



McMaster University

Research Data Management Institutional Strategy Report

The Current State of Research Data Management (RDM) at McMaster

Prepared by the McMaster RDM Institutional Strategy Working Group

Version 3

Last Revision: November 3, 2022

Table of Contents

Executive Summary	3
Introduction and Context.....	5
Summarized Environmental Scan	6
McMaster RDM Stakeholders	6
Current RDM Requirements	7
Data Collection Methodology.....	7
RDM Maturity Assessments	7
The McMaster RDM Needs Survey	8
RDM Focus Groups.....	8
Results – MAMIC.....	9
Institutional policies and procedures	9
IT Infrastructure.....	9
Service and Support	10
Financial Support.....	11
Results - Survey.....	11
Section 1: Demographics of the survey respondents	11
Section 2: Data management in research grants	12
Section 3: Data storage and backup	13
Section 4: RDM roles and responsibilities	17
Section 5: Data sharing	18
Section 6: RDM support services at McMaster	19
Section 7: Challenges and barriers in managing data.....	20
Results – Focus Groups	21
Theme Overview	21
Data Storage	22
Data Archival	23
Group Storage	23
File Transfer.....	23
Applications.....	23
Data Security.....	24
Sensitive Data.....	24
Data Ethics & Privacy	24
Encryption & Cyber Hygiene.....	24
Data Sharing	25
McMaster RDM Services.....	26
Awareness.....	26
Equitable Access.....	26
Integrating Local and Central Services.....	26
Training.....	27
Consultations	27
Conclusions and Next Steps	27
References	28
Appendices	30
Appendix 1: Survey design	30
Appendix 2: Focus Group script (researcher).....	30
Appendix 3: Focus Group script (research support staff)	30
Appendix 4: Environmental Scan - McMaster RDM Stakeholders and Service Providers	31
McMaster University Library	33
Academic.....	34
Research.....	34
Research Centres and Institutes.....	35
UTS – University Technology Services.....	36
Faculty of Health Sciences	37
Appendix 5: RDM - National + International Context.....	38
National Level.....	38
International Level.....	39
Appendix 6: Current RDM Requirements	39
Data Management Plans.....	39
IT Security.....	39
Data Deposit	40
Appendix 7: Definitions	40

Executive Summary

Research Data Management (RDM) is a suite of connected processes and practices, which are applied throughout the research lifecycle—i.e., as data is planned for, collected, organized, documented, stored, preserved, shared, and reused—in support of analysis, research, and dissemination that is beneficial to society. This process increases research visibility, generates new collaborations, enables verification of results, and fosters a culture of reproducible research. RDM is carried out by researchers—whether faculty, students, staff, or community members—and is supported by staff in research and privacy offices, institutional IT units, and libraries.

To support high professional and disciplinary standards that make good use of public funds, the Tri-Agencies have released a [Research Data Management Policy](#). (I. Government of Canada, n.d.) This policy has three pillars: data deposit, Data Management Plans (DMPs), and a requirement for each institution eligible to administer Tri-Agency funds to develop and share an **Institutional RDM Strategy** by March 2023. Strategies are expected to emphasize best practices and outline the requisite institutional support to apply them; aligning throughout with disciplinary norms, ethical/legal/commercial obligations, and principles of Indigenous data sovereignty.

With the goal of developing McMaster’s Institutional RDM Strategy, a Working Group (ISWG) was assembled in April 2021 with representatives from faculties, research support offices, ethics boards, IT support units, libraries, and McMaster University’s partner research hospitals. The ISWG is following steps laid out in the Digital Research Alliance of Canada (DRAC)’s Institutional Research Data Management Strategy [Development Template](#). (I. R. S. T. R. W. Group, 2021)

This document represents the culmination of Stage 2 in the development process, which relates to understanding and describing **the current state of RDM at McMaster**. It situates McMaster RDM in context and shares results from four assessment activities: [an environmental scan](#) of RDM stakeholder groups, requirements, and services within and beyond McMaster, [a scale and maturity assessment](#) of McMaster’s extant RDM services, [a survey of researchers’ RDM needs](#), and [focus groups with researchers and research support units](#).

Our high-level findings on the current state of McMaster Research Data Management reveal several areas of potential clarification and growth. Although many of these areas overlap, they are organized here around frameworks which support research—governance and policy, funding and support, culture, community, and collaboration—as well as needs related to RDM practices, tools, and infrastructure.

Current State: Frameworks

- **Governance and Policy:**
 - **Policy:** There are a variety of McMaster policy, procedural, and guideline documents that pertain to RDM, depending on the nature of data being managed and the sources of research funding. McMaster has no singular policy that aggregates RDM obligations set by external organizations, connects them to institutional expectations and requirements, and aligns them with a unified strategy.
 - **Strategy:** Developed collaboratively and iteratively, an institution wide RDM strategy will help to clarify expectations, roles, and responsibilities for researchers and supporting units alike, as well as support future policy development and assessment activities.
 - **Communication and Assessment:** McMaster needs more effective methods of information collection and dissemination—better ways of gathering information, assessing services, and communicating to researchers—to build responsive RDM services and strategies.
- **Funding and Support:**
 - **Equitable Funding:** Currently, students, academics with smaller research budgets, and units with lower funding struggle to afford certain aspects of RDM services, such as HR capacity and data management infrastructure.
 - **Scalable Support:** Research projects with complex requirements—including those applying data-intensive research, complex sharing and collaborative arrangements, or data from human participants—all have significant and unmet RDM needs.

- **Culture, Community, and Collaboration:**
 - **Incentives:** Considering many other demands on them, researchers face challenges finding the time to learn, adapt, and apply evolving RDM best practices.
 - **Awareness:** Many respondents were not yet aware of the existence of RDM services.
 - **Communities and Champions:** McMaster’s researchers stand to benefit from more inter-group knowledge exchange about RDM practices. McMaster has an opportunity to identify and support champions to help promote RDM principles and best practices in academic units.

- **RDM Services and Training**
 - **Service Coherence:** Currently, there is an unevenness of scale and maturity in RDM-related services at McMaster. A wide variety of campus units support research, however services are not always equitable, comprehensive, coordinated, nor well-communicated.
 - **Research Commons:** Researchers need a single place to find accurate and current information, i.e., a hub connecting outwards to discipline-specific help to navigate discipline-specific challenges.
 - **Central RDM Services:** There is strong support amongst the McMaster research community for more comprehensive RDM training and services.
 - **Training Across Levels:** There is a need for more training for all researchers, but especially a need for equitable training across levels that is available to and appropriate for early careers researchers, graduate students, and postdoctoral fellows to meet the increasing expectations for data management.

RDM Practices, Tools, and Infrastructure

- **Data Management Planning:** Researchers need resources, guidelines, training, and support to build DMPs (Data Management Plans) for research projects and grant applications. While some researchers have completed a DMP for a grant, most are currently unaware of resources available to support DMP creation.
- **Data Storage and Preservation:** Secure data storage with adequate backups is an important facet of data management. Currently the most common storage location for research data at McMaster is personally managed internal hard drives. As a result, a considerable portion of the university’s research output is at risk of theft, corruption, and loss.
- **Sensitive Data Management:** Protecting research participants and sensitive datasets requires easily accessible and understandable resources, training, and practical support for ethical data management, both for researchers and service providers.
- **Documentation, Sharing, and Access:** Evolving publisher, funder, and disciplinary expectations for data sharing present new methodological, technical, and cultural challenges for research. Data curation is a critical activity for creating datasets that are FAIR (findable, accessible, interoperable, and reusable), but there is insufficient expertise and capacity across all disciplines and approaches to support current and anticipated future research needs.
- **Data Security:** IT Security is a profound concern for researchers, who are not always experts in technology and must navigate a complex and evolving web of requirements, best practices, resources, services, and tools to keep their research and research data safe.
- **Indigenous Data:** Researchers and services need to better understand support for data sovereignty, Indigenous researchers, and those working with Indigenous communities to align with Tri-Agency priorities, McMaster’s [Strategic Plan](#) and [Strategic Plan for Research](#), and community developed principles ([OCAP](#), [CARE](#)).
- **Software & Applications:** Researchers need consistent and transparent ways to access existing software and support for developing, maintaining, and sharing their own custom software.

With the completion of the second stage of our strategy development, we have a fuller understanding of RDM at McMaster: our current scale and maturity of services and researchers’ RDM needs. Our next step is to finalize a document outlining the outcomes of Stage 3: Articulating the **Ideal State** of RDM at McMaster, which will be released in July 2022. From there, we will address Stage 4: Develop **McMaster’s Institutional RDM Strategy** by outlining our path forward from the current state of RDM towards our desired future state. The draft Institutional Strategy will undergo multiple rounds of review, further engagement, and revision in July 2022 (RITC + VPR), Fall 2022 (engagement sessions, focus groups, and faculty town halls), and January 2023 (RITC + VPR), in advance of publishing the strategy in March 2023.

Introduction and Context

Research Data Management (RDM) describes a series of practices and activities—planning, collection, organization, documentation, storage, preservation, and data sharing—that are carried out by researchers—whether faculty, students, staff, or community members—and supported by staff in research and privacy offices, institutional IT units, and libraries. Applying good RDM practices improves the efficiency and impact of research, increases research visibility, facilitates collaboration, protects intellectual property, enables verification of research results, and supports a culture of reproducible research (Popkin, 2019).

Digital research data is a valuable asset in and of itself, and governments, funders, researchers, journals, research communities, and research institutions are recognizing a need to articulate policies and plans for data stewardship to support its reuse and sharing. Canada has demonstrated its commitment to data access and open research through several important national and international commitments (I. Government of Canada, 2016).

Building upon foundational documents acknowledging the importance of RDM, which include the [Tri-Agency Statement of Principles on Digital Data Management](#) (I. Government of Canada, 2016) and the [Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans](#) (Canadian Institutes of Health Research et al., 2018), the Tri-Agency published its [Research Data Management Policy](#) in March 2021. The Policy communicates expectations around three pillars of activity:

1. Implementing **Data Management Plans** as a part of grant competitions (beginning with a [pilot program for select competitions in 2022](#))
2. Building a culture of **Data deposit and sharing** where appropriate and aligned with disciplinary norms (beyond current requirements for CIHR and SSHRC). Phased implementation will begin after 2022.
3. Requiring each institution eligible to administer Tri-Agency funds to develop and publicly share an **Institutional RDM Strategy** by March 2023.

