Institutional Research Data Management Services Capacity Survey INSIGHTS Report #1

RDM supports within Organizations: Budget, Structure, and Strategies

Prepared by the Portage Network, Research Intelligence Expert Group (RIEG) on behalf of the Canadian Association of Research Libraries (CARL)

JUNE 2020

www.carl-abrc.ca

Table of Contents

Introduction

Background

Over the past few decades, Research Data Management (RDM) has become a major policy initiative of research funding agencies as a mechanism to enhance research excellence. The Canadian *DRAFT Tri-Agency Research Data Management Policy for Consultation*[1], released in May 2018, promotes both RDM and data stewardship practices as integral to achieving research excellence. The draft policy is directed at researchers as well as the institutions that administer Tri-Agency funds, and it establishes the need for individual institutions to develop and implement strategies to support researchers in achieving the goals of the new policy (Government of Canada, 2018). The draft policy follows the 2016 *Tri-Agency Statement of Principles of Digital Data Management*, which articulates expectations for stakeholders at all levels of the research enterprise (Government of Canada, 2016). Since the release of the draft policy, research institutions across Canada have begun examining their ability to support RDM in the context of establishing institutional RDM strategies. In response, the Portage Research Intelligence Expert Group (RIEG) has initiated a series of surveys to examine institutional readiness for the rollout of the new Tri-Agency RDM policy in 2020 and the capacity to support researchers more broadly. The first survey, *Institutional RDM Strategy Survey,* examined the state of development of institutional RDM strategies. The summary report, released in November 2019, indicated that a majority of research institutions had begun the process of developing a strategy, although many were only in the initial stage of reviewing support materials (Portage Research Intelligence Expert Group, 2019).

As a continuation of information seeking by RIEG, the *Institutional Research Data Management Services Capacity Survey* was conducted in order to benchmark the current capacity of Canadian research institutions to provide necessary RDM support before the final Tri-Agency RDM policy is implemented. Key findings from the summary report, released in February 2020, illustrate that RDM capacity is largely being built from the ground up, primarily through the efforts of libraries in collaboration with other internal partners. Human resource capacity is the greatest need identified, which institutions are addressing through hiring practices and professional development for technical skill requirements (Cooper et al., 2020).

The purpose of this report is to examine in greater detail the findings reported in the *Institutional Research Data Management Services Capacity Survey* as they relate to

---

[1] Tri-Agency includes the three major federal funding agencies in Canada: Social Sciences and Humanities Research Council (SSHRC), Canadian Institutes of Health Research (CIHR), and Natural Sciences and Engineering Research Council (NSERC) (Government of Canada, 2018).

organizations and how they support RDM through budgets, collaborations, organizational structures and strategies.

The Survey

The bilingual questionnaire surveyed the current state of Canadian research institutions regarding development and allocation of human, organizational, infrastructure, and fiscal resources for RDM on their campuses. It also solicited suggestions for additional support that the Portage Network and other stakeholders could provide to assist with these efforts.

The survey was administered from September 3, 2019 – October 18, 2019 using SimpleSurvey software. It was distributed through listservs and a contact list of identified stakeholders. Each institution was asked to gather information from across campus and provide a coordinated response, although multiple responses from a single institution were also accepted. The survey consisted of 27 questions[2] ranging from general demographic information to detailed questions about current infrastructure and services in place across institutional stakeholder groups.

Summary of Findings

In this Insights Report, we focus on the organizational contexts through which Canadian institutions are building their RDM capacities. Topics include:

- Policies
- Budgeting
- Staffing Positions
- Internal collaborations
- Work structures
- External collaborations

Across institutions, the most engaged partners for RDM initiatives are libraries and IT (on the service provider side) and senior leadership and research administrative units (on the administrative side.) As leadership partners, faculty and graduate students are less often represented. Future investigations should include exploring the perspectives of these stakeholders and whether there are different approaches to priorities, policy development, budget, and capacity considerations.

---

[2] The survey questions can be found in the Data Dictionary in the Appendix of the *Institutional Research Data Management Services Capacity Survey: Executive Summary* (Cooper et al., 2020).

Most institutions have some kind of dedicated personnel involved in working on RDM, whether through the creation of new positions or reassignment of responsibility. Universities are typically creating new RDM positions, most of which are created in libraries. The relatively small group of colleges reporting RDM positions indicated that these are personnel reassigned to RDM responsibilities. Further investigation is needed to determine why this is the case.

The majority of institutions reported some type of internal structure to support RDM initiatives, whether formal or informal; libraries and research administrative bodies are more likely to be the leaders in these structures. As RDM policy in Canada evolves, it is anticipated that the internal structures will become more formal.

Canadian universities are actively participating in regional and national RDM collaborations at a high rate; however, they are relatively less active in disciplinary or international RDM networks. This activity may be happening at a researcher level, which might not be detected by our survey. Colleges are reporting very few external collaborations; these are mostly at the regional level. It seems that the existing regional and external collaborative natures of Canadian institutions, especially among university libraries, have provided channels for the RDM networks. There is an opportunity to increase and extend representation at the disciplinary and international levels. The New Digital Research Infrastructure Organization (NDRIO)[3] could play a role here in the future. Institutions with limited resources will especially benefit from increased inclusion in external resources available through collaboration.

Financial commitments to RDM support are not yet widely seen at the institutional level. Some universities are able to make some funding streams available for RDM through particular units, but very few colleges have done this. It would be useful to revisit this in the future to see if there are changes. If RDM financial needs cannot be met with institutions' internal resources, external funding or collaborative investment may become necessary.

Completed RDM strategy and policy development levels are still low, as the final Canadian Tri-Agency Research Data Management Policy is still anticipated, but many institutions are currently engaging in the process. Based on our results, the strategies that are forming have been concentrating on data storage, personal and sensitive data, and FAIR data sharing, with less emphasis on the legal and licensing aspects, or in the disciplinary areas. Canadian institutions currently lack consistent guidance or policy on data ownership, data retention periods, and data licensing, which are essential for researchers to share research data. This could align with our results showing limited participation of institutional legal offices in current strategy

[3] See further information about NDRIO on Engage DRI website: https://engagedri.ca/.

development. Institutions would need to provide support and interpretation in these areas as well. National guidance will be key in ensuring appropriate strategy and policy development in these areas.

Detailed Results

For each of the topics covered in this report, overall distributions and comparisons between types of research institutions[4] are reported.

Demographics

The survey received 85 responses from 77 institutions across Canada (Table 1). Eight institutions submitted two responses, which were combined into one response per institution, and the responses from the 77 institutions will be used in this report[5]. Based on the institutions' names, we were able identify their types (universities, colleges/CEGEPs[6], research centres, and governments) and the region they are part of:

- West (British Columbia, Alberta, Saskatchewan, Manitoba)
- Ontario
- Quebec
- Atlantic (New Brunswick, Nova Scotia, Prince Edward Island, Newfoundland & Labrador)[7]

---

[4] "Institutions" refers to the 77 institutional respondents – universities, colleges, research centres, and government organizations for the remainder of the report.

