Institutional Research Data Management Services Capacity Survey
INSIGHTS Report #3
Future of Institutional RDM Support: Resources Prioritized, Investments, Challenges, and Accelerators

MAY 2021

Table of Contents

Introduction

The Research Intelligence Expert Group (RIEG) is a pan-Canadian working group of NDRIO-Portage that gathers evidence to guide the development of best practices in Research Data Management (RDM) in Canada, and communicates with stakeholder communities about existing and arising issues in related policies and practices.[1]

From September 3 to October 18, 2019, RIEG administered a bilingual questionnaire that surveyed the current state of Canadian research institutions in developing infrastructure and allocating human, organizational, and fiscal resources for RDM on their campuses. The survey also solicited suggestions for additional support that NDRIO-Portage,[2] and other stakeholders, could provide to assist these efforts. The goal of the survey was to assess institutional capacity to support RDM in response to the Tri-Agency Research Data Management Policy, which at that point in time was still in draft form.[3] The survey was distributed through several RDM-focused Canadian listservs and a contact list of identified institutional stakeholders. Each institution was asked to gather information from across campus and provide a single coordinated response, although multiple responses from a single institution were also accepted. The survey instrument consists of 27 questions ranging from general demographic information to detailed questions about current infrastructure and RDM services across institutional stakeholder groups.[4]

An Executive Summary of the survey results was released in January 2020,[5] followed by the first Insights report in June 2020,[6] and the second Insights report in March 2021.[7] This third and final Insights report explores the following topics in RDM capacity development:

---

[1] Research Intelligence Expert Group.
https://portagenetwork.ca/network-of-experts/research-intelligence-expert-group/
[2] New Digital Research Infrastructure Organization (NDRIO), https://engagedri.ca/; Portage Network, https://portagenetwork.ca/. Portage will become part of NDRIO as of March 31, 2021.
[3] The Tri-Agency Research Data Management Policy was released in 2021, replacing the 2018 draft Policy. http://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html.
[4] The survey questions can be found in the Data Dictionary in the Appendix of the Institutional research data management services capacity survey: Executive summary by Cooper, A., Perry, C., Szwajcer, A., Wang, M., & Khair, S. (2020), Portage Network, https://doi.org/10.14288/1.0388722.
[5] Institutional Research Data Management Services Capacity Survey: Executive Summary. http://hdl.handle.net/2429/73607
[6] Abel, Jennifer, Cooper, Alexandra, Dearborn, Dylanne, Perry, Carol, Szwajcer, Andrea, & Wang, Minglu. (2020, June 24). Institutional Research Data Management Services Capacity Survey INSIGHTS Report #1. RDM Support within Organizations: Budget, Structure, and Strategies. Zenodo. http://doi.org/10.5281/zenodo.3906443.
[7] Cooper, Alexandra, et al. Institutional Research Data Management Services Capacity Survey INSIGHTS Report #2 Current Capacity within Institutions: Highly Qualified Personnel, and Infrastructure and Services. Zenodo, 24 Feb. 2021. http://doi.org/10.5281/zenodo.4570030

- RDM Resources Prioritized
- Investment in RDM Technology and Human Resources
- Barriers and Challenges to Support RDM
- Accelerators for RDM Development

Respondent Demographics

The survey received 85 responses from 77 institutions across Canada (Table 1), including universities, colleges/CÉGEPs, [8] research centres, and governments organizations. Eight institutions submitted two separate responses, which were combined into one response per institution.[9] The responses from the 77 institutions are summarized in this report.

Based on the institutions' names, we classified the responses into the institutional types of universities, colleges/CÉGEPs, research centres, and government organizations by geographical region:

- West (British Columbia, Alberta, Saskatchewan, Manitoba)
- Ontario
- Quebec
- Atlantic (New Brunswick, Nova Scotia, Prince Edward Island, Newfoundland & Labrador)[10]

	Atlantic	Quebec	Ontario	West	Total
Types of institution					
University	9	14	14	15	52
College/CÉGEP	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."

[8] The principles for the consolidation are 1) If the original values are consistent from the two respondents, the value for the combined case is the same as the original; 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 was either "chosen" or other affirmative answers, for example, "yes" or "no".

[9] CÉGEPs, or Colleges d'enseignement general et professionnel, are publicly funded, post-secondary, pre-university, collegiate technical colleges exclusive to the Quebec provincial educational system.

[10] Institutional responses from provinces other than Ontario and Quebec were combined geographically into 'West' and 'Atlantic' to anonymize the relatively low number of responses.

RDM Resources Prioritized

Respondents were asked to rank, by priority level, the resources required to streamline RDM services at their institution. Resources included human, [11] financial, technical, together with policy and legal guidance, with an option to provide additional requirements in the 'other' category. Figure 1 summarizes the priority of resource requirements across respondent institutions where the percentage of positive-sentiment responses (Level 1-3) will appear as positive and the percentages of negative-sentiment responses (Level 4-6) as negative values.

Overall, human resources were identified as the highest priority (level 1) by 37.7% of institutions responding, with 75.3% of all respondents ranking it in their top three priorities. Policy guidance was indicated as a top three priority by 55.8% of institutions and it was the highest reported priority among 22.1% of institutions. The only significant difference in responses from universities and colleges was in the area of policy where universities placed policy as the highest priority (28.9%) compared to highest priority in colleges (9.5%).

The need for financial (66.2%) and technical (68.8%) resources scored higher than policy overall as the top three priorities. However, lower percentages of institutions rank them as their highest priority (Level 1) (financial (16.9%); technical (14.3%)).
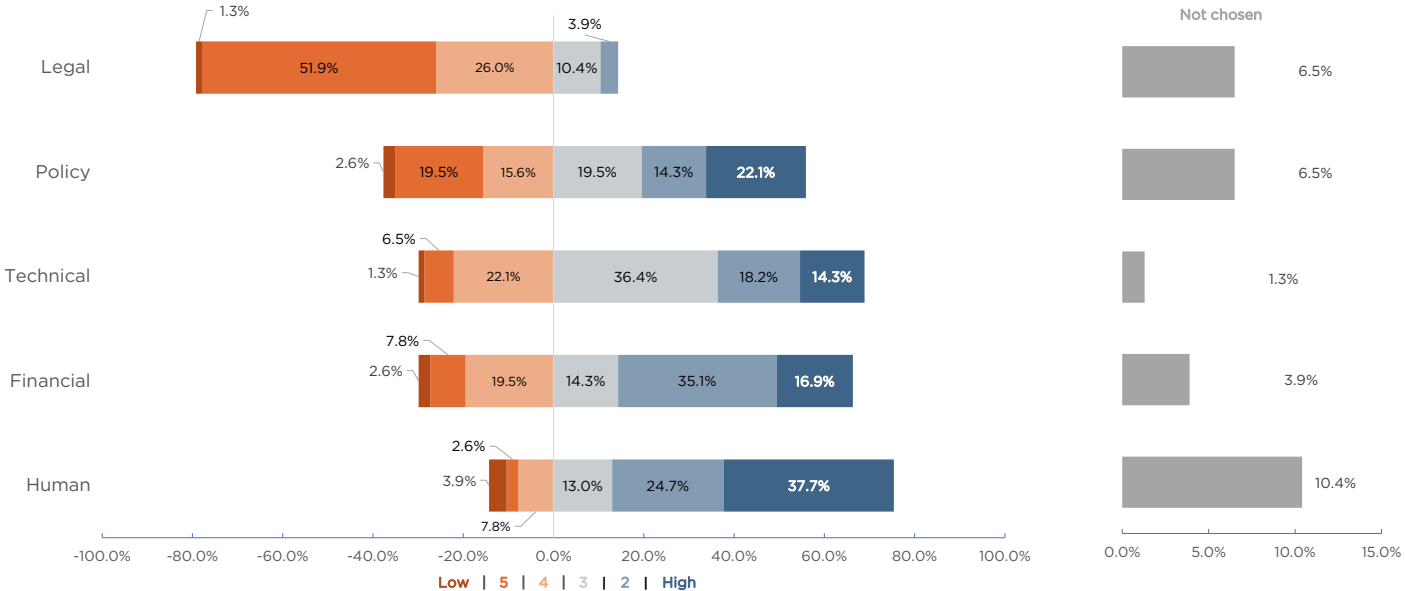