Institutional RDM Strategies are intended to emphasize good RDM practices and express how the institution will support its researchers to apply them in coherence with disciplinary norms, in alignment with ethical, legal, and commercial obligations, and according to principles of Indigenous data sovereignty. The agencies expect the institutional strategies to be high quality and outline how their policies and supports will enable researchers to follow RDM best practices.

Strategies are expected to be reviewed and revised regularly as practices evolve, while also considering the multiple facets of RDM within the larger context of our national research enterprise. These strategies center around recognizing data as a valuable research output. Practices to support this include promoting the importance of data management to researchers, staff, and students; guidance on data management and planning; developing institutional standards and policies; and providing access to repository services. It also requires alignment with ethical, legal, and commercial obligations, principles of Indigenous data sovereignty, and Tri-Agency requirements.

Under the governance of its Research Information Technology Committee and Office of the Vice President, Research, McMaster University assembled an Institutional Strategy Working Group (ISWG) in April 2021 to lead its Strategy development. The working group has followed the Digital Research Alliance of Canada (DRAC)'s [Institutional Research Data Management Strategy Development Template](#), (I. R. S. T. R. W. Group, 2021), which comprises the following activities:

- Stage 1: Assemble a Strategy Development Team
- Stage 2: Assess the Current State of RDM**
- Stage 3: Envision the Future State of RDM
- Stage 4: Articulate the Institution's Path Forward
- Stage 5: Assemble and Launch the Strategy

This document represents the culmination of Stage 2 in the development process, which relates to understanding and describing the current state of RDM at McMaster. It situates McMaster RDM in context and shares results from four assessment activities: [an environmental scan](#) of RDM stakeholder groups, requirements, and services within and beyond McMaster, [a scale and maturity assessment](#) of McMaster's extant RDM services, [a survey of](#)

[researchers' RDM needs](#), and [dozens of focus group interviews](#) with researchers and research support units. The assessment activities were built upon, informed by, and contextualized against previous IT services and research technology reviews that were conducted at McMaster in 2016 and 2019, respectively.

Summarized Environmental Scan

This section provides a quick overview of the RDM environment at McMaster in the context of this Current State report. Further details are included in the appendices. This includes several focuses. [Appendix 4: Environmental Scan – McMaster RDM Stakeholders and Service Providers](#) shares information about the many different groups providing RDM training, support, and services. [Appendix 5: RDM - National + International Context](#) details how RDM at our university fits into the bigger picture of RDM services federally and globally. We also highlight current expectations for RDM planning and practices in [Appendix 6: Current RDM Requirements](#).

McMaster RDM Stakeholders

While Research Data Management is an established aspect of any research process, RDM needs are expanding due to research developments (amount and complexity of data) and external pressures (funding agency requirements, data privacy legislation, data sharing agreements, and more). Because RDM is a multifaceted endeavour connecting across all disciplines and multiple services, the stakeholder groups responsible for RDM at McMaster are expansive and diverse. They include:

- **Researchers:** Faculty Members, Postdoctoral Fellows, Graduate Students, Research Staff
- **IT:** [Research and High-Performance Computing Support \(RHPCS\)](#), [University Technology Services \(UTS\)](#), [Computer Services Unit \(CSU\)](#), [IT Security](#), Faculty IT units
- **Research Services:** [Research Office for Administration and Development Support \(ROADS\)](#), [McMaster Industry Liaison Office \(MILO\)](#), [Health Research Services \(HRS\)](#), School of Graduate Studies (SGS)
- **Ethics:** [McMaster Research Ethics Board \(MREB\)](#), [Hamilton Integrated Research Ethics Board \(HiREB\)](#), [Animal Research Ethics Board \(AREB\)](#)
- **Research Centres & Institutes:** [McMaster Indigenous Research Institute \(MIRI\)](#), [Sherman Centre for Digital Scholarship \(SCDS\)](#), [MacData](#), [Spark](#), [Secure Empirical Analysis Lab \(SEAL Lab\)](#), [Canadian Research Data Centre Network \(CRDCN\)](#), [Research Data Centre \(RDC\)](#)
- **Libraries:** [Health Sciences Library \(HSL\)](#), [McMaster University Library \(MUL\)](#), [McMaster Dataverse](#), [Data Analysis Support Hub \(DASH\)](#), [MacSphere](#)

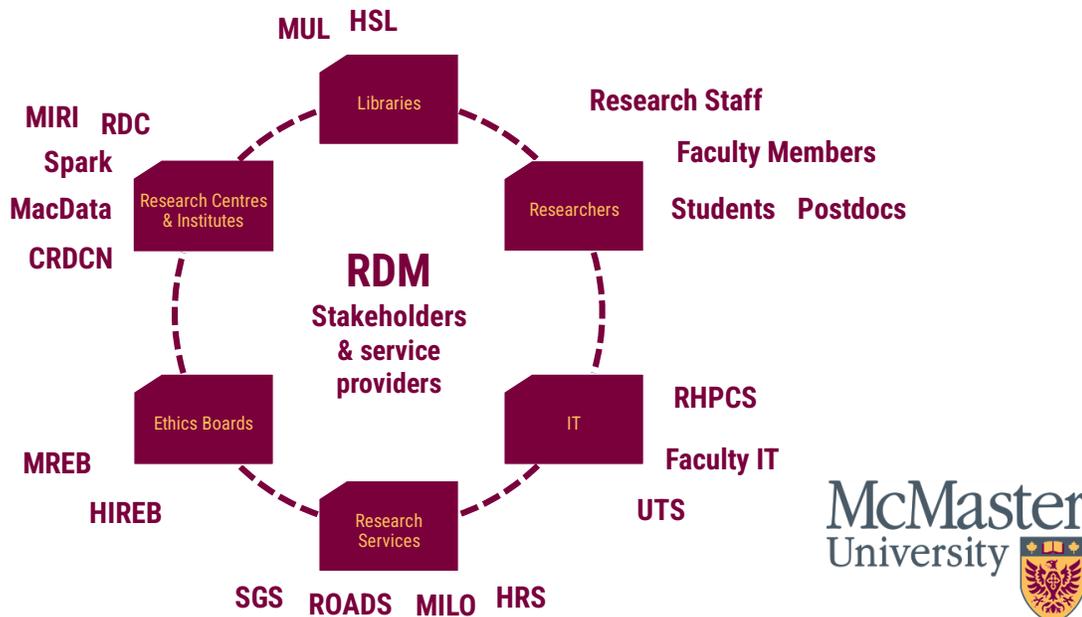


FIGURE 1: OVERVIEW OF McMASTER'S INSTITUTIONAL RDM LANDSCAPE

More information on research service providers and their intersection with RDM support is included in Appendix 4: Environmental Scan – McMaster RDM Stakeholders and Service Providers.

In addition to institutional-level stakeholders and service providers, McMaster’s research enterprise also connects with national and international stakeholders. These are outlined in [Appendix 5: RDM - National + International Context](#). Within this national context, this research is also informed by several documents and needs assessment summaries conducted first by Portage and then by the Digital Research Alliance of Canada (DRAC). These include the [Portage Research Intelligence Expert Group Institutional RDM Strategy Survey](#) (P. R. I. E. Group, 2019), [Institutional Research Data Management Services Capacity Survey](#) (Cooper et al., 2020), DRAC’s [The Current State of Research Data Management in Canada](#) (Shahira Khair et al., 2020), and the DRAC [Researcher Needs Assessment](#) (Felipe Pérez-Jvostov et al., 2021)

Current RDM Requirements

In addition to institutional requirements, researchers face other requirements for RDM depending on project specifications and external pressure. There is increasing scrutiny from research ethics on data management as more research becomes digital and funders and journals have increasing RDM requirements. Requirements could include requesting Data Management or Information Security Plans as part of the competition application process or requiring research to be deposited in a trusted data repository. National funders like the Tri-Agencies, provincial and non-profit funding agencies, industry partners, as well as several major publishers and journals consider RDM an important facet of research. An overview of the current state of these requirements is included in [Appendix 6: Current RDM Requirements](#).

Data Collection Methodology

The information presented in this current state document was acquired and synthesized from four assessment activities: environmental scan, maturity and scale self-assessment, RDM survey, and RDM focus groups). This data collection process builds on the extensive groundwork done to outline the landscape of IT services at McMaster, both IT services more broadly, as well as specific considerations for research IT. These fact-finding exercises resulted in the [IT Services Review \(November 2016\)](#) and the Research IT Services Review documents. Our work builds on these two review documents. The methodologies undertaking to collect information for this synthesis are described below.

RDM Maturity Assessments

In an effort to build a baseline evaluation of the current state of McMaster’s RDM policy, procedure, infrastructure, services, and funding landscape, the ISWG carried out assessments using the Research Infrastructure Self Evaluation (J. Rans & A. Whyte, n.d.) (RISE) and the [Maturity Assessment Model in Canada](#) (Fry et al., 2021) (MAMIC) frameworks. Released in late 2021, the MAMIC builds on the RISE model to provide a specific, contextually relevant, and standard tool for evaluating the **maturity** and **scale** of RDM supports within Canadian institutions.

The MAMIC was developed into a survey using REDCap software and initially distributed to members of the ISWG, who further shared it with contacts in their respective units or Faculties. The survey was also shared with IT leads within UTS and the Faculties, with a similar request to forward to others knowledgeable enough to assess the MAMIC elements. A total of 12 completed responses were aggregated to provide a mean **maturity** and **scale** score for each of the 26 elements, as well as free text comments relating to them. Maturity and scale assessments were refined following further discussion within the ISWG. The maturity and scale scores and score classification ranges are shown in Table 1.

TABLE 1: MATURITY AND SCALE CLASSIFICATION METHODS FOR AGGREGATED MAMIC RESULTS.

Maturity level	Classification
0 - Does not exist	Low (0 – 2)
1 - Element is not formalized or is ad hoc.	
2 - Element is under development.	Moderate (2 – 3.5)

- 3 - Element is operationalized and launched.
- 4 - Element is robust and focuses on continuous evaluation.
- N/A - Not applicable
- DNK - Do not know.

High (3.5 – 4)

Scale	Classification
0 - Not applicable (i.e. if 0 or NA were chosen for <i>Maturity Level</i>)	Low (0 – 1.5)
1 - Offered only to specific users upon request.	Moderate (1.5 – 2.5)
2 - Available within certain units or cohorts.	Moderate (1.5 – 2.5)
3 - Available to everyone.	High (2.5 – 3)

The McMaster RDM Needs Survey

From the start of the Strategy Development process, we saw a clear need to hear directly from researchers. We chose to run an online survey to try to understand the present and future Research Data Management (RDM) practices, challenges, and needs of researchers working within the institution.