[5] The principles for the consolidation are 1). If original values are consistent from the two respondents, the value for the combined case is the same as the originals; 2). Text answers from the original two responses are combined into the value for the combined case; 3). If one of the original values is either "not chosen" or "don't know", the value for the combined case will be the same as the other original value, which is either "chosen" or other affirmative answers, for example, "yes" or "no"; 4). Q7: A higher level of participation from the double entries is assigned to the combined case; 5). Q11: priority number is first averaged from the duplicates, and then reordered accordingly. Arbitrary ordering is only used when two factors happen to have the same average priority score for a combined case, which will not have a significant influence on the aggregated result.

[6] CEGEPs, or Collèges d'enseignement général et professionnel, are publicly funded, post-secondary, pre-university, collegiate technical colleges exclusive to the Quebec provincial education system.

[7] Institutional responses from provinces other than Ontario and Quebec were combined geographically into "West" and "Atlantic" to anonymize the responses, which were relatively low in number.

Across all regions of Canada, 52 universities (52.5% of the 99 Universities Canada members[8]) and 21 colleges[9] (16.5% of the 127 Colleges and Institutes Canada members[10]) represented the majority of the respondents in our survey[11].

The survey also gathered information about respondent affiliations at their institution; respondents could choose all that applied. The following departments/offices are represented in the survey: Library (50); Research Office (39); CIO (8); Ethics Board (9); Researchers (11); IT (10); and Other (7).

	Atlantic	Quebec	Ontario	West	Total
Types of institution					
University	9	14	14	15	52
College/CEGEP	0	9	6	6	21
Research Centre	0	0	0	1	1
Government	0	0	1	2	3
Total	9	23	21	24	77

Table 1. *Number of institutional responses by region and institution types. Region and institution type are derived from Q1. "Name of institution."*

Results Across Institutional Types

Organizations
The development and implementation of institutional RDM strategies benefits from collaboration between multiple stakeholders within the institution (Portage Network, 2020). Our survey investigated the engagement levels of different stakeholders and their level of leadership.

Partners
We examined which internal organizational units are leading the development of institutional RDM capacity. Figure 1 summarizes the participation patterns of the top four most frequently involved units (libraries, research administrative units, IT departments, and senior leadership) by institution type.

---

[8] See the "Member Universities Archive" compiled by Universities Canada: https://www.univcan.ca/universities/member-universities/
[9] Colleges refers to both colleges and CEGEPs for the remainder of the report.
[10] Total number of colleges is gathered from the Colleges and Institutes Canada: https://www.collegesinstitutes.ca/our-members/member-directory/
[11] The survey asked respondents to estimate full-time equivalent researchers within given ranges; however, as a majority of respondents did not answer this question, we cannot present an accurate composite of researcher numbers across institutions.

Among the 52 universities, 65 reported either leading or participating involvement from all of the above four units, specifically 49 Libraries (94.2%); 45 Research Administrative units (86.5%); 37 Information Technology (IT) departments (71.2%); and 36 Senior Institutional Leadership (69.2%).

Among the 21 colleges, 45% reported either leading or participating involvement by all of these 4 units: 15 Research Administrative bodies (71.4%); 10 Research Ethics Boards (47.6%); 10 Senior Leadership (47.6%); and 10 Libraries (47.6%).
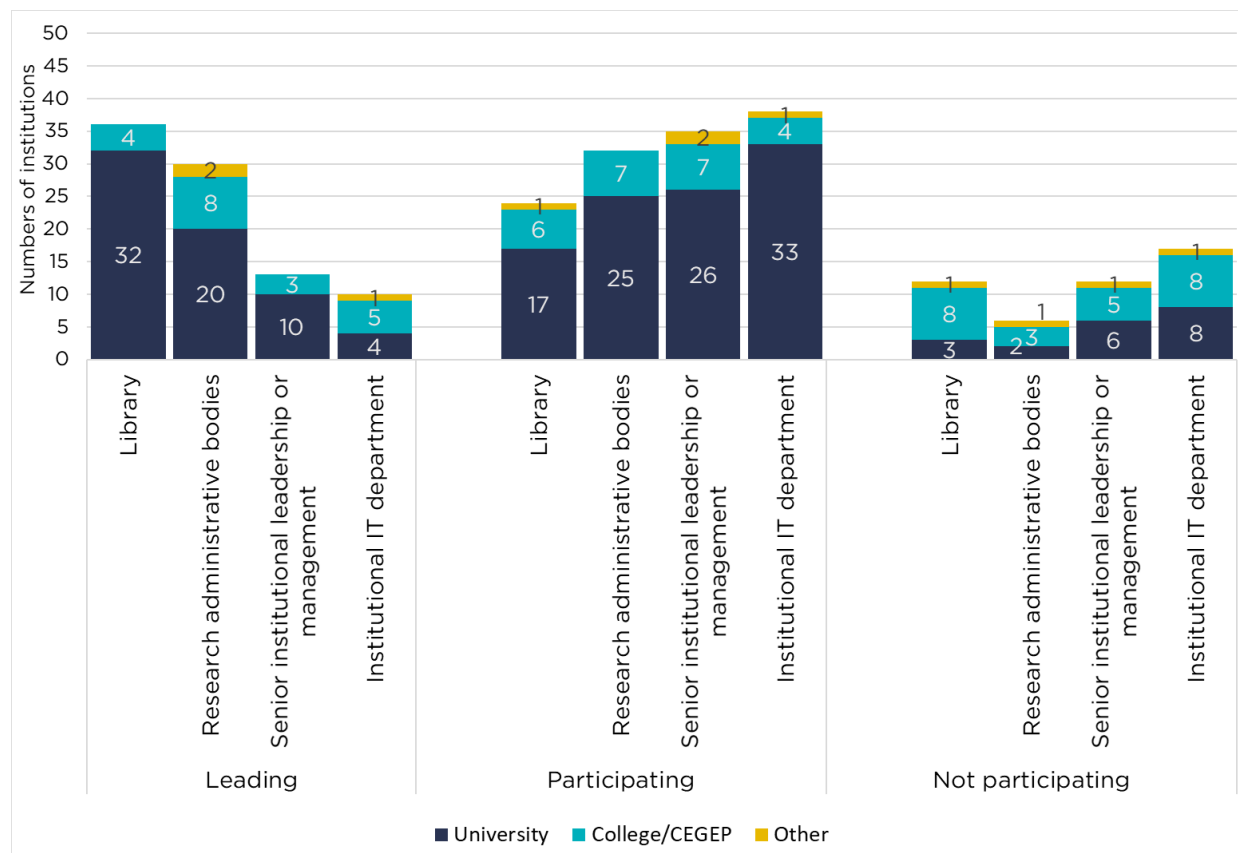
Figure 1. *Organizational units supporting RDM within an institution. Q7. "Which department(s) is/are involved in leading and participating in research data management at your institution?" (n=77)*

Within institutions, other units are comparatively less engaged: only 16.8% of Faculties of Graduate Studies, 19.5% of institutional legal departments and 24.7% institutional big data initiatives or data science centres are leading or participating in RDM activities at this point in time. Given that these stakeholders are potentially key to successful RDM initiatives, conscious engagement with these bodies should be sought.