Figure 1. RDM Resources Required. Q11. "What resources are required at this time to streamline research data management at your institutional level? Please rank them in order of priority.". For each resource, "Not Chosen" is included as part of the figure. Also excluded is "Other" resources. (n=77)

---

[11] The questions addressed in Insights #2 referred to the skills required by personnel to deliver RDM support and referenced as 'highly qualified personnel' or HQP. The question in this report refers to personnel in the broader sense and so referred to as 'human resources'.

Investment in RDM Technology and Human Resources

Institutions were asked for their plans related to future investment in RDM technology. Figure 2 summarizes the investment in RDM technology across all respondent institutions. Overall, 46.8% of institutions either had no plans for further investment in RDM technology or did not know at this point. There were 24.7% that indicated they would adapt their investment strategy to meet researchers' needs, while 19.5% planned to institute base level investments in RDM technology. Some 15.6% of respondents indicated that they would coordinate institutional RDM investment through centralized IT services. Prioritizing select stages of data life cycle investment and support for all aspects of data life cycle needs were the least selected options at 10.4% and 9.1%, respectively.

There were differences between colleges and universities in a few areas of RDM investment. A higher percentage of universities (25.0%) than colleges (9.5%) were committed to investing at the base level of RDM technical infrastructure and a higher percentage of universities (17.3%) than colleges (9.5%) planned to invest in institutional level coordinated through centralized IT services. Although several universities either plan to invest by prioritizing select stages of the data life cycle (15.4%) or are dedicated to support all aspects of the data life cycle (9.6%), no colleges reported plans to take these approaches to investment in RDM technology.

The survey included an open-ended question about RDM positions that institutions are planning to create within the next two years. Although human resources are identified as the highest priority by institutions to streamline RDM support, only a small number of institutions plan to create new dedicated RDM-related positions over the next two years. Among the responding institutions, 29.9% do not expect to create any RDM positions and 28.6% do not know yet if they would. Of the institutions that did note they were planning to create dedicated RDM positions (7.8%), they also mentioned the following titles: "research data management librarian", "data management coordinator", data curation specialist", repository manager".

[Q20] How institutions will invest in
RDM technology? (n=77)

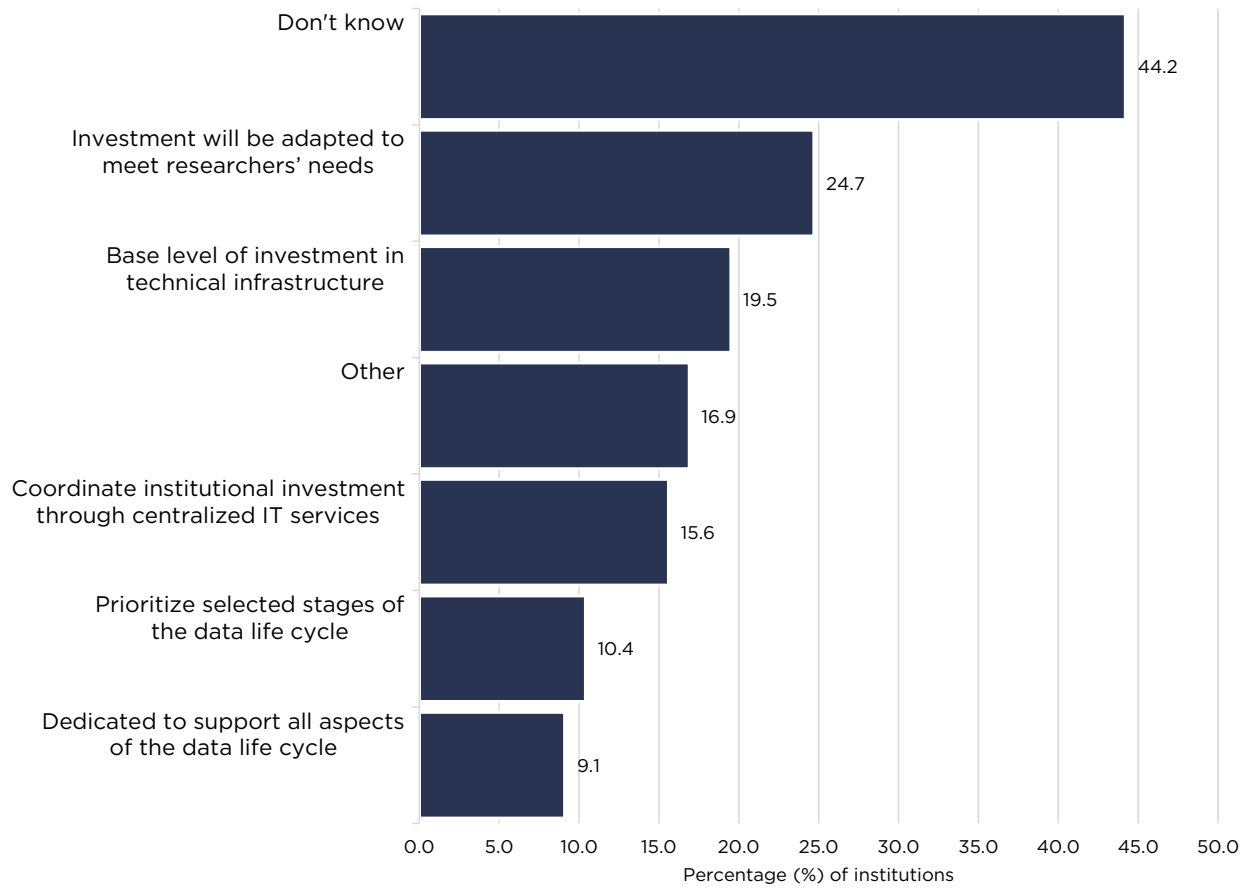| Category	Percentage
Don't know	44.2
Investment will be adapted to meet researchers' needs	24.7
Base level of investment in technical infrastructure	19.5
Other	16.9
Coordinate institutional investment through centralized IT services	15.6
Prioritize selected stages of the data life cycle	10.4
Dedicated to support all aspects of the data life cycle	9.1

Percentage (%) of institutions

Figure 2. Investment in RDM technology (all institutions). Q20. "How is your institution planning to invest in research data management technology? (check all that apply)." (n=77)

RDM Professional Development Programs for Staff and Researchers

Results revealed that institutions may need to make ongoing investments in the professional development of existing staff and researchers to stay abreast of developments in the RDM fields. A low proportion of institutions have established a professional development program for either staff or researchers (Figure 3). Universities reported some level of professional development for staff (23.5%) and researchers (19.6%) while colleges reported no professional development for staff (0%) and very little for researchers (5.3%).