The survey asked researchers about their awareness of funding mandates for research data management, how they store and back up their data, how they and their trainees and collaborators work with research data, how they share their data, the barriers and challenges they face working with research data, and their awareness and interest in RDM services at McMaster. The full survey structure can be found in [Appendix 1: Survey Design](#).

The survey was launched on November 24th, 2021 and closed on January 7th, 2022. It was advertised through McMaster news stories, emails to faculty, graduate students, and staff through the office of the Vice-President, Research, through the Working Group members, and through the faculties. The survey received 156 responses, of which 105 were complete.

The survey also asked respondents if they would be willing to meet with the McMaster RDM Institutional Strategy Working Group for a short follow-up interview to discuss their data management practices. We received 26 responses indicating they would be willing to meet in a focus group setting. This directly led us into our next phase of information gathering.

RDM Focus Groups

While the survey data provided an overview of current RDM practices, it lacked context and we recognized a need to capture a fuller picture of the diverse challenges and needs of researchers at McMaster. To add depth and context to the survey results, we decided to run a series of focus groups. These conversations and discussions with researchers and other stakeholders provided a space for researchers to articulate their context and their priorities. Other institutions such as Concordia (Dennie & Guindon, 2017) have had success a similar approach.

The response to the initial set of interview requests was high so we chose to structure our interviews as focus groups of 4-5 individuals. This allowed us to capture responses from all attendees while accommodating everyone who expressed interest in order to get the broadest range of perspectives possible. We interviewed faculty, graduate students, postdoctoral fellows, research staff, and research and IT support staff from all 6 Faculties at McMaster.

In total, we ran 21 researcher focus groups with 58 participants and 8 research support focus groups with 33 participants. Focus groups were aimed at gathering information on researchers' RDM challenges and needs, and how McMaster services could help them meet those needs. Each focus group followed a predetermined script (Appendices 2 and 3), although we were flexible and often allowed participants to direct the conversation according to their own interests.

Results – MAMIC

The aggregated survey results for each of the four MAMIC categories are shown in Tables 2 through 5, accompanied by a summary that incorporates additional feedback from respondents and further discussion within the ISWG. While most estimations of maturity and scale have considerable uncertainty (respondents noted challenges identifying and contextualizing the relevant aspects of elements), the exercise helped identify areas of notable weakness and strength, and also provided a comparative benchmark for future assessment activities. General outcomes for each category are summarized below:

Institutional policies and procedures

TABLE 2: SUMMARIZED AGGREGATED MAMIC ASSESSMENT FOR ELEMENTS WITHIN THE INSTITUTIONAL POLICIES AND PROCEDURES CATEGORY

Element	Maturity	Scale
Institutional RDM Strategy: As defined by the Tri-Agency. This includes any Institutional RDM roadmap detailing how the strategy will be implemented.	N/A	N/A
Institutional RDM-related Policies: Includes all relevant policies at the institution that may address RDM or components related to RDM.	Low	High
Data Management Planning-related Procedures and Guidelines: Any institutional procedures or guidelines that outline how researchers should address data management plans (e.g., expectations of DMP creation, submission and/or review).	Low	Moderate
Security and Risk Assessment Policies and Procedures: Any institutional procedures or policies that address security and risk assessment related to research data (e.g., legal and privacy issues, vulnerability assessments, etc.).	Low	Moderate
Communication and Outreach Plan: Any plans for the promotion of RDM. This may include raising awareness of national policies and guidelines that affect RDM (e.g., Tri-Agency policies, funder policies, journal policies), and providing links and resources for best practices and tools.	Moderate	High

Summary: The current environment is complex and uneven, as a variety of policy, procedure, and guideline documents pertain to RDM, depending on the nature of data being managed and sources of research funding. Furthermore, awareness of relevant policies and procedures varies significantly across the institution. Currently, McMaster has no singular policy that aggregates RDM obligations set by external organizations (e.g., provincial and national legislation, tri-agency policy), connects them to institutional expectations and requirements, and aligns them with a unified strategy. Efforts by research administration offices (ROADS, HRS, MILO), UTS, RHPCS, the University Library, and Faculty IT units have helped to address some IT security and risk assessment needs of researchers, but more development and coordination is required.

IT Infrastructure

TABLE 3: SUMMARIZED AGGREGATED MAMIC ASSESSMENT FOR ELEMENTS WITHIN THE IT INFRASTRUCTURE CATEGORY

Element	Maturity	Scale
Access / Authentication / Authorization: Availability of infrastructure and associated services covering identity management and verification, and mediated access to research data according to security needs.	Moderate	Moderate
High Performance Computing: Offering aggregate computing power that delivers higher performance than a typical desktop workstation.	Moderate	Moderate
Active Storage: Institutionally managed storage space for active projects.	Moderate	High
Active Storage - Sensitive Data: Institutionally managed secure storage space for active projects that have sensitive or confidential data requirements.	Low	Moderate

Collaborative Digital Workspace: Private digital space for researchers to collaborate on research projects. This may be within the institution or across institutions.	Moderate	Moderate
File Transfer service: Mechanism for securely moving large files within and outside of the institution.	Moderate	Moderate
Repository for Data Deposit: Provided by the institution for data deposit and dissemination upon completion of the research project.	Moderate	Moderate
Sensitive Data Deposit: Provided by the institution for secure data deposit and restricted access upon completion of the research project.	Low	Moderate
Preservation / Archival Storage: Commitment and space to maintain data integrity and access over a specified period of time.	Low	High
Governance of Enterprise Software: Management of various software packages needed at different stages of the research life-cycle.	Low	Moderate

Summary: The technical requirements for research computing and RDM vary significantly across disciplines, and solutions are not universally discoverable, available, nor accessible. Many services and tools are provided by McMaster IT units, but researchers often struggle to understand what exists and is available to them. Generally, solutions are either not mature enough to meet the particular and diverse needs of McMaster researchers, or they are not scaled appropriately to all who require them. For example, units like RHPCS and SEAL provide custom solutions for researchers, but their cost-recovery model make them inaccessible to researchers without an appropriate technology budget. Cost-free national services like Compute Canada, DMP Assistant, the Federated Research Data Repository (FRDR), and the Borealis Dataverse Repository address critical research needs, but awareness of them is variable across the institution. Activities such as data transfer for large files and collections, managing and archiving sensitive data, and sharing data with collaborators outside of the university are common challenges without broad, appropriate solutions at the institution.

Service and Support

TABLE 4: SUMMARIZED AGGREGATED MAMIC ASSESSMENT FOR ELEMENTS WITHIN THE SERVICE AND SUPPORT CATEGORY

Element	Maturity	Scale
General RDM Training: Builds RDM skills among researchers through workshops, webinars, lectures, etc.	Low	High
General RDM Consultation: Individual or group sessions relating to RDM needs.	Moderate	High
DMP Training: Builds skills for researchers to effectively put together a DMP through workshops, webinars, lectures, etc.	Moderate	High
DMP Consultation: Individual or group sessions relating to developing a DMP.	Low	Moderate
Data Curation Training: Builds skills for researchers to effectively curate their data through workshops, webinars, lectures, etc.	Moderate	Moderate
Data Curation Consultation: Individual or group sessions relating to curating and preparing data for deposit (including metadata and documentation).	Low	Moderate
RDM-related software training: Builds tool and software skills for working with data across different stages of the research life-cycle through workshops, webinars, lectures, etc.	Low	Moderate
RDM-related software consultation: Individual or group sessions relating to the use of enterprise tools and software.	Low	Moderate

Summary: Components of RDM services, support, and training are offered by discrete IT and research support units, Libraries, and research centres and institutes across McMaster. While general training and consultation services are available to researchers through these groups, they can be difficult for potential users to find. Training offerings are perceived as “scattered, inconsistent, and uncoordinated across units”, and their irregular delivery schedules

don't always align with researcher needs. There is doubt that sufficient expertise exists across all disciplines and approaches to support researcher RDM needs—particularly in specialized areas like data curation, active data management, and the application of specific software to manage data.

Financial Support

TABLE 5: SUMMARIZED AGGREGATED MAMIC ASSESSMENT FOR ELEMENTS WITHIN THE FINANCIAL SUPPORT CATEGORY

Element	Maturity	Scale
Resource allocation for HR capacity: Includes RDM service design, that is, costs for things such as employees to provide the services, professional development costs to expand expertise in RDM, etc.	Low	Moderate
Resource allocation for infrastructure: Includes service development, cost models, and investments in infrastructure.	Low	Low

Summary: Generally, central RDM services are either cost-recovery (and therefore unavailable to all researchers) or not significant enough to support needs RDM needs across campus. Some researchers pay for their own infrastructure because McMaster cannot meet needs; it is unclear how these needs will be addressed considering changing national Digital Research Infrastructure and service provision.

Results - Survey

Section 1: Demographics of the survey respondents

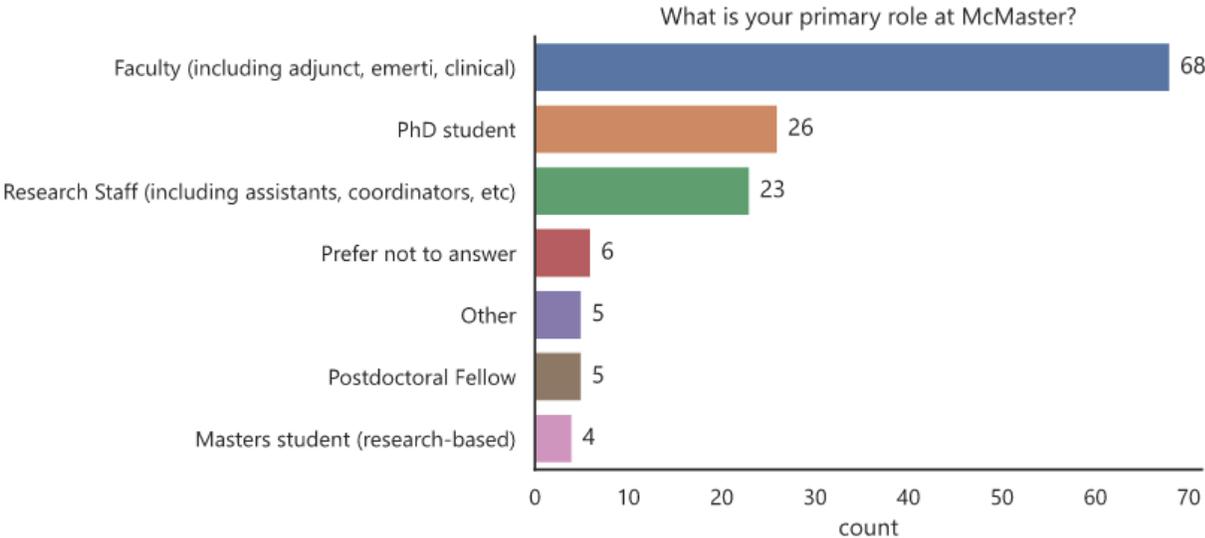


FIGURE 2: PRIMARY ROLE OF RESPONDENTS WITHIN THE UNIVERSITY

Most of the survey respondents were faculty, followed by PhD students, and then research staff (figure 2). The faculty of Health Sciences was the top Faculty of origin for our respondents, followed by the DeGroot School of Business, and the Faculty of Engineering (figure 3).