Positions

We surveyed whether institutions have created or reassigned any RDM-related positions. The results (Figure 2) show that overall, 33.8% created new positions and 28.6% reassigned some positions to take up RDM responsibilities, as either fully or partially added to an existing job description. However, more than a third of institutions (37.7%), have neither full-time nor part-time dedicated RDM positions. Among the 52 universities reporting, 71% have created or reassigned RDM positions; most are new positions. For the 21 colleges responding, only 38.1% reassigned current positions with RDM responsibilities; the majority (61.9%) neither have RDM dedicated positions, nor created new positions.

The types of RDM-supporting positions vary across institutions, but most are located within library units. The job titles mentioned for new RDM positions include RDM coordinator, RDM librarian, and data manager/steward/specialist/curator. RDM responsibilities were also assigned to data/statistics librarians, preservation and scholarly communication librarians, research services librarians, and liaison/subject librarians. In other units, IT/security and research office positions were also mentioned occasionally.
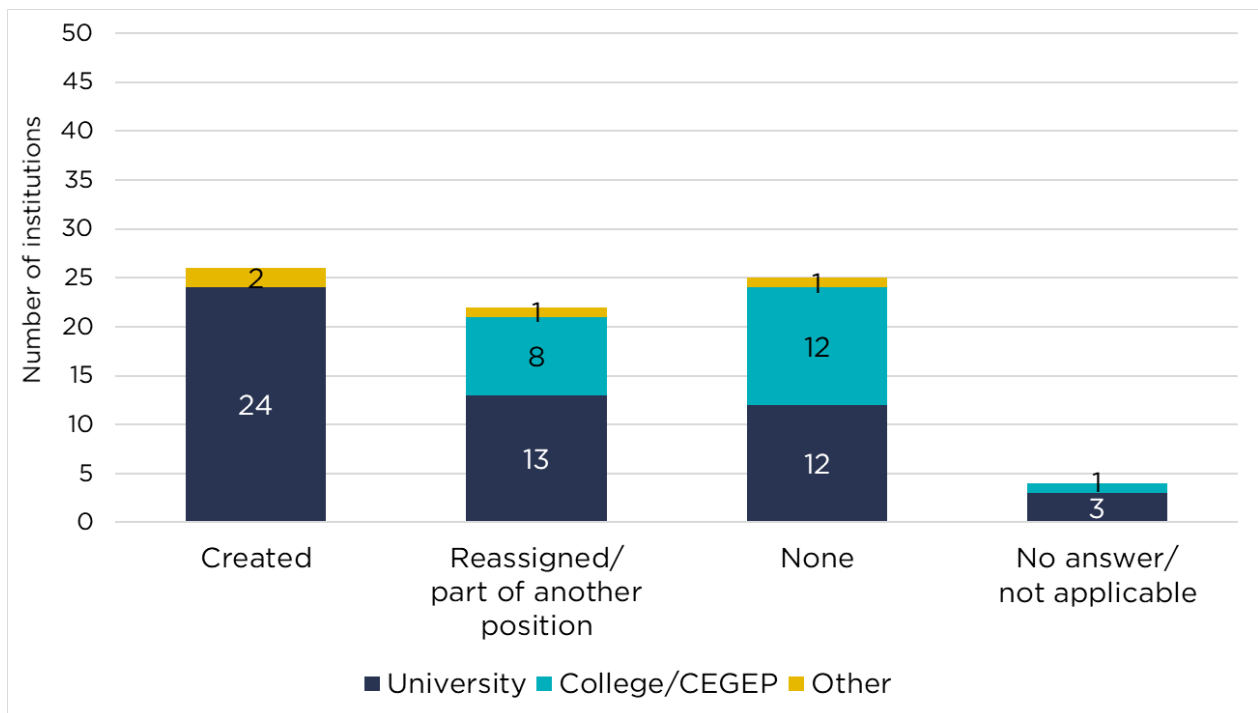

Figure 2. *Data management positions created or reassigned by institution type. Q8. "What data management positions (if any) have your institution created or reassigned (please provide associated job titles)?" (textual responses are coded into categories of created, reassigned/part of another position, none, and no answer/not applicable). (n=77)*

Working Structure

We explored how organizations are collaborating internally to develop capacity and identified a continuum of structures emerging across institutions (Figure 3). The responses were coded into three categories: institutional RDM groups, informal multi-unit groups/discussion within an institution, or no current stakeholder collaboration within an institution.

Almost half of reporting institutions (46.8%) have formed RDM working groups, and one-quarter (24.7%) work collaboratively via institution-level informal/ad-hoc groups, stakeholders discussions, or smaller-scale collaborations within single units of the institution. Less than one-third of institutions (28.6%) have not formed any institutional stakeholder working structure or did not report any in the survey.

Over half of 52 universities that we surveyed (55.8%) have already worked on RDM through formal institutional groups, and another quarter (25%) have various levels of informal collaboration or discussions established. Among the 21 colleges, slightly less than half (42.9%) have formed either formal or informal multi-stakeholder RDM working structures.
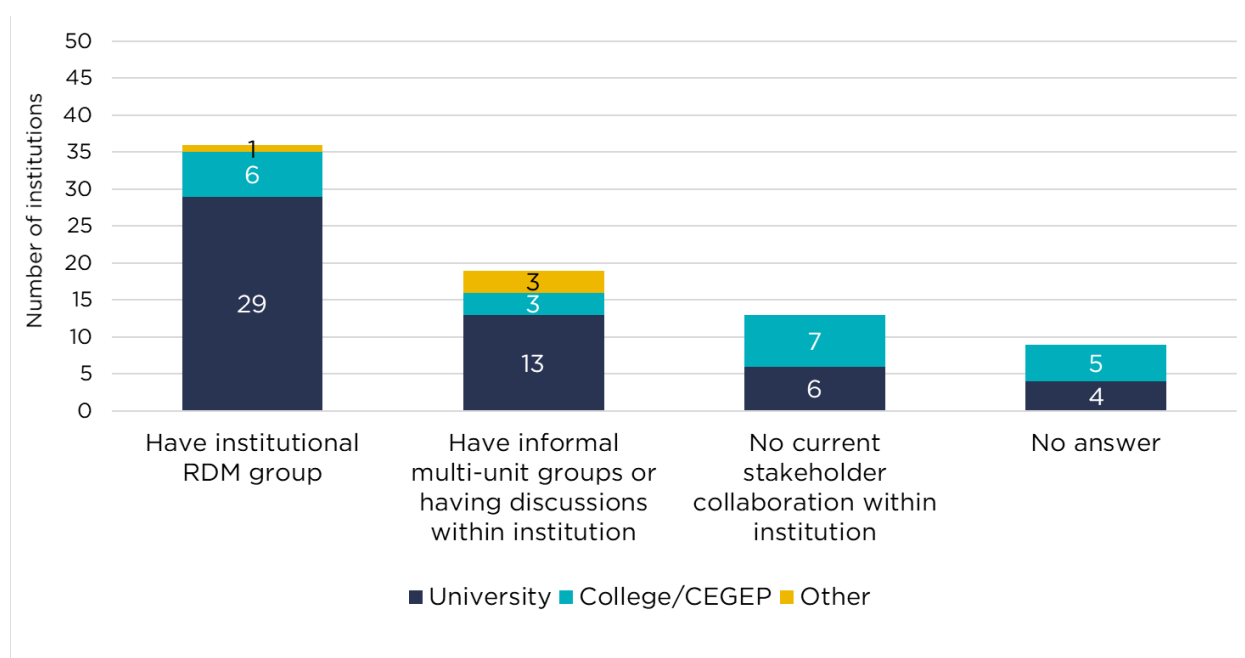