Respondents who indicated that they had implemented professional development RDM programs for their staff or researchers were also asked to describe their programs. Common themes in the text responses concerning the audience and delivery are summarized below:

Staff Training
Existing RDM staff training is provided across several delivery modalities including one-to-one, hands-on workshops, webinars and asynchronous courses. Audiences included library staff as well as support staff from other units across the institution. Offerings are often led by the library, with some training offered by cross-sector units (e.g. the research office). In addition to internal offerings, institutions also funded external opportunities for staff to attend in-person or online RDM training.

Researcher Training
The commonly mentioned format of RDM training offered to researchers is free, internal and at the point of need. Respondents indicated that RDM researcher training is delivered through either hands-on sessions, training days/discussion meetings, or self-directed learning modules. There were also indications that this type of training is being directed at graduate students.
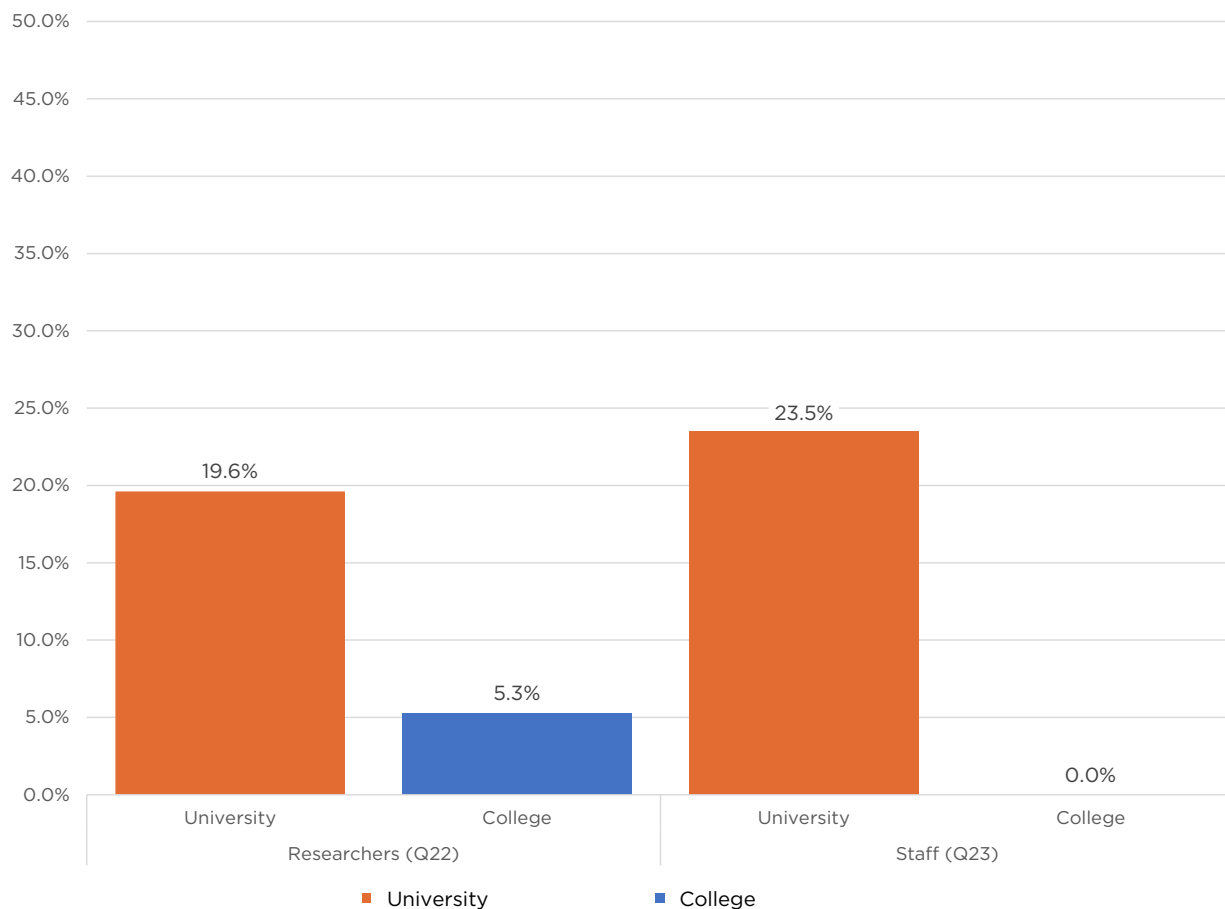
Figure 3. Professional development program on RDM for staff and researchers by Institution Type. Q22a. "Does your institution currently have a professional development program focusing on research data management for staff (for example, reskilling or new skill development)?" (n=77); and Q23a. "Does your institution currently have a professional development program focusing on research data management for staff (for example, reskilling or new skill development)?". (n=77)

Barriers and Challenges to Support RDM

Institutions were asked to identify the barriers and challenges to support RDM and adherence according to the Findable, Accessible, Interoperable, and Reusable (FAIR) principles. Figure 4 demonstrates the responses by institution type. The responses selected were categorized by the following themes:

- Awareness and resistance: the most common responses were "limited awareness of the benefits of RDM" (71.4%), "lack of awareness raising, including training opportunities" (62.3%), and "resistance to making data available or to share data" (63.6%).
- Time: the next most common response was "lack of time to clean and prepare data and metadata" (71.4%).
- Lack of incentive and funding: majority of institutions, 68.8% identified "the absence of incentives to promote RDM", 67.5% identified "lack of funding", and 59.7% identified "concerns over increased costs".
- Technical and support: over half of institutions we surveyed (63.6%) chose "Lack of skilled staff with expertise on the topic", 61.0% chose "lack of support for researchers", and 59.7% chose "technical complexity" as a challenge.
- Policy and legal: there were still certain levels of concerns in policy or legal areas: 61.0% chose "limitations of confidentiality clauses", 58.4% chose "absence of policies or guidelines at the national level", and 44.2% chose "concerns over legal frameworks".
- Coordination: coordination was reported as least concerning: 50.6% of institutions chose "different disciplinary practices", 50.6% chose "lack of coordination among the relevant actors", and 18.2% chose "lack of coordination at the regional or national level".

When examining barriers and challenges by institution type, there were some differences. The three most common barriers for universities identified were:

1. the resistance to sharing or making data available,
2. the lack of time to clean and prepare the data and metadata, and
3. the absence of incentives to promote RDM.