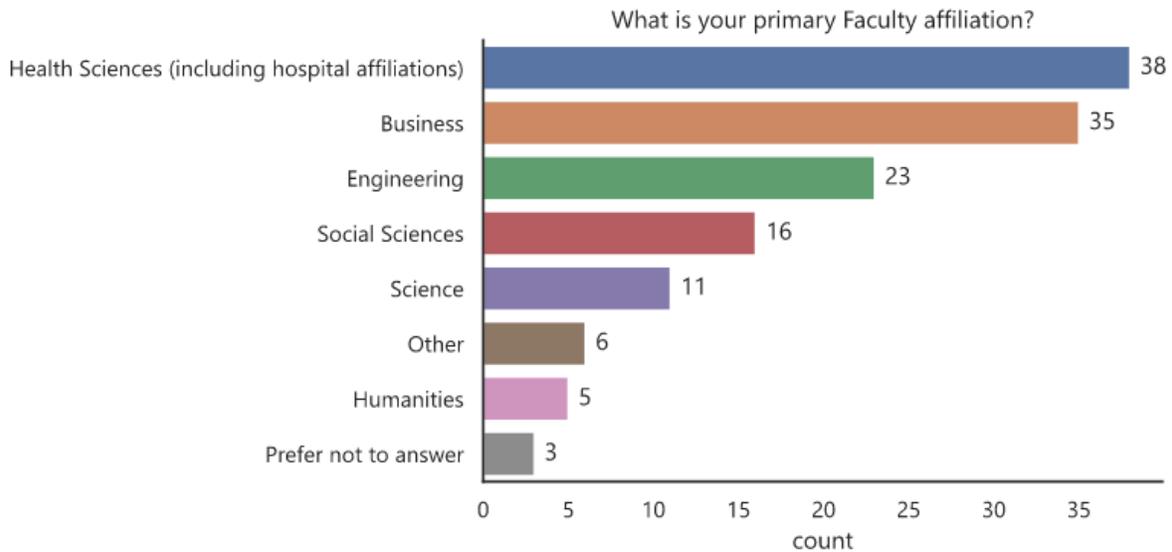


FIGURE 3: PRIMARY FACULTY AFFILIATION OF RESPONDENTS

Section 2: Data management in research grants

Over a third of respondents had currently or previously held a grant with data management requirements (figure 4), indicating that although the requirements from the Tri-Agency RDM policy have not yet come into effect, McMaster researchers are already managing RDM-related requirements related to funding sources.

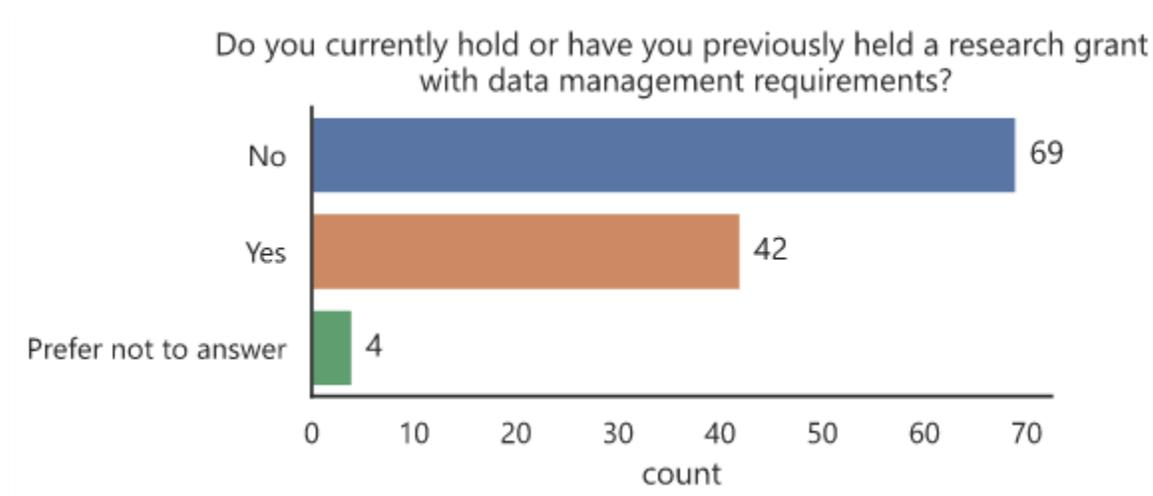


FIGURE 4: RESEARCH DATA MANAGEMENT IN GRANTS

Of the requirements, the most common included a requirement to create a data management plan (DMP), a data sharing plan (DSP), or special security requirements (figure 5).

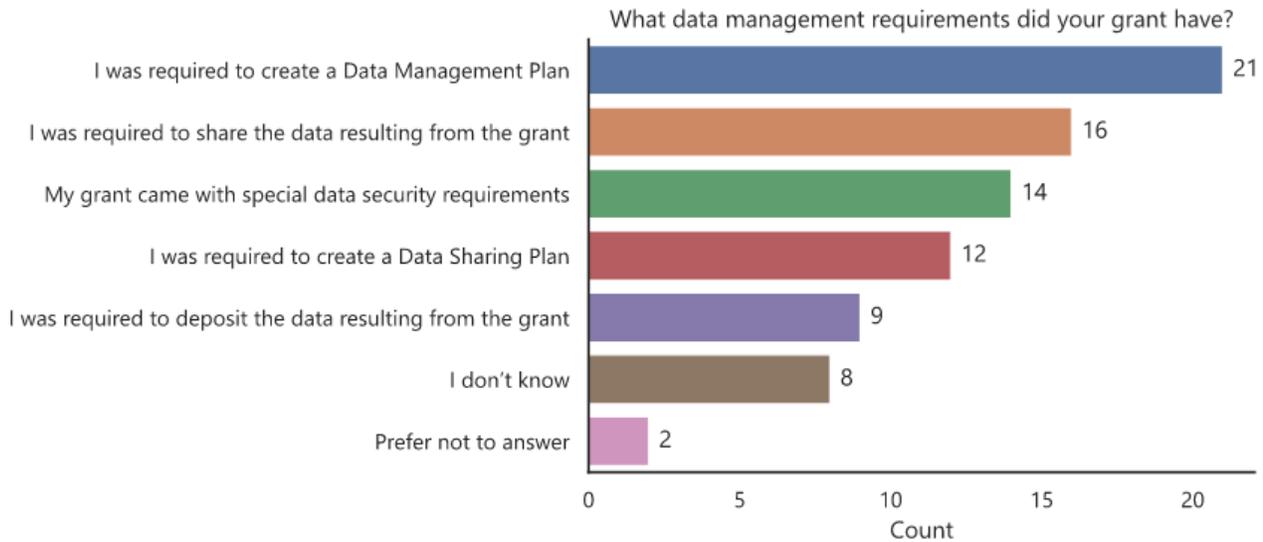


FIGURE 5: GRANT DATA MANAGEMENT REQUIREMENTS

For those who were required to create DMPs, the majority did not use existing tools and templates that are freely available online (figure 6).

Section 3: Data storage and backup

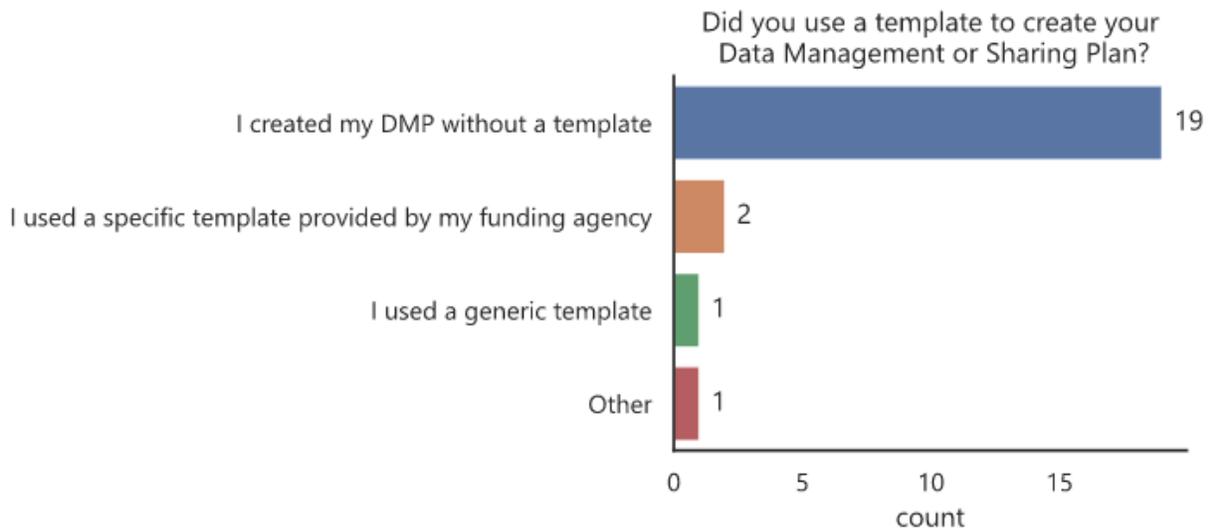


FIGURE 6: DATA MANAGEMENT PLAN CREATION

The most common storage location for research data was personally managed internal hard drives (figure 7). While internal hard drives are easy for researchers to use, they are typically the least reliable option, exposing researchers to higher risks of data loss unless carefully backed up. The next most common options were McMaster provided cloud storage systems and McMaster managed computers or servers, better options for security since professional IT staff can ensure computers and servers are properly updated and secured.