Figure 3. *Stakeholder collaboration mechanisms within institutions by institution type. Q9. "How are different stakeholders within your institution working collaboratively to tackle research data management challenges?" (textual responses are coded into categories of institutional RDM group, informal multi-unit groups/discussion within institution, no current stakeholder collaboration within institution, and no answer). (n=77)*

Figure 4 further examines the level of involvement of institutional partners in working structures across all institution types. Of the 53 institutions those reporting the establishment of formal groups, libraries were most often leading those groups (39.6%), followed by research administrative bodies (28.3%). On the other hand, from the 22 institutions reporting informal multi-unit groups, most often led by research administrative bodies (45.4%), followed by libraries (40.9%).
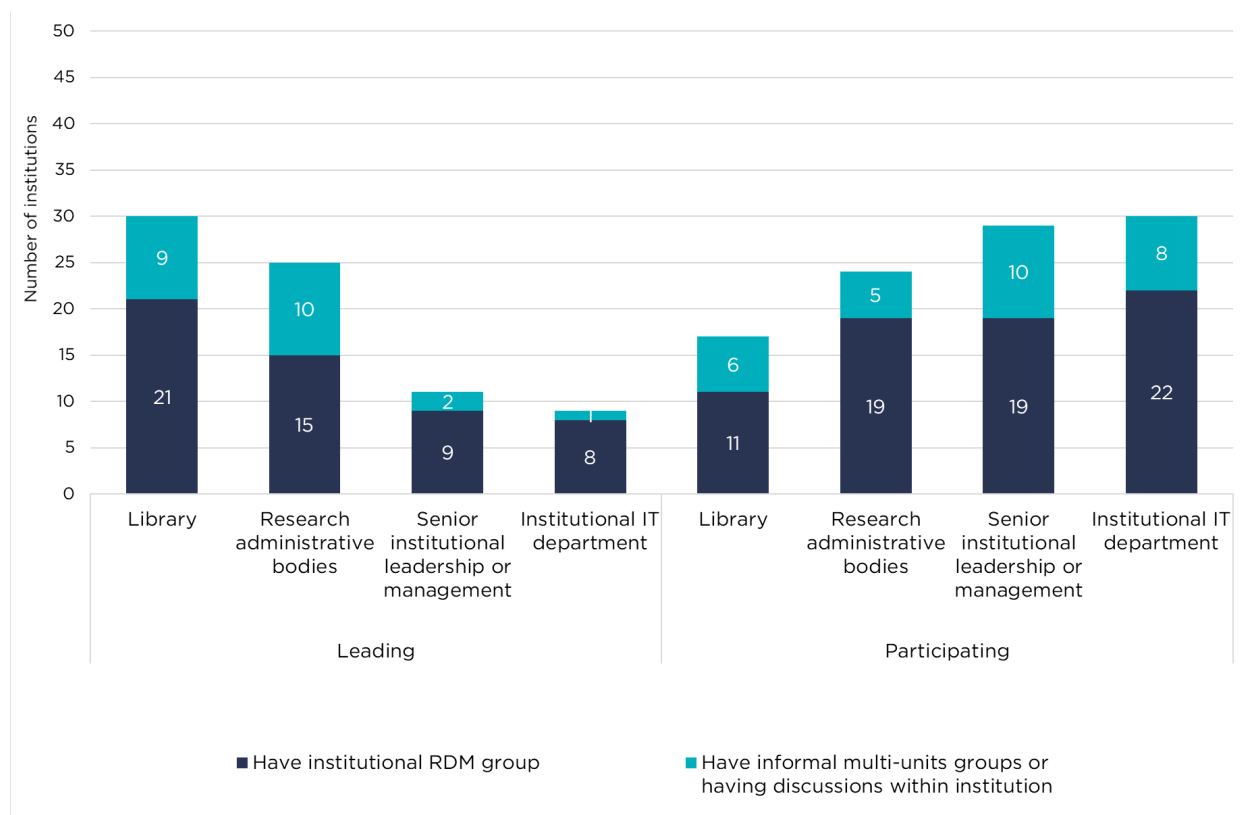

Figure 4. *Stakeholder participation in RDM working groups. Q7. "Which department(s) is/are involved in leading and participating in research data management at your institution?". Respondents could check all that applied. Not included in this figure are the respondents that selected "not participating" (34). Q9. "How are different stakeholders within your institution working collaboratively to tackle research data management challenges?". Textual responses are coded into categories: have institutional RDM group, have informal multi-unit groups or having discussion, no current stakeholder collaboration within institution, or no answer (9, not shown in figure). (n=77)*

External Collaboration

Building institutional RDM capacity may be assisted by collaborative relationships to support infrastructure and services, with both internal and external partners. External partnerships may be especially valuable where internal resources are more limited (Bryant et al., 2018). Institutions were asked to report the nature of their external collaborations: regional, national, international or discipline-specific. These types were

not strictly defined and therefore open to interpretation. Regional collaboration could include consortia or multi-institutional RDM initiatives within a province or smaller geographical area but could also include interprovincial initiatives or consortia (e.g. Council of Prairie and Pacific University Libraries). Nationally-led initiatives by such organizations like the Portage Network or the Canadian Association of Research Libraries (CARL) are examples of national collaboration. However, these initiatives can also be influenced by or include larger regional consortia. Similarly, there are discipline-specific initiatives that could also be interpreted as either national or international in scope.