On the other hand, colleges cited lack of awareness-raising (including training opportunities) as the most common barrier, followed by limitations of confidentiality clauses, lack of skilled staff with expertise, limited awareness of the benefits and lack of support for researchers. Of particular note, lack of coordination at the regional or national level was the least cited barrier for both colleges and universities.
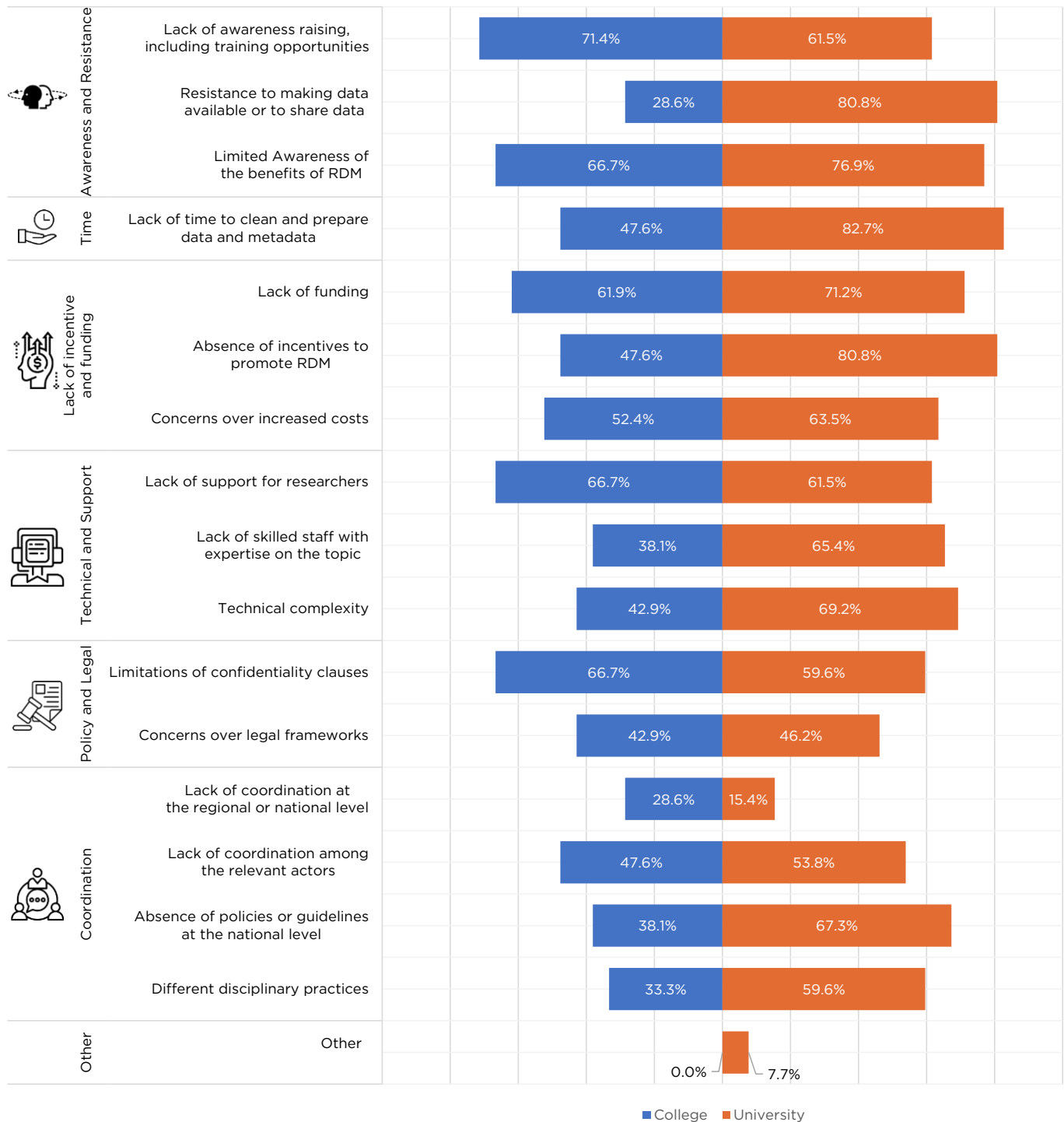
Figure 4. Challenges to supporting RDM by Institution Type. Q26. "What are the barriers/challenges for your institution to support research data management and adherence to FAIR principles (i.e. making data Findable, Accessible, Interoperable, and Reusable)? Check all that apply." (n=77)

"Awareness" icon by Luis Prado from the Noun Project., "Coordination" icon by Surya Cannavale from the Noun Project., "Save time hand" icon by Vectors Point from the Noun Project., "Incentive" icon by Eucalyp from the Noun Project., "Tech support" icon by Ben Davis from the Noun Project., "Policy" icon by Lars Meiertoberens from the Noun Project.

Level of Institutional Stakeholder Awareness

Respondents were asked to rate the level of awareness of Tri-Agency RDM policy development by various stakeholder groups at their institution. Institutional and library stakeholders had a higher level of awareness of policy development than researchers (including early-stage, mid-career, and senior researchers).

Among universities respondents, 59.6%[12] of libraries reported high levels of awareness of policy development with the institutional leadership reporting that 34.6% are aware of these developments. Awareness at the college level is lower with libraries and institutional leadership reporting on 14.3% being awareness of policy development. For researchers, it was reported that 9.6% of university researchers and 23.8% of senior college researchers are aware of policy development.[13]

Figure 5 summarizes the level of awareness of the RDM policy development by the Tri-Agencies by institutional stakeholders groups, where the percentage of positive-sentiment responses (Level 1-3) will appear as positive and the percentages of negative-sentiment responses (Level 4-6) as negative values.
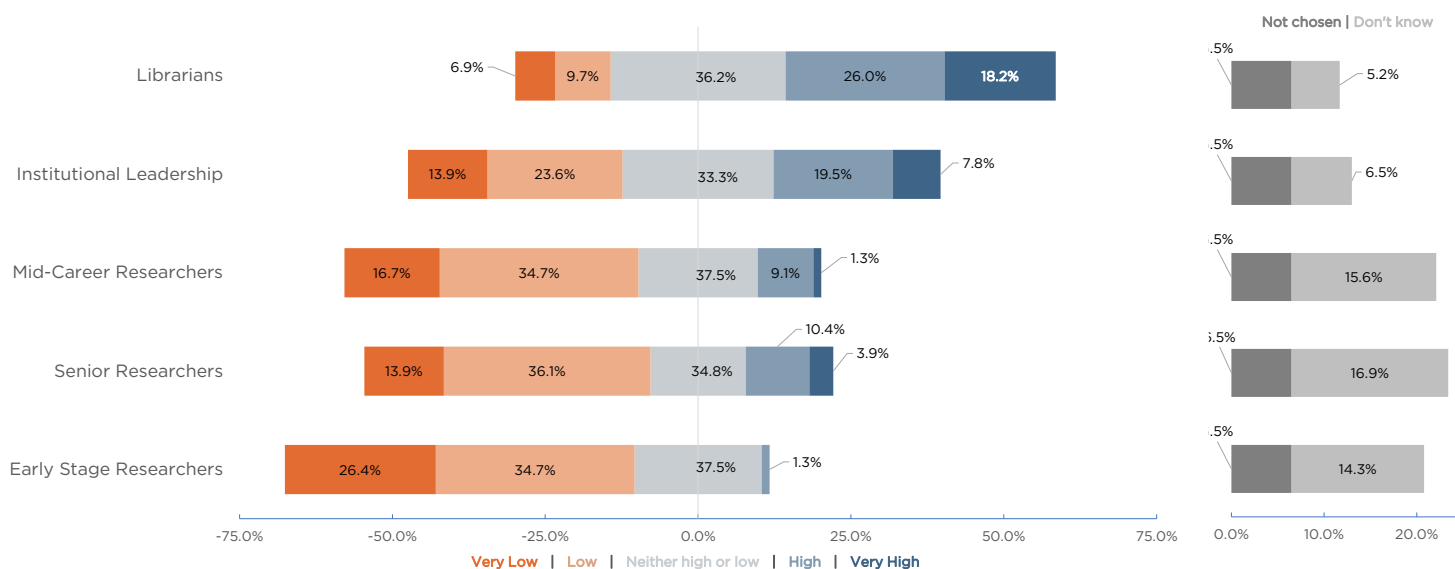

Figure 5. Level of awareness of stakeholder groups at institution. Summary of responses across all institutions Q25."Rate the level of awareness of each of the following stakeholder groups at your institution regarding Tri-Agency research data management policy development". (n=72)

---

[12] In reporting levels of awareness, high and very high were combined for all reported responses.