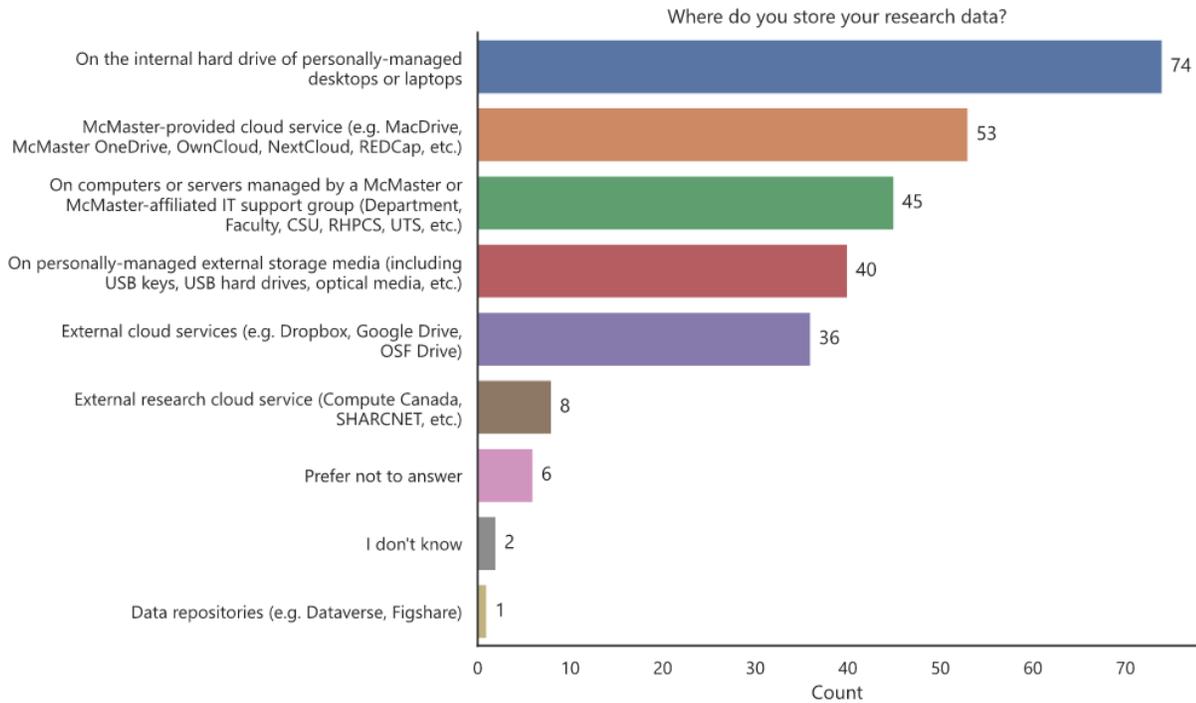


FIGURE 7: DATA STORAGE LOCATION

The amount of storage researchers need varies a lot, with the majority of researchers having minimal storage needs under 100 GB that are easily met. The current default allocation for the institutional OneDrive is 1 TB, and 300 GB on MacDrive. These amounts should satisfy the majority of researchers on campus (see figure 8). However, there is a long tail of researchers with higher needs, over 4 TB and potentially in the hundreds of TB, who will require more support.

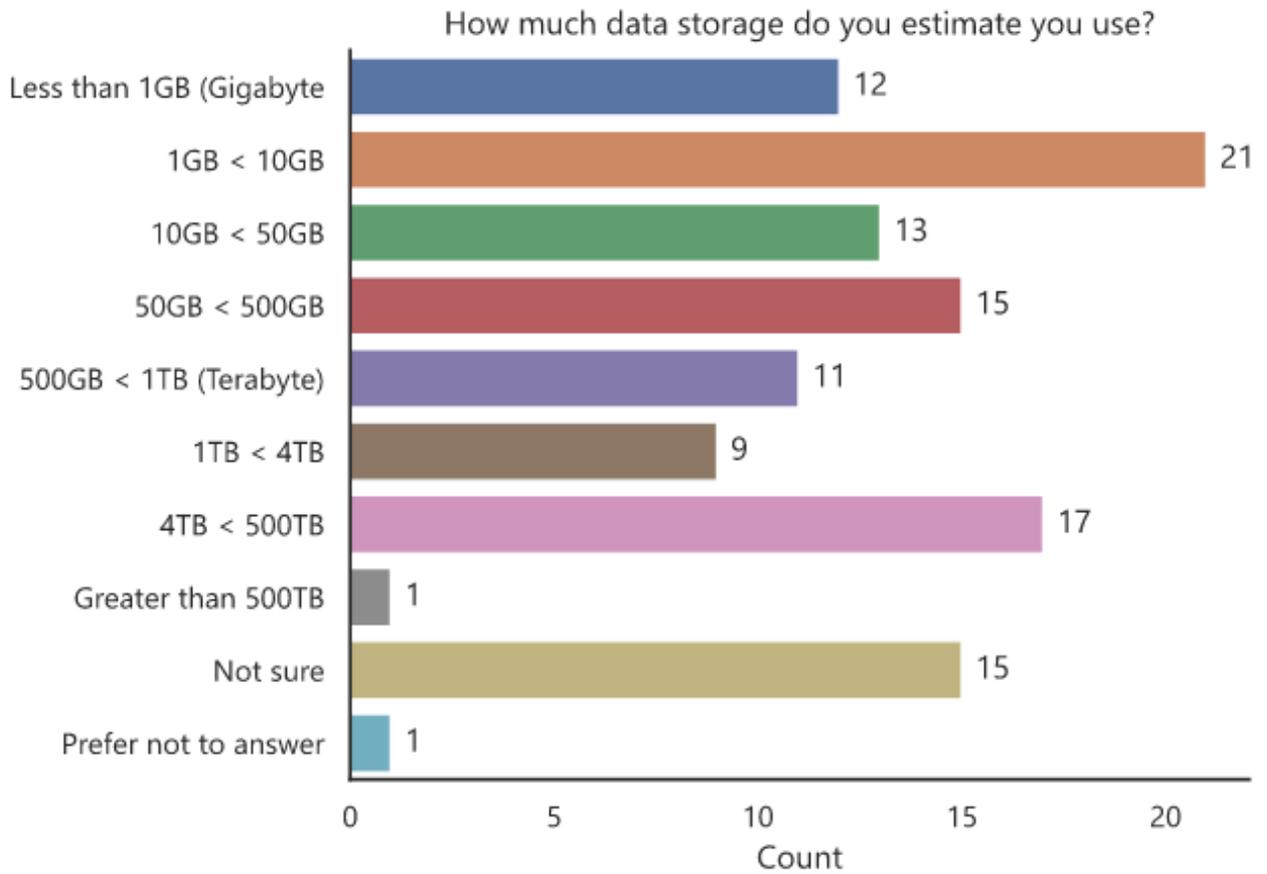


FIGURE 8: DATA STORAGE VOLUME NEEDS

By multiplying the average of a respondents' current data storage needs bucket with their expected growth, we estimate that the current storage requirement for only the respondents to this survey is around 5 PB and expect this to grow to around 28 PB in 5 years, solely from survey respondents. We should expect a large increase in storage requirements across campus. There is a clear cluster of high data storage need users (see figure 9), but most researchers expect to stay within what is provided by McMaster's OneDrive platform.

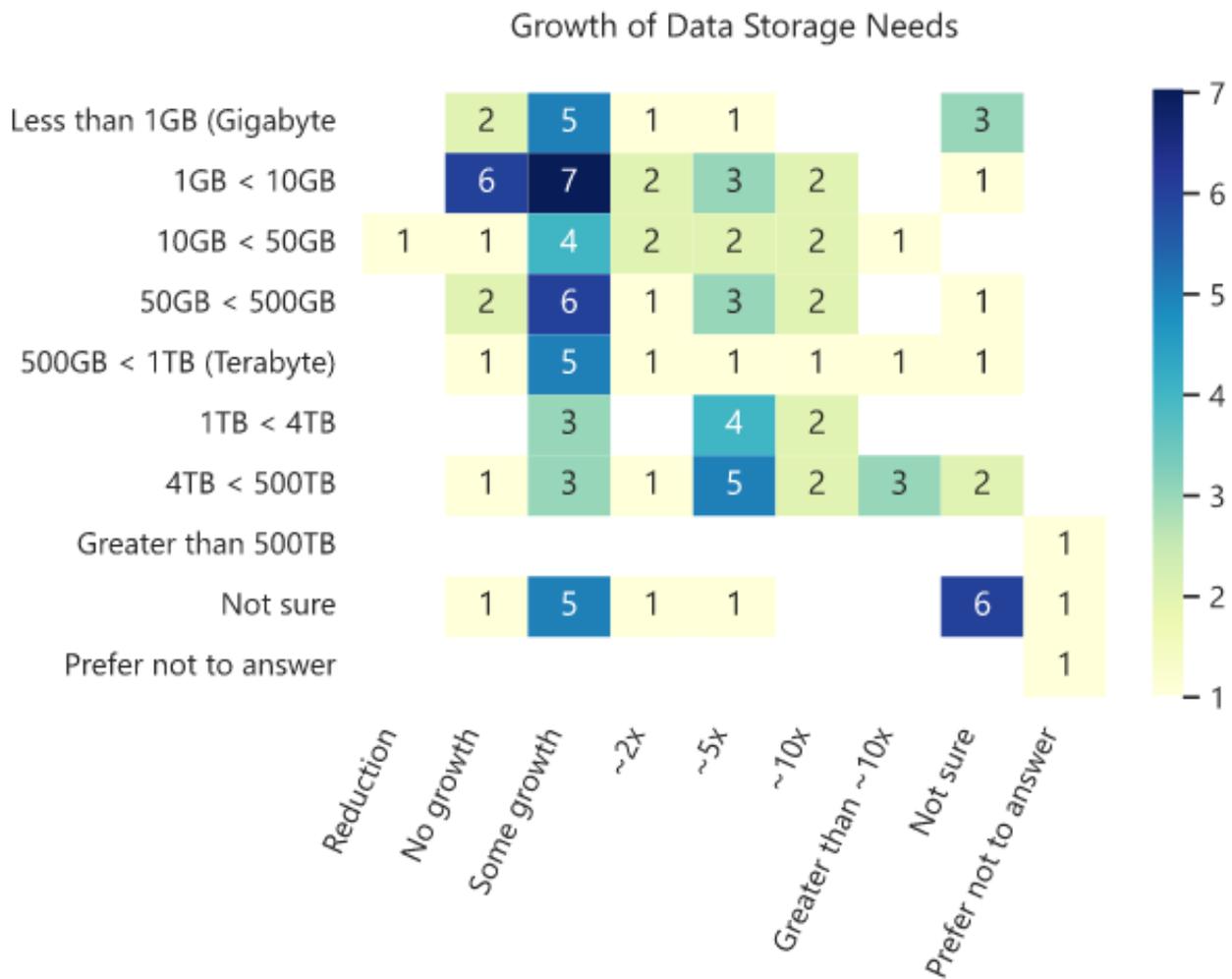


FIGURE 9: GROWTH OF DATA STORAGE VOLUME NEEDS, GROUPED BY CURRENT STORAGE REQUIREMENTS (AS PER FIGURE 8)

A key requirement for safely storing data is keeping backups. If data is stored on institutional servers or cloud platforms, backups are often performed automatically on a regular basis by the storage provider. If researchers are storing data on a computer or mobile device, they may make use of backup functions built into their operating system like Time Machine on MacOS or Windows Backup or using 3rd party software. 68% of respondents reported that backup copies of their data exist (see figure 10).