More than half of surveyed institutions contribute to regional and national collaboration initiatives: 72.1% reported regional collaborations and 65.6% reported national collaborations. Only 16.4% of institutions collaborate internationally for RDM, and 14.8% participate in disciplinary collaborations. The low rate of reporting of international or disciplinary collaborations may be due to the nature of the survey respondents positions (i.e., librarians or research administrators rather than researchers) and their knowledge of collaborations at the broader institutional level. For the 52 universities, both regional and national initiative categories were equally reported and with high responses (69.2%). For the 21 colleges, regional collaboration (28.6%) was the most commonly identified type of external collaboration; however, the majority (71.4%) did not identify any type of external collaborations.

As shown in Figure 5, external collaboration patterns are similar for institutions across the regions, and all align with the general findings of more national and regional collaborations.
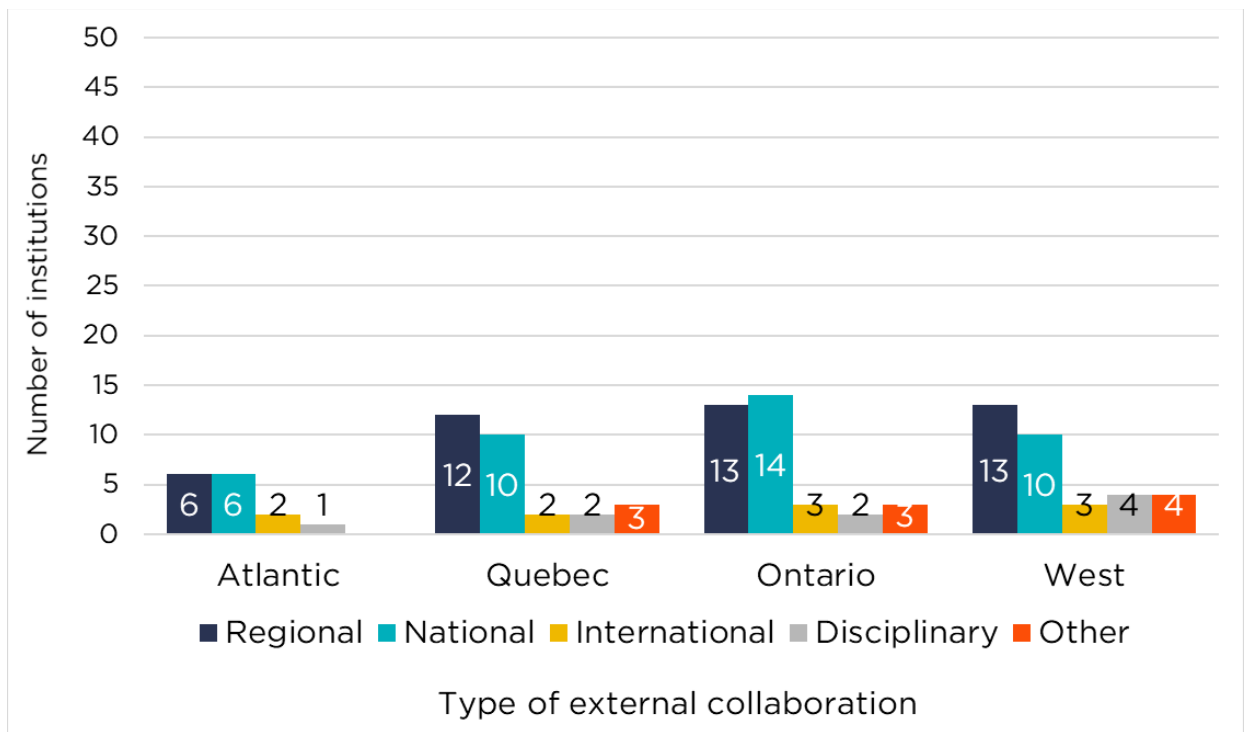
Figure 5. *External collaborations by region. Q24a-d. "Does your institution partner with or contribute to any of the following research data management initiatives?" (respondents could choose more than one option). (n=77)*

Institutions with formal RDM groups are more likely to have all forms of external collaborations, compared to institutions with informal RDM groups of stakeholders or no current stakeholder collaborations. Figure 6 compares external collaboration types across internal collaboration structures within university respondents. The 29 universities that indicated they had formal RDM groups identified both regional and national collaboration equally (82.8%) as the most common partnership.
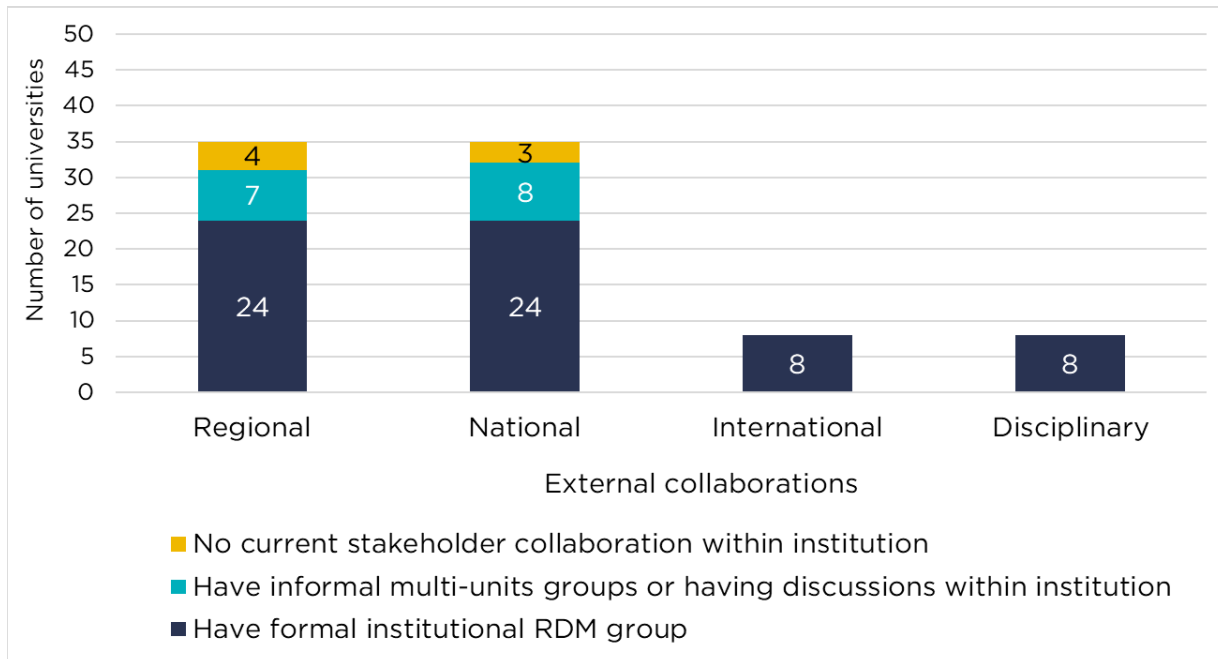
Figure 6. *University respondents' external collaborations and internal RDM collaboration mechanism. Q24a-d. "Does your institution partner with or contribute to any of the following research data management initiatives?" (respondents could choose more than one option). (n=52)*

Financial Investment in RDM

Budget allocation is one indicator of organizational commitment or recognition that RDM capacity development requires dedicated funding. Institutions were asked to report how they budget for RDM services, and their responses were coded into five categories: a dedicated institutional budget; a unit operational budget; a cross-institutional collaboration/consortium-type budget; budget funded by external source, such as a grant; or no dedicated budget.