[13] It should be noted here that the survey was sent to only library and institutional leadership. The responses reporting on researchers are from the library and institutional leadership's point of view.

Accelerators for RDM Development

Implementation of RDM best practices may need coordination of many stakeholder groups within the research ecosystem. Our survey asked respondents to list the most important actions that could help promote RDM and adherence to the FAIR principles at their institutions. They were also asked to reflect on the actions at their own institutions, as well as regionally, nationally, internationally, and within the disciplinary communities. Respondents were not required to comment on each level but could provide as many answers as relevant in the open-text fields. Few respondents commented on the international and discipline levels. Analyzed text was grouped into the categories shown in Figures 6-10.

Overall, funding was the most frequently identified driver overall with national funding as the most effective action to promote RDM. Training, collaboration, and policy all shared the second most frequent response and the third most common category of reported accelerators include IT/infrastructure, leadership and direction and increased staff/staff time, although responses in these areas were low.

Accelerators with more than 10 mentions at each stakeholder level are reported below, and the number in parenthesis refers to the number of mentions.

Institutions
- Policy (21)
- Funding, staff, and training (16), respectively
- Leadership (13)

Regional
- Collaboration (12)
- Training (11)

National
- Funding (37)
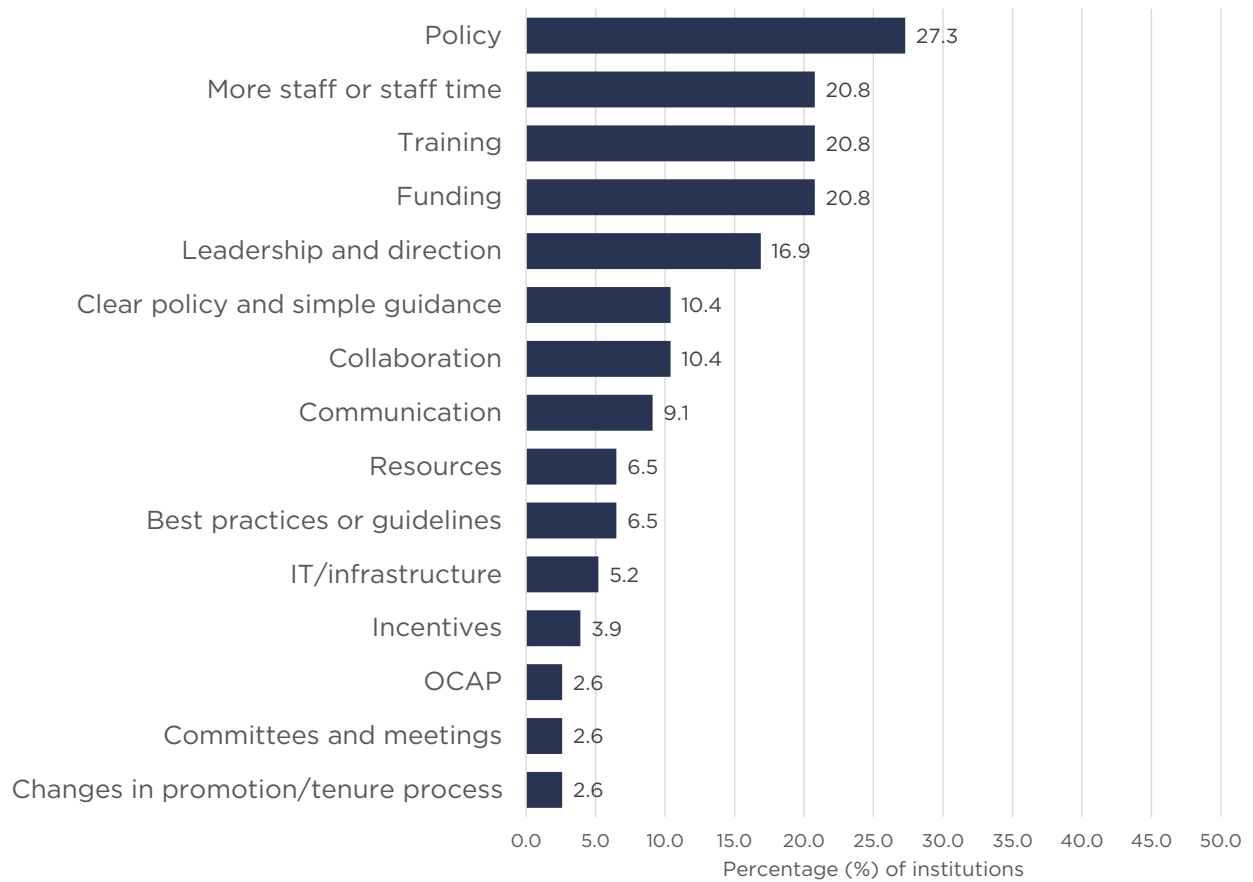- Collaboration (13)
- IT infrastructure (12)