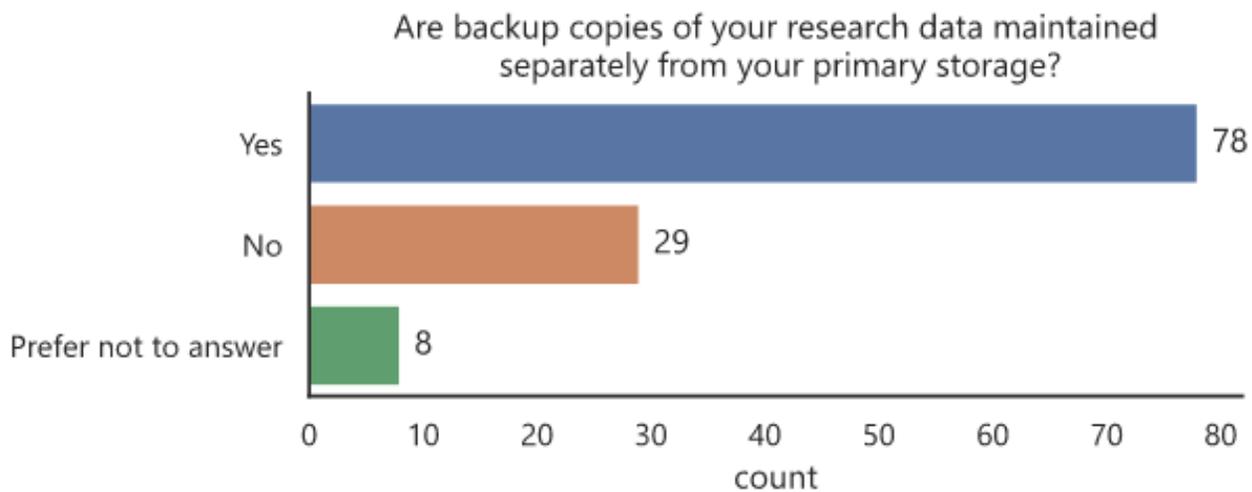


FIGURE 10: DO RESEARCHERS BACKUP THEIR DATA?

Of those backups, the most common option reported was using personally managed external media, including hard drives and USB drives. While this type of backup offers some protection, it has numerous limitations of which the most significant is that backups must be done by the individual regularly or they go out of date. As with data storage, the second and third most common options were McMaster provided or affiliated servers and cloud services. McMaster provided servers and cloud services is the ideal solution, where backups can be automated and kept separately from the data collection and storage location.

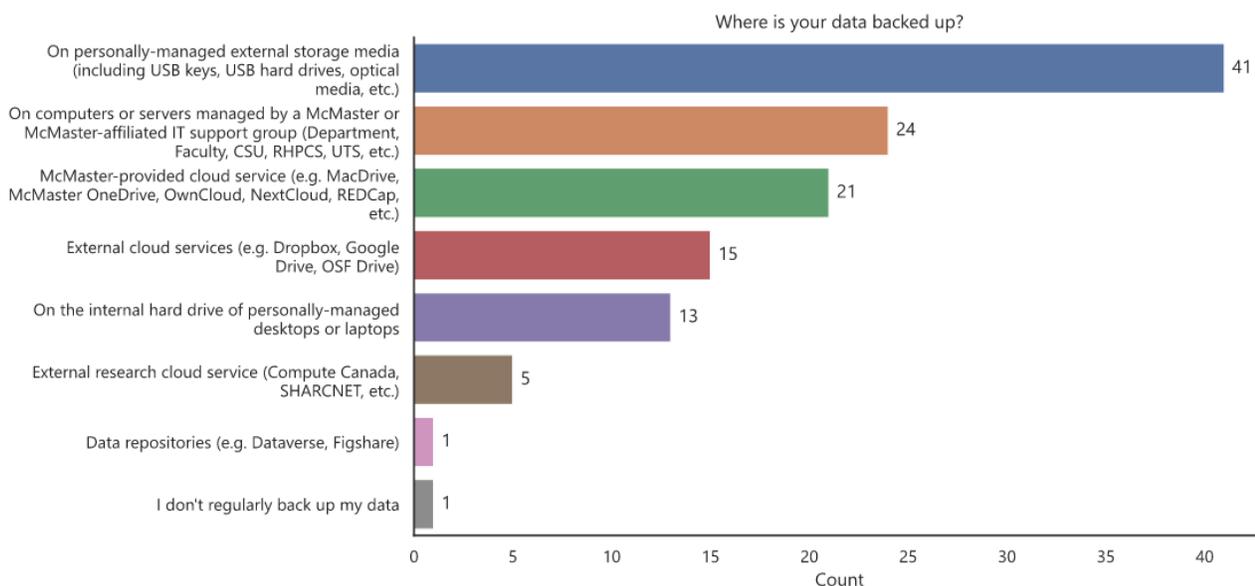


FIGURE 11: WHERE IS RESEARCH DATA IS BACKED UP

Section 4: RDM roles and responsibilities

There are several interesting divisions in who performs which data management tasks within a research group. The roles that perform the most data management tasks are faculty, graduate students and postdoctoral fellows, and research staff, but the tasks are not equally divided. Faculty are more likely to do the more “strategic” tasks like data management planning, data management procedure and policy development, data and system security management, and data archival and sharing. Graduate students and postdocs are more commonly responsible for proximal data management tasks like data collection, cleaning and processing, analysis and visualization, and organisation and documentation. Research staff (staff working directly on research including research coordinators, research assistants, etc.) seem to be involved in all tasks relatively equally. We also see undergraduate students playing a role in data collection, cleaning, and analysis. Research support staff (staff

working to support research efforts including grant officers, RDM Specialists, IT personnel) are most commonly involved in data security, and data storage.

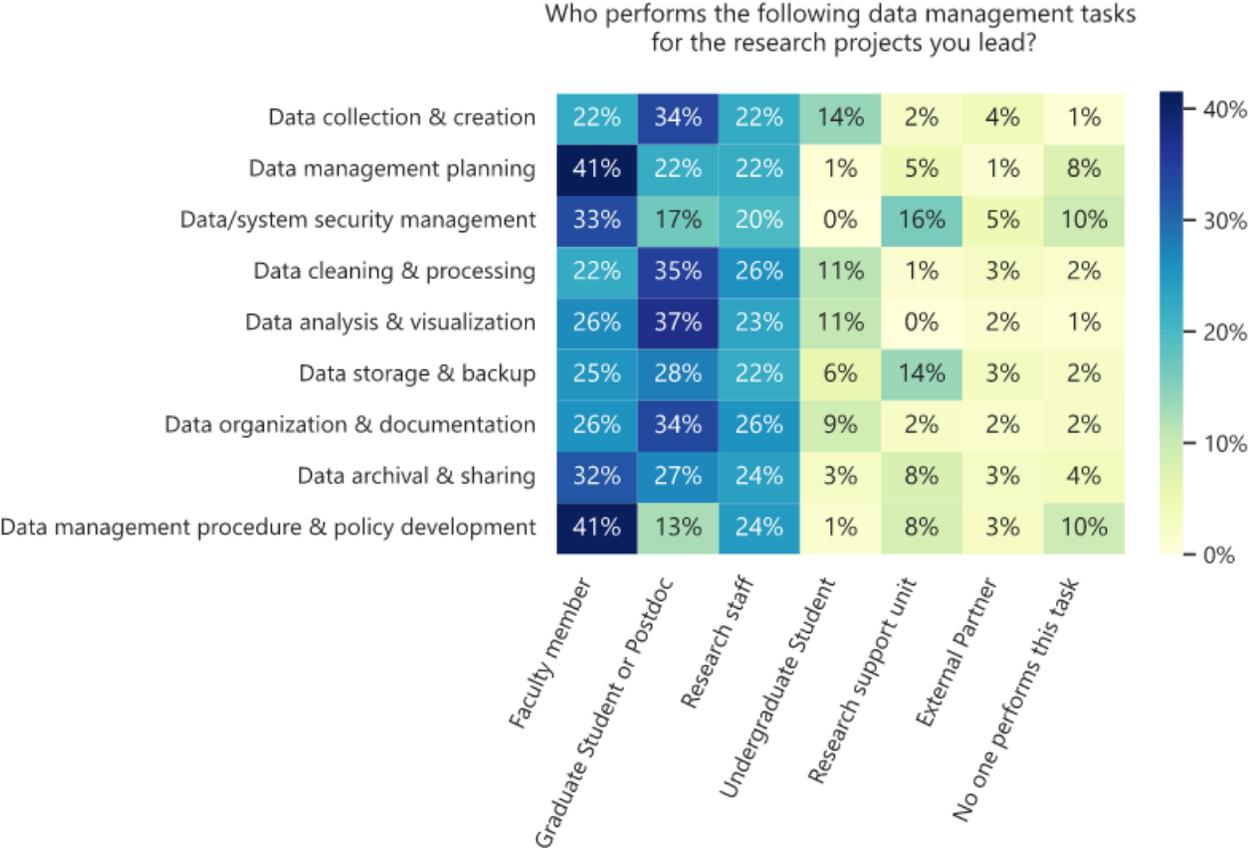


FIGURE 12: WHO PERFORMS THE FOLLOWING DATA MANAGEMENT TASKS FOR THE RESEARCH PROJECTS YOU LEAD?