As shown in Figure 7, just over one quarter of institutions (27.3%) either did not answer the question or replied unknown; 37.7% of the institutions indicated there is no dedicated RDM budget in their institutions. Another quarter of institutions (27.3%) reported using unit operational budgets for RDM services, which has thus become the most current common way of funding RDM among Canadian institutions. Only one institution reported a dedicated institutional RDM budget. Three institutions contribute financially to cross-institutional collaborations or consortia, and two institutions acquired RDM funding from outside of their institutions.
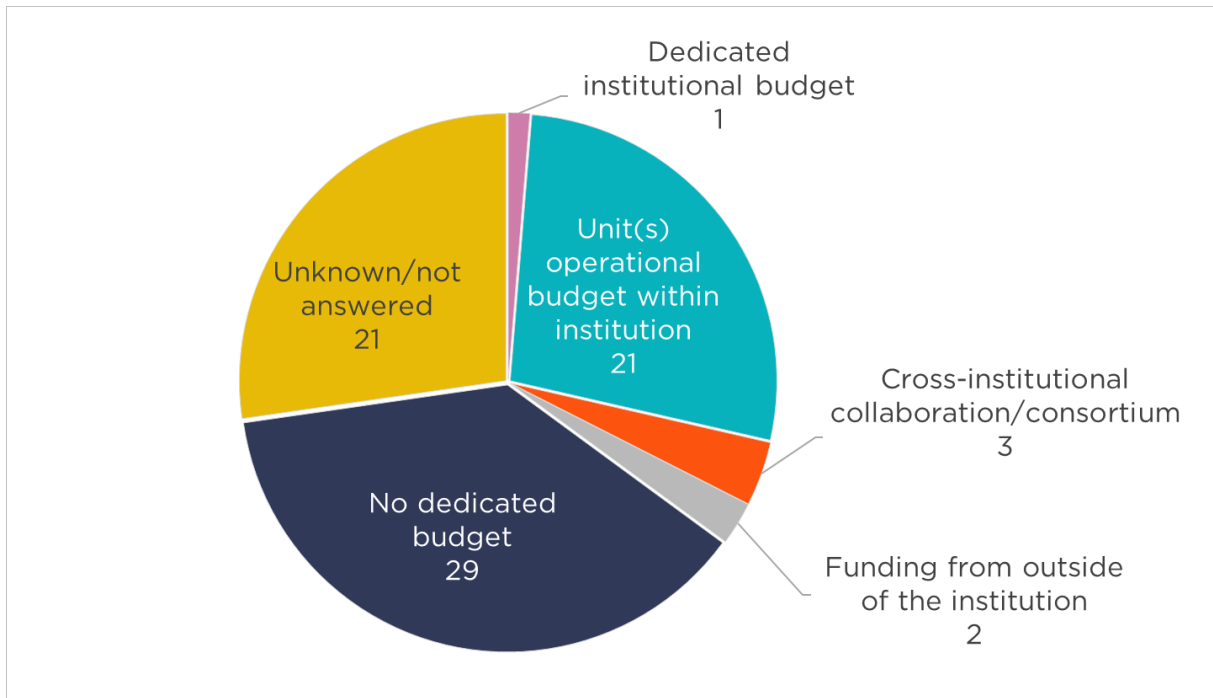
Figure 7. *Institutional budgeting mechanisms for RDM services. Q19 "How does your institution budget for research data management services?" (Write-in question, text answers were coded by the authors into categories of dedicated institutional budget, unit(s) operational budget within institution, cross-institutional collaboration/consortium, funding from outside of the institution, no dedicated budget, and unknown/not answered). (n=77)*

Figure 8 shows the breakdown of RDM budgeting models across institutional types. Approximately one third of the 52 universities (34.6%) allocated some funding for RDM through their unit budgets, compared with only 9.5% of the 21 colleges. It is clear that most institutions, especially colleges, do not have committed funding for RDM at the institutional level.
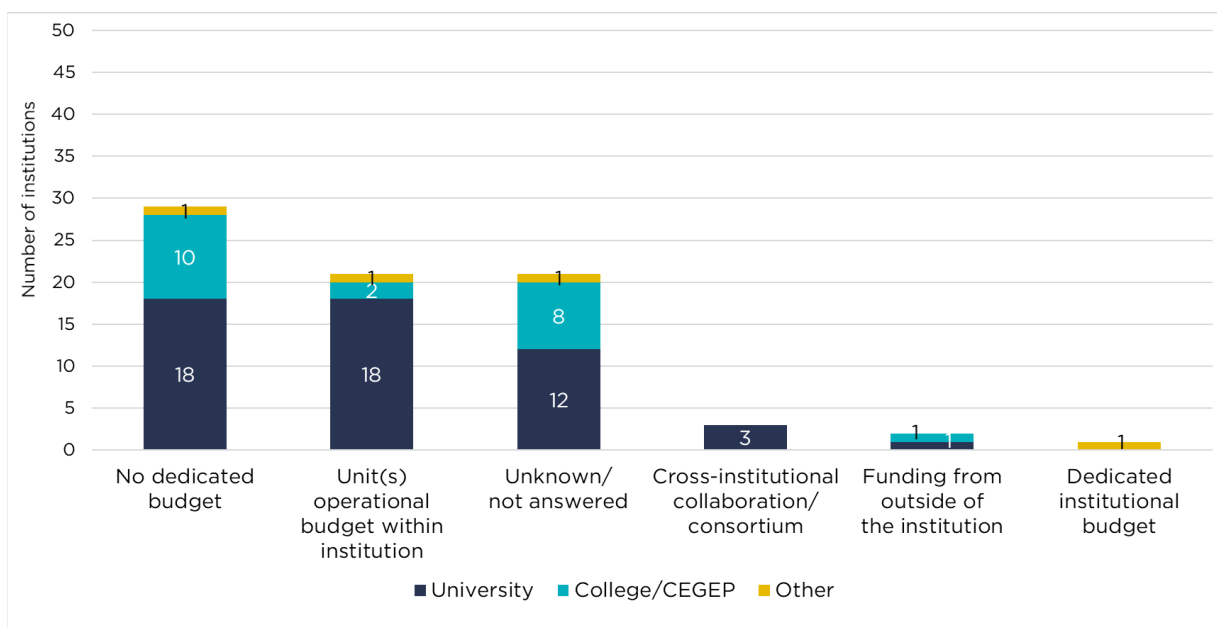
Figure 8. *Institutional budgeting mechanisms for RDM services by institutional type. Q19 "How does your institution budget for research data management services?" (write-in question, text answers were coded into categories) (n=77)*

Strategy/Policy Development

Developing strategies or policies have been defined as the initial step of institutional RDM engagement (Pinfield et al., 2014). The Canadian Tri-Agency draft Research Data Management Policy begins with a requirement for institutions to create their own institutional RDM strategy, which is defined as an outline of "how the institution will provide its researchers with an environment that enables and supports world-class research data management practices". The draft policy further suggests that such institutional strategies should include a commitment to "develop their own data management policies and standards for data management plans" (Government of Canada, 2018) . The *Institutional RDM Strategy Survey* previously conducted by RIEG revealed that many Canadian research institutions have already started to prepare for their RDM strategy development; some even have drafted documents (Portage Research Intelligence Expert Group, 2019).

