities for discussion.
- In terms of restrictions or embargoes preventing the sharing of data, understanding differences noted between the disciplines can help to mediate concerns or develop training and infrastructure to ensure that these restrictions or embargoes are respected.
- A majority of respondents showed a level of interest in all research data services queried, with responses highest for communication about funding and journal requirements, assistance preparing data management plans, and an institutional repository for data.

The explosion and complexity of research data have brought researchers, and the institutions they work for, new challenges in the management, curation, preservation, and long-term storage of those data. The aggregate data gathered by the participating institutions of this study provides valuable insight into the current RDM environment on those campuses and can inform planning and implementation of services and infrastructure to support researchers. Results of the surveys may also help inform decisions and improve understanding of disciplinary trends among Canadian researchers. The remaining four institutions (University of Ontario Institute of Technology, Dalhousie University, University of Ottawa, and McGill University) expect to run their survey iterations for science and engineering researchers later this year. The nine institutions are also currently adapting the survey instrument to use on their respective social science and humanities researcher populations. The additional data gathered will add considerably to the understanding of the Canadian RDM landscape.

5.0 References

Barsky, E., Mitchell, M., Buhler, J. (2016). *UBC Research Data Management Survey.* Retrieved from http://hdl.handle.net/11272/10364. V1 [Version].

Government of Canada (2015). *Draft tri-agency statement of principles on digital data management.* Retrieved from http://www.science.gc.ca/default.asp?lang=En&n=547652FB-1

Hwang, C. (2016). *Research Data Management Faculty Survey, University of Alberta*. Retrieved from https://drive.google.com/file/d/0B0Y3YAsi7aKpdkxzMkVfVDdadzg/view.

Queen's University (2016a). Quick facts: Enrolment (fall 2014) [table]. Retrieved from http://www.queensu.ca/about/quickfacts.

Queen's University (2016b). Quick facts: Faculty [table]. Retrieved from http://www.queensu.ca/about/quickfacts.

Sewerin, C., Dearborn, D., Henshilwood, A., Spence, M., & Zahradnik, T. (2015). *Research Data Management Faculty & Postdoctoral Survey, University of Toronto.* Retrieved from http://hdl.handle.net/1807/69144.

Social Sciences and Humanities Research Council (2013). *Capitalizing on big data: Toward a policy framework for advancing digital scholarship in Canada*. Retrieved from http://www.sshrc-crsh.gc.ca/about-au_sujet/publications/digital_scholarship_consultation_e.pdf.

Szigeti, K. (2016). *University of Waterloo Survey Instrument Dec. 2015.* Retrieved from http://hdl.handle.net/10864/11649. V2 [Version].

University of Alberta (2016a). Summary statistics [table]: Faculty. In *2014-2015 Annual report*. Retrieved from https://uofa.ualberta.ca/-/media/ualberta/reporting/annual-reports-and-financial-statements/annual-report-20142015.pdf.

University of Alberta (2016b). Students by the numbers 2014-15. Retrieved from https://uofa.ualberta.ca/about/facts/students.

University of British Columbia (2016). UBC overview & facts: UBC at a glance 2015/16. Retrieved from https://www.ubc.ca/about/facts.html.

University of Toronto (2016a). Quick facts: Faculty & staff (fall 2013). Retrieved from https://www.utoronto.ca/about-u-of-t/quick-facts.

University of Toronto (2016b). Quick facts: Student enrolment (fall 2014-15). Retrieved from https://www.utoronto.ca/about-u-of-t/quick-facts.

University of Waterloo (2016). Waterloo facts: Our people. Retrieved from https://uwaterloo.ca/about/who-we-are/waterloo-facts.

Zaraiskaya, T., Cooper, A., Moon, J., Murphy, S., Saleh, N. (2016). *Queen's University RDM Survey.* Retrieved from http://hdl.handle.net/10864/11651. Data Services [Distributor] V1 [Version].