Section 5: Data sharing

Data sharing can be challenging for researchers to determine whether and how they want to share data. By far the most common way our respondents choose to share their data is through personal requests (see figure 13). This may be the easiest option for researchers since it allows them to do data curation informally and in a way that fits into their existing workflows. It also allows researchers to maintain the highest level of control over who is using their data. However, it is becoming less accepted in the academic literature, and more journals and funders are requiring data to be shared openly with publications in trusted data repositories. Direct sharing by request is also the most likely to be subject to inappropriate practices, such as privacy violations (through sharing itself or the mechanism of sharing) or storage technologies (data breach is more likely). The second most common option our researchers use to share data is as attached to journal publications. This may reflect the increase in requirements mentioned above. Many researchers still do not share their data at all—this was the third most common choice.

We received several comments on data sharing, with several researchers saying that they are unsure about how to proceed with data sharing, or that they had not yet shared data but likely would in the future. Many researchers shared that they worked with sensitive data (such as health data from human participants) and that the data were unsuitable for sharing or they did not know how to share the data while preserving participant privacy and consent. Data sharing is a growing requirement and an area where we expect to see practices and related requirements change rapidly.

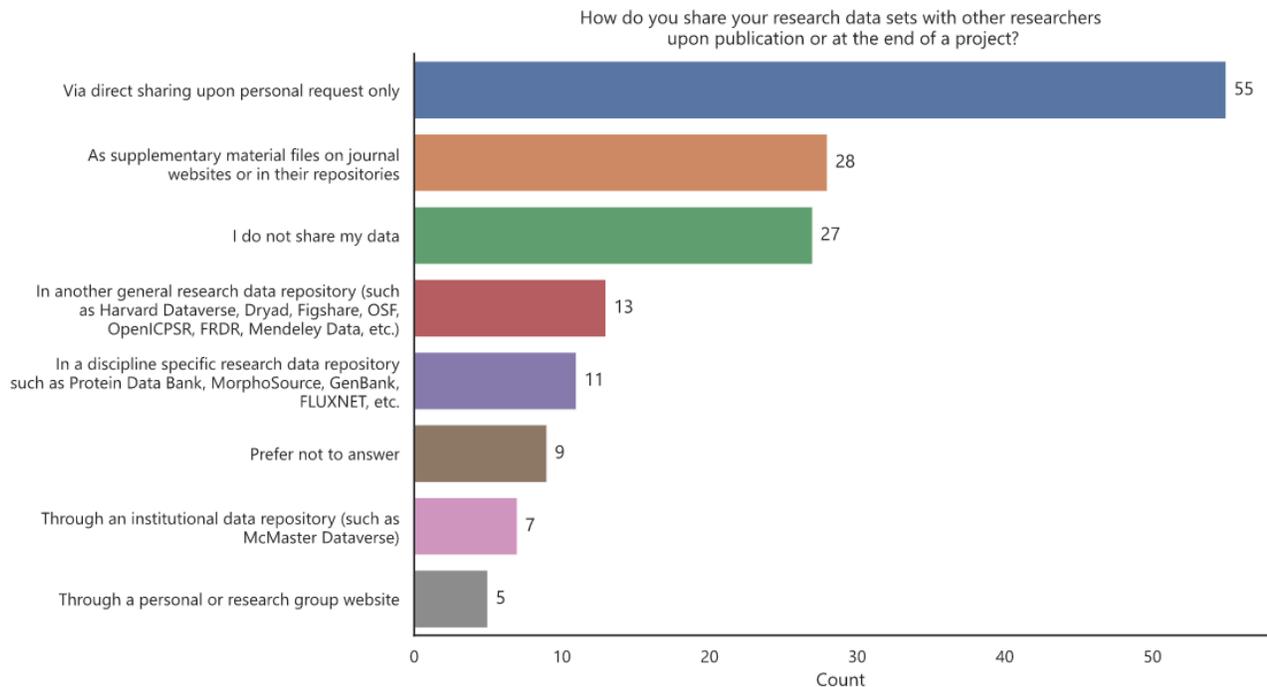


FIGURE 13: HOW DO YOU SHARE YOUR RESEARCH DATA SETS WITH OTHER RESEARCHERS UPON PUBLICATION OR AT THE END OF A PROJECT?

Section 6: RDM support services at McMaster

Building awareness of RDM support is a key area of work for us, with 59% of respondents not aware that services existed and only 13% having used RDM supports before (see figure 14).

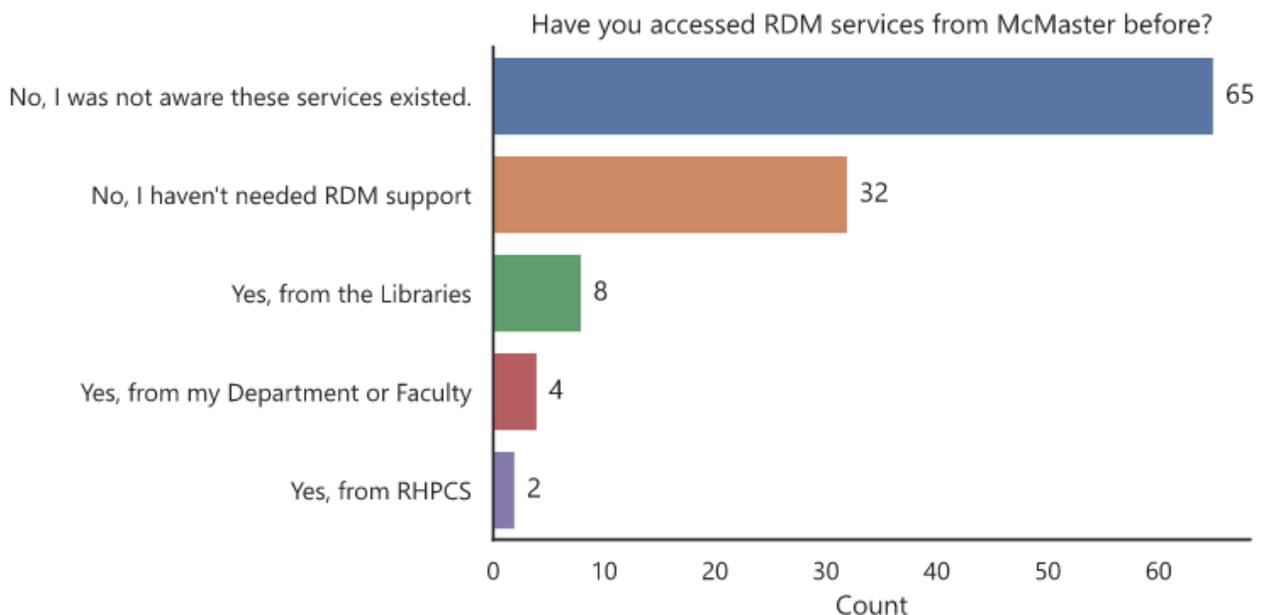


FIGURE 14: HAVE YOU ACCESSED RDM SERVICES FROM McMASTER BEFORE?

We gauged interest in a range of RDM support services and found consistent broad interest from researchers (see figure 15). Averaging across all categories, 69% of respondents were at least slightly interested in RDM support

services, and 44% of respondents were at least moderately interested. 75% of respondents are very or extremely interested in accessing at least one of the RDM services listed.

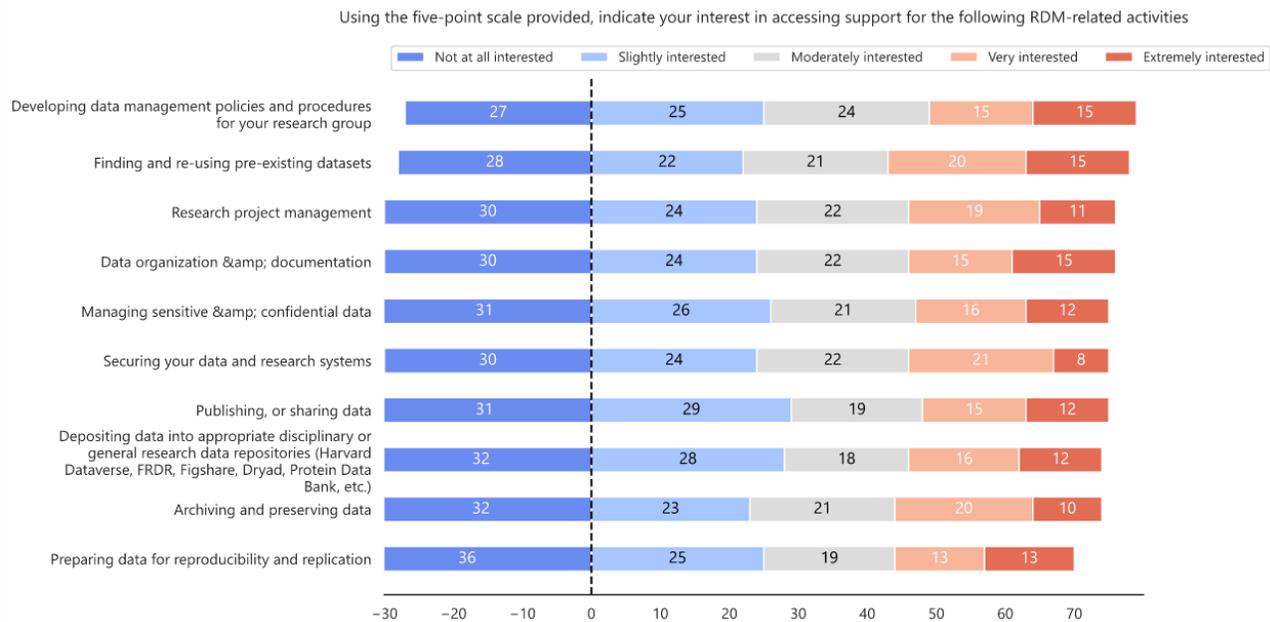


FIGURE 15: INDICATE YOUR INTEREST IN ACCESSING SUPPORT FOR THE FOLLOWING RDM-RELATED ACTIVITIES

Similarly, there was broad interest in accessing support in different modes (see figure 16), with an average of 69% of respondents at least slightly interested in all service modes and 44% at least moderately interested. 72% of respondents are very or extremely interested in accessing RDM services through at least one of the modes.