To build on these results and to further investigate the state of both the institutional strategy and policy developments, we asked whether strategy and/or policy development was underway or completed and if so, which of the following elements were addressed: FAIR principles[12]; sensitive data guidance; provisions for data storage, specific discipline needs, and handling of personal identifying information;

---

[12] For a definition of FAIR principles (Findable, Accessible, Interoperable, Reusable data) see https://www.go-fair.org/fair-principles/

legal support; and licenses. The survey also examined related policies that may already exist, including ownership of research data, minimum retention periods for research data, and requirements for licensing research data.

Figure 9 shows that most responding institutions have not started developing RDM policies (71.4%). In contrast, 5.2% of responding institutions - two colleges and two universities - already have RDM strategies, and 45.5% are currently developing them, over half of which are universities (51.9%).
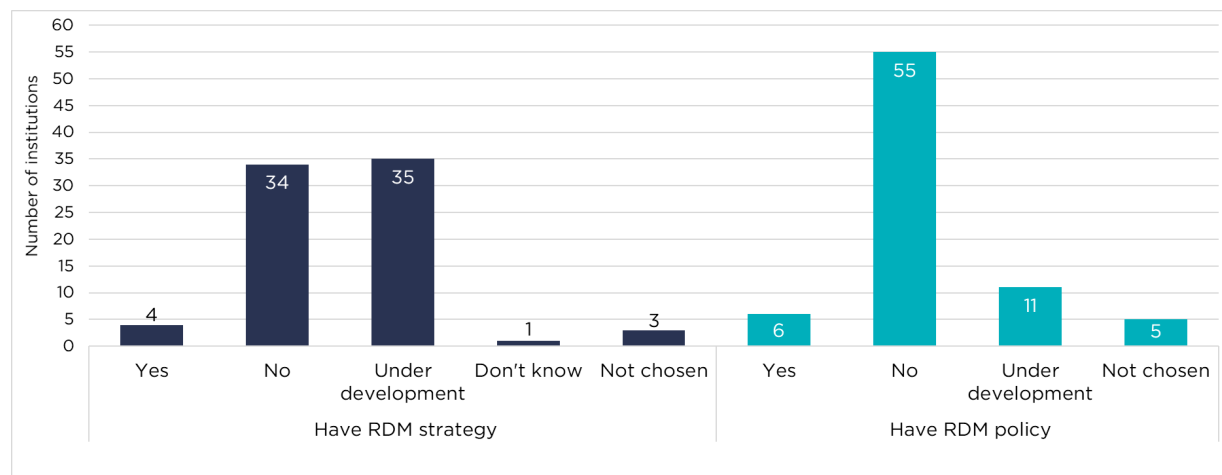

Figure 9. *Status of RDM strategy and policy development. Q14. "Does your institution currently have a research data management strategy?" and Q15. "Does your institution currently have a research data management policy?". (n=77)*

As shown in Figure 10, four common elements are covered by over half of the strategies and policies, whether existing or in development: provisions for data storage, FAIR principles, provisions for handling personal data, and guidelines for sensitive data. For both strategy and policy development, the majority of the institutions have not been including content on licensing guidelines, legal support, or provisions for specific disciplinary areas. Universities and colleges show very similar patterns, although the FAIR data principles seem to be less emphasized in colleges.
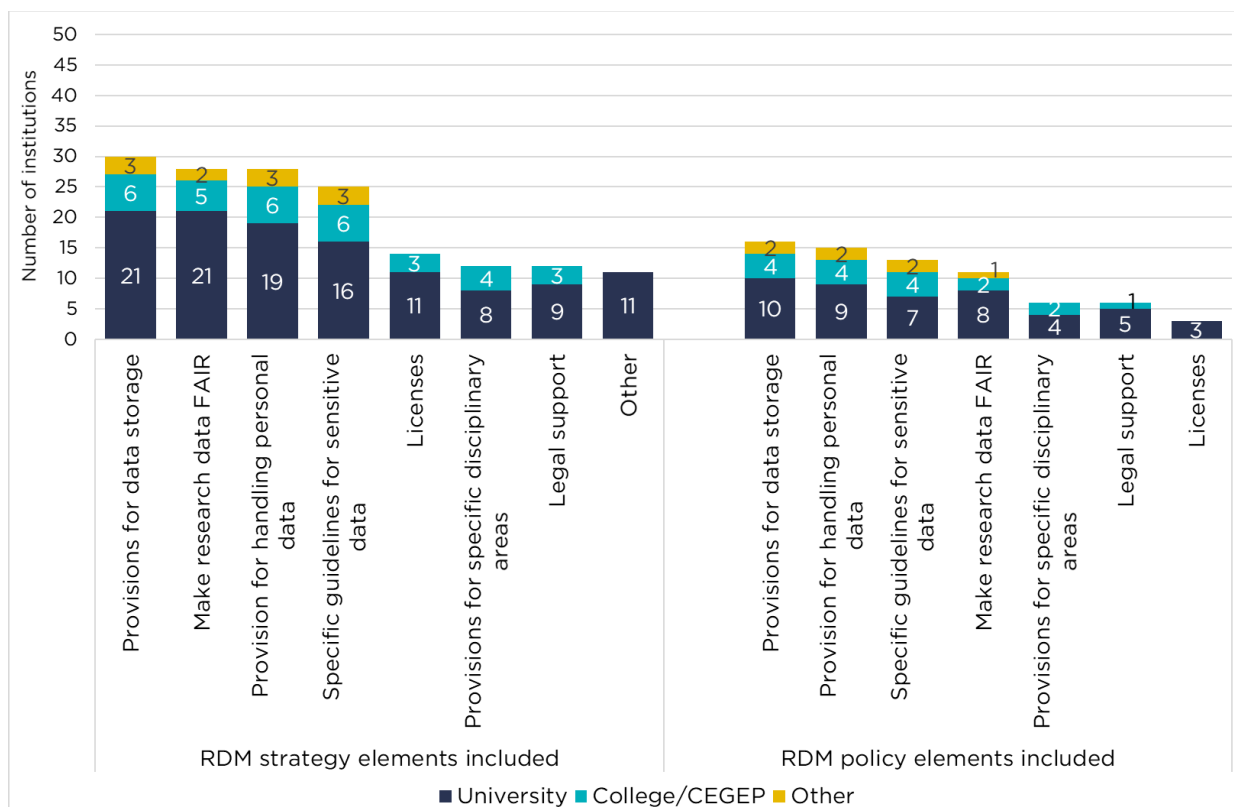
Figure 10. *Elements included in institutional RDM strategies and RDM policies. Q14b. "If you selected 'Yes' or 'Under development' in Question 14a, what elements are covered by your institutional strategy on research data management? Check all that apply." and Q15b. "If you selected the response 'Yes' or 'Under development' in Question 15a, what elements are covered by your institutional policy on research data management? Check all that apply." (n=77)*