Figure 6. Institutional accelerators for RDM development (all institutions). Accelerators are derived from text answers to Q27. "What are the most important actions that could help promote research data management and adherence to FAIR principles (Findability, Accessibility, Interoperability, and Reusability) of research data at your institution?" (n=77)
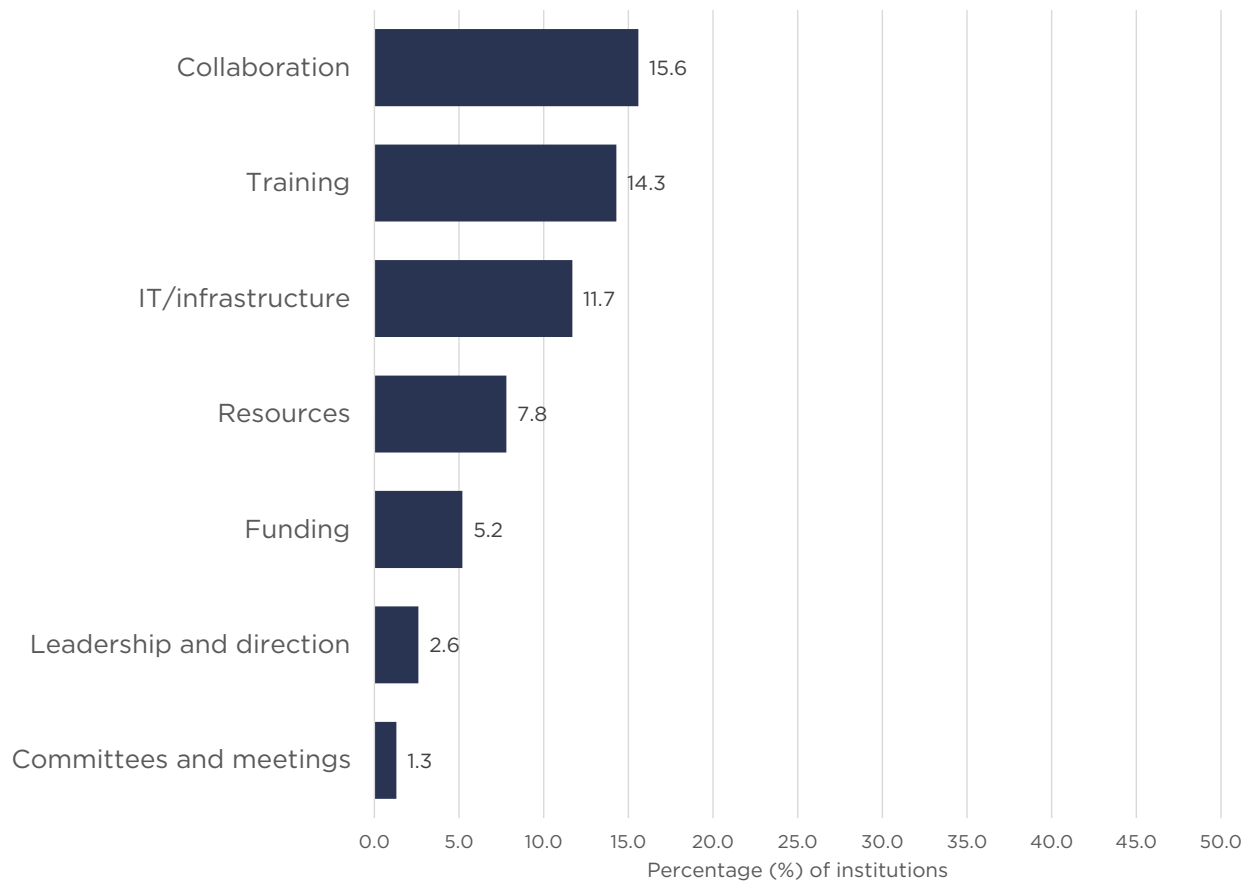
Figure 7. Regional accelerators for RDM development (all institutions). Accelerators are derived from text answers to Q27. "What are the most important actions that could help promote research data management and adherence to FAIR principles (Findability, Accessibility, Interoperability, and Reusability) of research data at your institution?" (n=77)
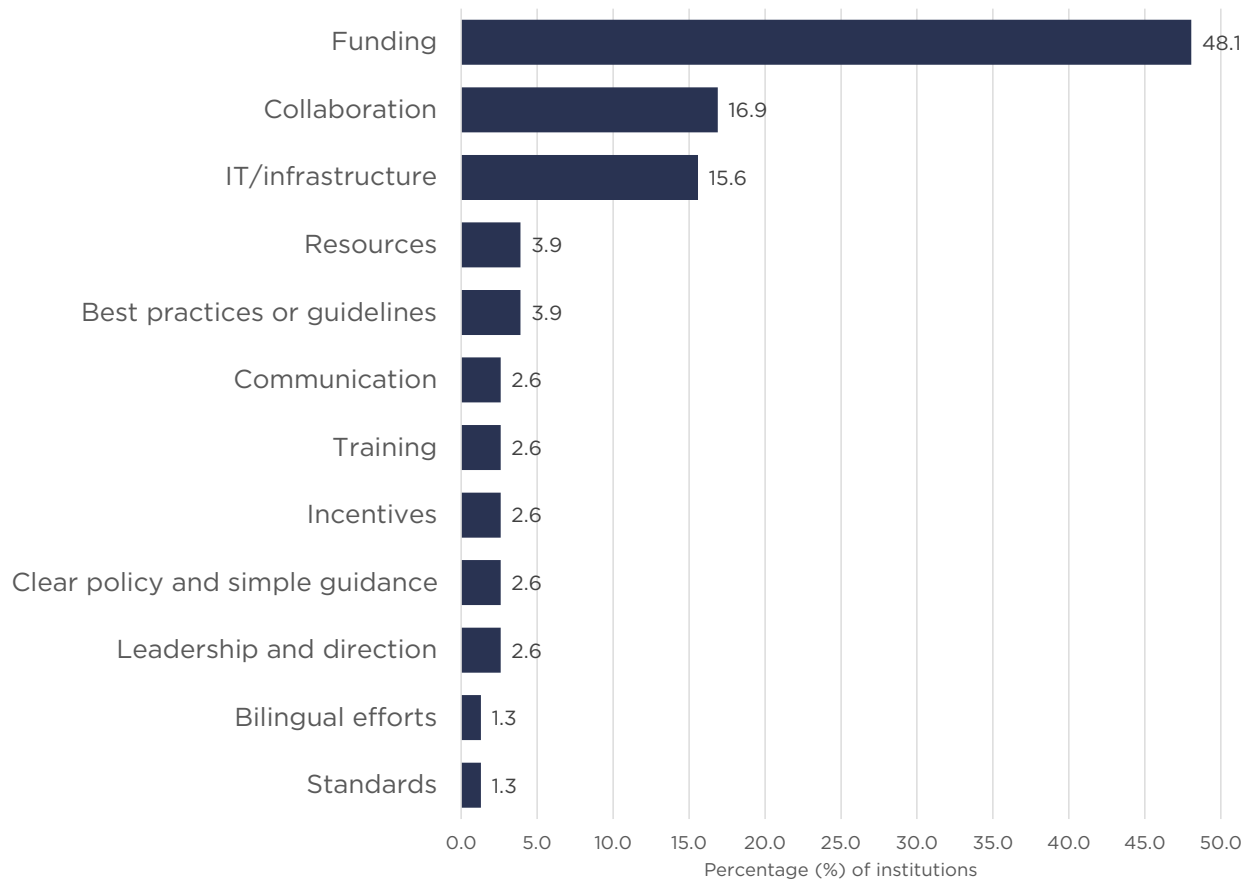
Figure 8. National accelerators for RDM development (all institutions). Accelerators are derived from text answers to Q27. "What are the most important actions that could help promote