Data ownership, minimum data retention periods, and licensing of research data are of particular interest, as these aspects are essential considerations for data sharing and reuse. Over one third of the institutions have guidelines or policies on data ownership (39%). Figure 11 shows that half of the 52 universities surveyed have regulations on data ownership and just under half reported that researchers own the data that they create. For colleges, the two institutions which reported having a data ownership policy stated the policy stipulated that the institution is the owner of research data. Eighteen institutions (23.3%), equally universities and colleges, chose "other" and provided more detail about their data ownership regulations. Some mentioned that their intellectual property policy on research products might be applied to research data. Others said that data ownership is stipulated in collective agreements with faculty associations as well as various other types of institutional governance documentation. Others pointed out that research data ownership could also be determined by grant or contract, and thus possibly belong to external research partners. Still others mentioned a shared or hybrid model of data ownership. The

means by which institutions can require researchers to abide with RDM strategies and policies may vary as a result.

Over one third of institutions have guidelines on retention periods (36.4%), while fewer institutions have guidelines on data licensing (13%). A slightly larger percentage of colleges (42.9%) than universities (32.7%) have data retention guidelines. Eight universities (15.4%) reported requirements for research data to be retained for at least five years, and another eight (15.4%) require a seven-year data retention period. Most colleges did not report a specific data retention period. Some institutions also noted that data retention periods could be project specific and dependent on corresponding disciplines, funders, and ethics review requirements. More colleges (23.8%) than universities (7.7%) reported guidelines for data licensing.
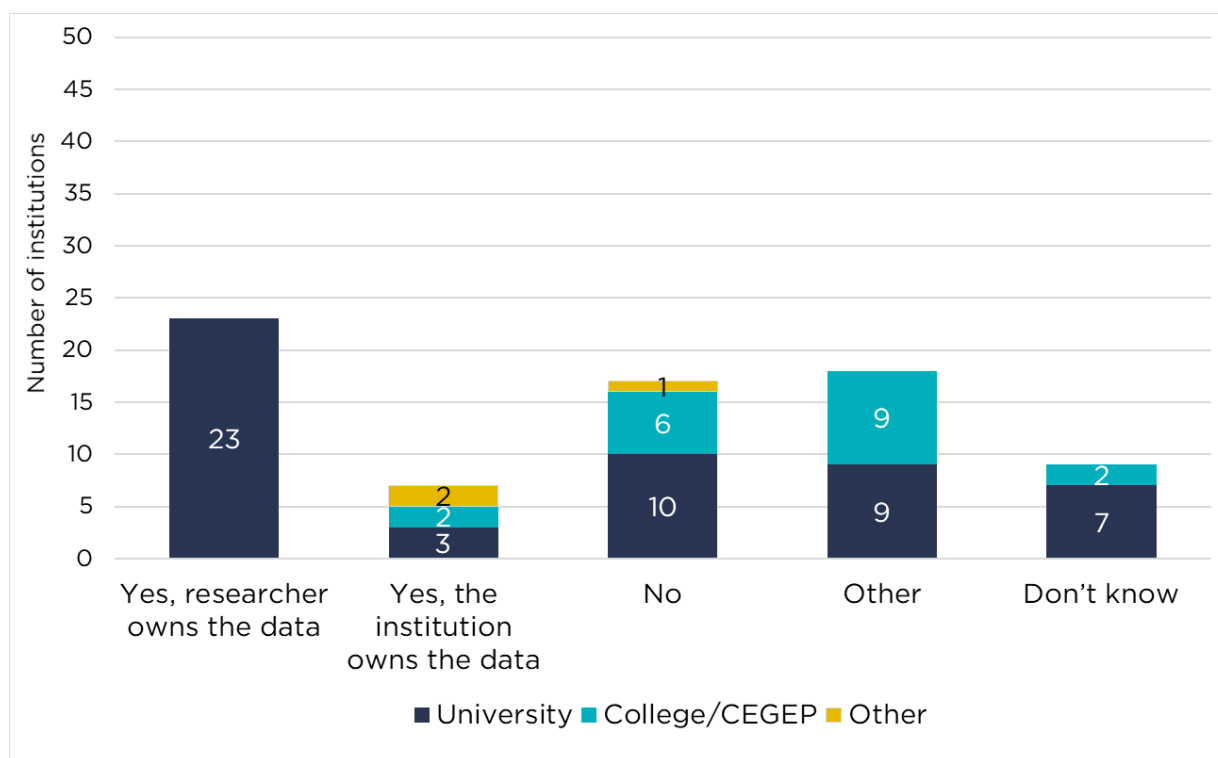

Figure 11. *Institutional guidelines or policies on ownership of data. Q16. "Is there a guideline/policy at your institution on the ownership of research data?" (n=74)*

References

Bryant, R., Lavoie, B., & Malpas, C. (2018). *Sourcing and scaling university RDM services* (No. 4; The Realities of Research Data Management). OCLC Research. https://doi.org/10.25333/C3QW7M

Cooper, A., Perry, C., Szwajcer, A., Wang, M., & Khair, S. (2020). *Institutional research data management services capacity survey: Executive summary*. Portage Network. https://doi.org/10.14288/1.0388722

Government of Canada. (2016, December 21). *Tri-Agency statement of principles on digital data management*. http://www.science.gc.ca/eic/site/063.nsf/eng/h_83F7624E.html

Government of Canada. (2018, May 25). *DRAFT Tri-Agency research data management policy for consultation*. https://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html

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

Portage Network. (2020). *Institutional research data management strategy guidance V. 2.0*. https://portagenetwork.ca/wp-content/uploads/2020/05/Portage-Institutional-Strategy-Guidance-v5-EN.pdf

Portage Research Intelligence Expert Group. (2019). *Institutional RDM strategy survey - summary of results*. Portage Network. https://portagenetwork.ca/wp-content/uploads/2019/11/RIEGStrategySurveySummary_EN.pdf