research data management and adherence to FAIR principles (Findability, Accessibility, Interoperability, and Reusability) of research data at your institution?" (n=77)
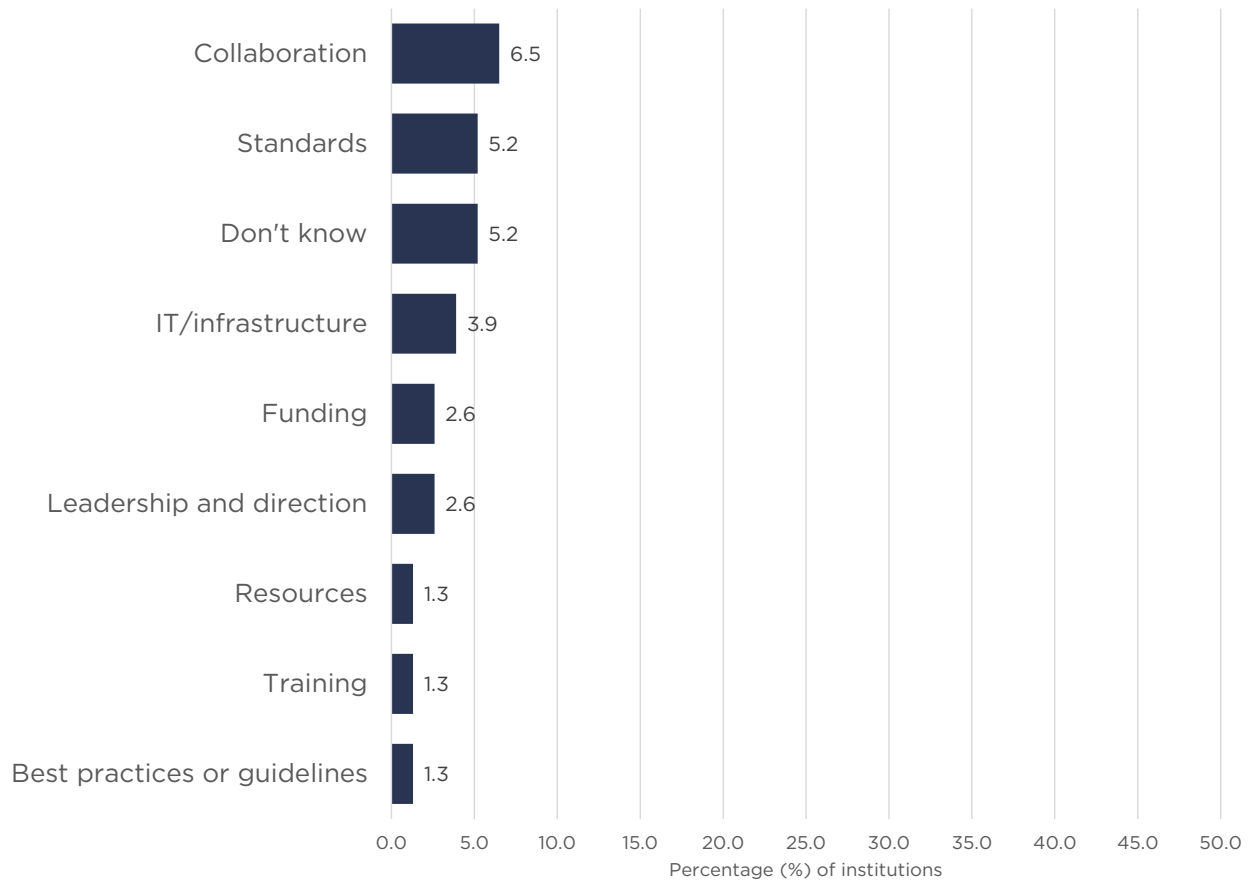
Figure 9. International accelerators for RDM development (all institutions). Accelerators are derived from text answers to Q27. "What are the most important actions that could help promote research data management and adherence to FAIR principles (Findability, Accessibility, Interoperability, and Reusability) of research data at your institution?" (n=77)
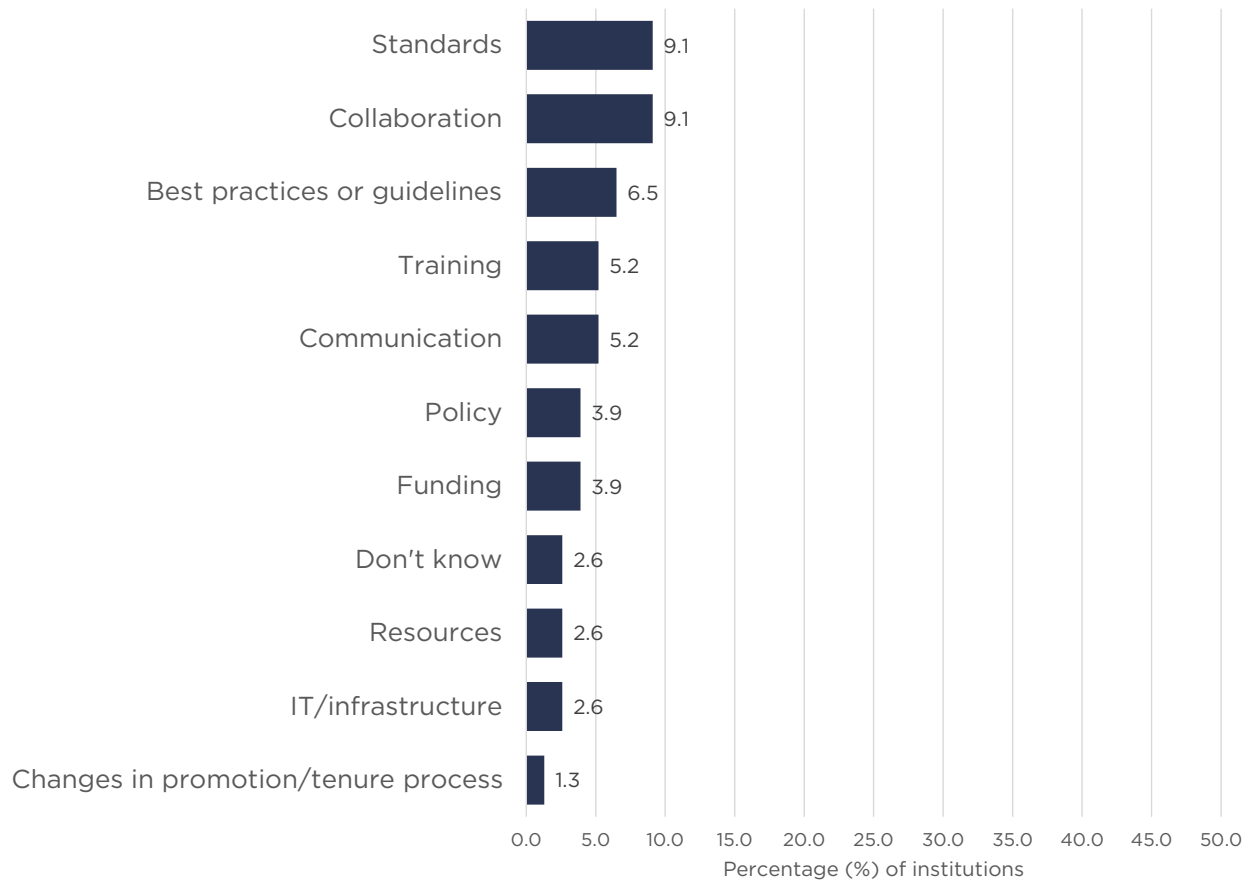
Figure 10. Within disciplines accelerators for RDM development (all institutions). Accelerators are derived from text answers to Q27. "What are the most important actions that could help promote research data management and adherence to FAIR principles (Findability, Accessibility, Interoperability, and Reusability) of research data at your institution?" (n=77)

Conclusion

This report is the final Insights Report for the Institutional RDM Capacity Survey. The survey was developed to examine RDM capacity at Canadian institutions. Throughout the three Insights Reports, survey results were examined to report on the existing status of RDM support development and to set a benchmark to compare future survey results after the Tri-Agency Research Data Management Policy implementation.

In general, the results have shown that RDM services are being built from the ground up. At the same time, institutions are actively participating in regional and national networks of infrastructure and training resources. Universities appear to be taking the lead in the development of these services, relying on rearranging current staff responsibilities to provide support to researchers. Engagement in the institutional RDM landscape is still new to researchers, disciplinary stakeholders, and regional, national, and international connections.

Few institutions have begun development of RDM strategies or policies and even fewer have finalized them. Institutions appear to be mobilizing in this direction however, and policies that are being formed concentrate on data storage, personal and sensitive data, and FAIR sharing, with much less emphasis placed on legal and licensing aspects. Limited involvement of institutional legal counsel in current strategy development has led to a lack of guidance or policies on data ownership, data retention periods and data licensing.

In terms of RDM infrastructure, college respondents are generally less equipped than university respondents. Active data storage is decentralized in most institutions and provided through researchers' departments or schools. Disciplinary repositories are widely used by researchers within most responding institutions, while some universities are also using consortial, journal, and institutional repositories. Long term data preservation infrastructure is widely lacking for all institutions.

Existing RDM services are offered through libraries often in conjunction with available national resources. While resources needed to provide RDM support are currently limited across Canada, universities are forging ahead in the development of RDM services and infrastructure. Colleges presently remain at the development and planning stage for RDM services and infrastructure. Universities engage more often than colleges in regional and national collaborative initiatives.

None of the responding institutions reported having a full complement of HQP in all areas under consideration. Specific skills shortfalls exist in the areas of managing sensitive data, data curation, and research software development.

Looking to the future, the top priorities to increase RDM capacity in institutions were identified as:

- Highly Qualified Personnel (HQP)
- Financial support
- Infrastructure and technology
- Policy guidance

When asked about specific plans for near-term investment in HQP and technology, however, very few respondents reported plans to hire new RDM personnel or improve technology/IT infrastructure at their institutions.

Institutions perceived the biggest challenges to RDM development to be a lack of awareness of RDM benefits, training opportunities, and resistance among researchers to participate in RDM activities - in particular, to share or make data openly available. Barriers and obstacles included:

- Lack of time
- Incentives
- Support
- Guidance for researchers to further engage in RDM best practices

Potential drivers that could help accelerate further RDM support development included:

- National and institutional level infrastructure and funding
- Coordination at the national and regional levels
- Policies, staff, and training
- Leadership at the institutional level

Recommendations

Recommendations already implemented or being implemented by NDRIO-Portage

- Low levels of RDM capacity for data curation were reported by respondent institutions.[14] To close this gap, NDRIO-Portage has initiated development of a national curation network.
- Technical skills associated with managing sensitive data were identified as lacking.[15] To increase institutional capacity for managing sensitive data, NDRIO-Portage established a Sensitive Data Expert Group to help develop guides on managing sensitive data, hired a policy, privacy and sensitive data coordinator, and is developing additional features to support sensitive data deposit using the [Federated Research Data Repository (FRDR)](16). A number of key outputs, including a [Sensitive Data Toolkit for Researchers](17), have been developed and deployed.
- Streamlined interdepartmental communication and coordination within an institution would accelerate both the formation and implementation of an RDM strategy framework, which in turn would raise awareness, fill service gaps, and improve service support. NDRIO-Portage has developed [guidance and training documents](18) for creating institutional RDM strategies and policies through the work of the [National Training Expert Group](19) and the [Research Intelligence Expert Group](20).

Recommendations not yet implemented

- Respondent institutions reported needing help to increase RDM capacity across the board. While institutions have been making good use of NDRIO-Portage training resources, more resources and directed training are needed.
- Colleges reported needing help to increase RDM awareness.[21] NDRIO-Portage can encourage participation of colleges in the national network of experts leading to more opportunities for collaboration and increased RDM awareness and capacity.

---

[14] Insights Report 2, page 7. http://doi.org/10.5281/zenodo.4570030
[15] Insights Report 2, page 12. http://doi.org/10.5281/zenodo.4570030
[16] Federated Research Data Repository. https://www.frdr-dfdr.ca/repo/
[17] Sensitive Data Toolkit for Researchers.
https://portagenetwork.ca/network-of-experts/sensitive-data-expert-group/
[18] Tools and Resources: Institutional Strategies.
https://portagenetwork.ca/tools-and-resources/institutional-strategies/
[19] National Training Expert Group.
https://portagenetwork.ca/network-of-experts/portage-training-expert-group/
[20] Research Intelligence Expert Group.
https://portagenetwork.ca/network-of-experts/research-intelligence-expert-group/
[21] See Insights Report 3, page 11.

- Access to and knowledge of available research software was flagged as an important area requiring resources to increase RDM capacity within respondent institutions.[22]
- While active in regional and national RDM collaborations, institutions reported being relatively less engaged in disciplinary and international networks.[23] Recognizing the importance of these networks, NDRIO-Portage has developed and recently launched disciplinary templates in further support of the DMP Assistant. [24] NDRIO-Portage should play a continuing role in advancing disciplinary and international connectedness, and encouraging engagement with key resources, standards, and initiatives
- Financial support for both HQP and technology were identified as important institutional RDM priorities. Currently, institutional financial support flows primarily at the departmental/unit level. [25] National-level RDM support provided by NDRIO-Portage will help free up institutional/departmental resources to respond to local RDM needs. National-level support also encourages cross-departmental cooperation at institutions -- the Portage RDM Strategy Template is an example of this, as it promotes multi-stakeholder engagement in the strategy-building process. As a further example of national level support for RDM, NDRIO-Portage has launched two funding calls. The first will support Canadian repositories in meeting CoreTrustSeal certification requirements. [26] The second will support improved curation of Canadian research data related to the COVID-19 pandemic, with the goal of making these data FAIR (Findable, Accessible, Interoperable, and Reusable).
- Local RDM champions and other outreach programs are needed to further engage researchers from all disciplines. NDRIO-Portage could develop specialized training to help RDM service providers create outreach programs to engage not only researchers, but the HQP who interact with them. (e.g. research and ethics offices, IT, and libraries).
- To measure progress in the evolution of the institutional RDM capacity across institutions, NDRIO-Portage plans to re-administer a similar survey sometime after the launch of the Tri-Agency Research Data Management Policy.[27] This will help NDRIO-Portage identify areas of success and areas in need of further attention.

---

[22] Insights Report 2, pages 12-14. http://doi.org/10.5281/zenodo.4570030
[23] Insights Report 2, page 18. http://doi.org/10.5281/zenodo.4570030
[24] DMP Assistant. https://assistant.portagenetwork.ca/
[25] Insights Report 1, page 4. https://doi.org/10.5281/zenodo.3906470
[26] CoreTrustSeal. https://www.coretrustseal.org/
[27] Tri-Agency Research Data Management Policy.
http://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html

Research institutions play a crucial role in developing and implementing RDM support in Canada and are essential to our achieving the culture change needed to bring RDM fully into the research life cycle. This survey highlights the complex nature of the RDM landscape, punctuated by a variety of specialized and often siloed stakeholders, disciplines, and jurisdictions, each with their own needs and priorities. NDRIO-Portage through its national RDM services, platforms, policies, and training is actively working to increase communication and understanding among these RDM players, and to provide practical support to researchers and to those responsible for supporting them.