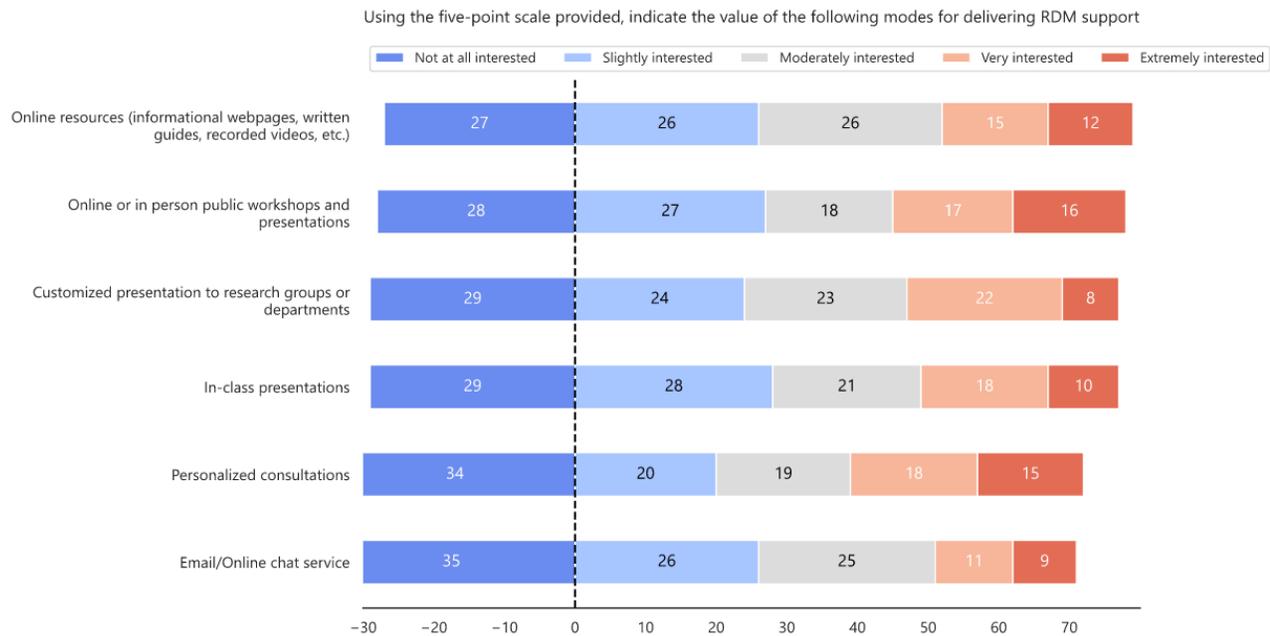


FIGURE 16: INDICATE THE VALUE OF THE FOLLOWING MODES FOR DELIVERING RDM SUPPORT

Section 7: Challenges and barriers in managing data

Researchers face many challenges in managing their data, and these challenges often compound. The number one challenge researchers reported was a lack of time. From the comments we received, this lack of time is often what leads to the lack of knowledge.

Researchers also wrote to us about their challenges meeting the changing and expanding RDM requirements. Expectations are increasing and researchers are worried about finding the time to learn and train their research teams about what is required of them. Some indicate they are overwhelmed with what they are expected to learn and implement.

The ever-increasing digitization of research presents new and growing challenges for researchers working with sensitive data. Increasing cyber threats lead to researchers feeling afraid they aren't doing enough to protect their data because they don't know what they need to do.

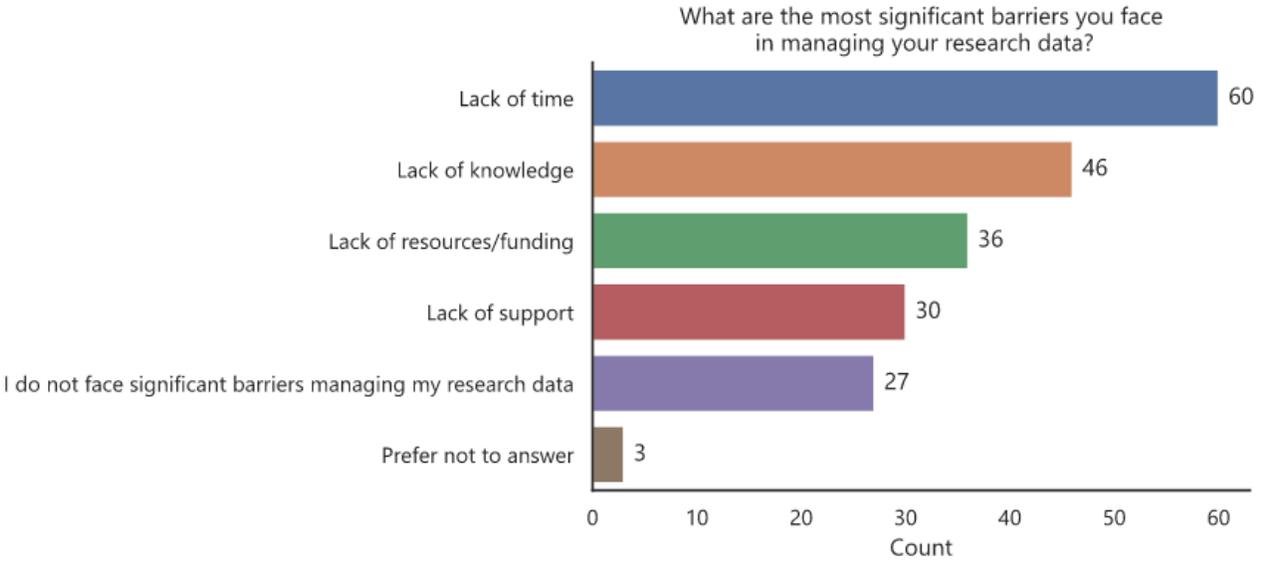


FIGURE 17: WHAT ARE THE MOST SIGNIFICANT BARRIERS YOU FACE IN MANAGING YOUR RESEARCH DATA?

Results – Focus Groups

The focus groups focused on exploring where researchers are currently struggling with managing their data, how we can help meet their needs, and how we can reach them. The focus groups were centred around four main questions:

1. What are the biggest challenges you face in managing your data?
2. What services and training options would be the most useful to help you meet those challenges?
3. Where do you go to find information and guidance for managing your data?
4. What would be the best way for us to advertise RDM services, workshops, and resources to you?

We took detailed notes during each focus group, and transcripts were obtained using the Zoom live transcription feature to ensure we captured the information from each focus group accurately. The notes were then used to identify patterns of commonality between focus groups and common themes expressed by the participants.

Theme Overview

Four major themes were identified from the focus groups: Data Storage, Data Security, Data Sharing, and McMaster Services. Each of these themes contains several sub themes. Each theme is an area where researchers are currently struggling, and an opportunity for RDM services to provide support.



FIGURE 18: FOCUS GROUP THEMES

Data Storage

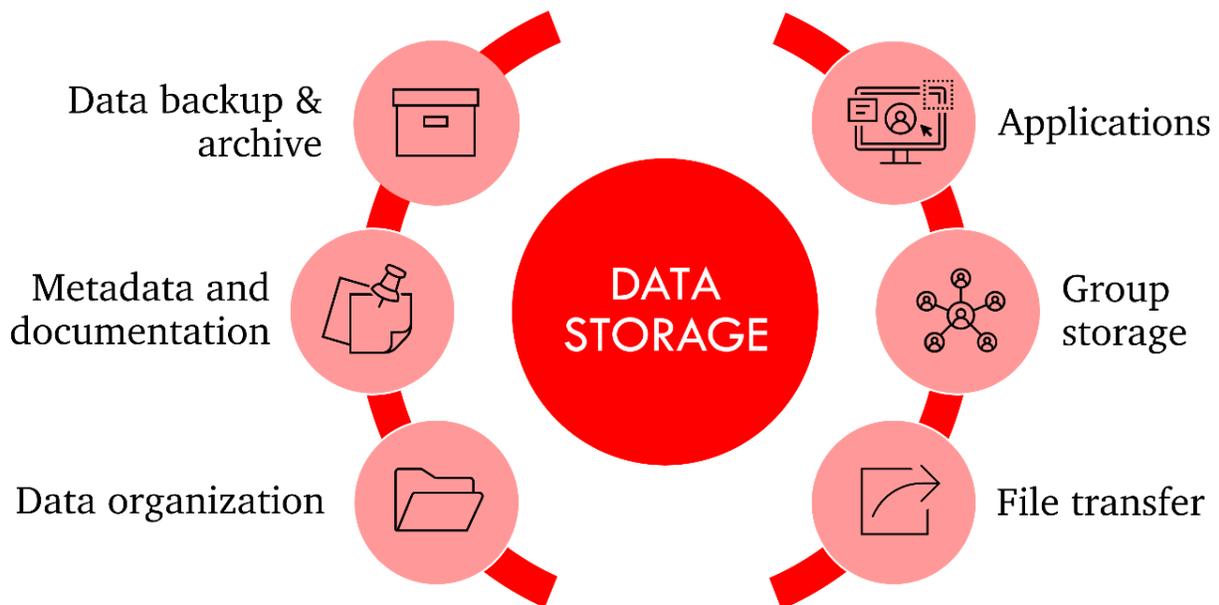


FIGURE 19: DATA STORAGE THEMES

Overall, the sentiment from researchers is that McMaster does a good job of providing basic file storage. File servers are available through Faculty IT units and RHPCS, and two cloud storage options are provided by UTS:

MacDrive, and OneDrive/SharePoint/Teams. However, there are several specialized needs that are not fully being met.

Data Archival

One of the challenges of a research lab is keeping track of and organizing data. As a researcher's career progresses, they naturally accumulate more and more data, and it is desirable to keep that data so that it may be integrated into future studies. This data may not be actively worked on, so it makes sense to keep it in a separate archival system. Several researchers expressed a desire for the University to provide a dedicated archival storage system like a magnetic tape storage system. Although cloud archive platforms exist including Amazon S3 Glacier, not all research data can be archived online in commercial systems.

Group Storage

A major challenge that many principal investigators (PI) expressed to us was managing the data for their research group. The members of a research group are always in flux, as students graduate and staff are rarely on permanent contracts. If the group doesn't have established policies, training, and documentation on data management practices, each member of the group is often left to figure out how to manage their data independently. If each member of the group is storing data in their own system, it's a major challenge for the PI to ensure that all the data is backed up and retained after members leave the group. Many PIs expressed concern that they are not always able to manage this and are looking for a way to organize their lab and store data as a group. Microsoft Teams is often a good option for these groups. Teams may not work for groups working with sensitive data and migrating data may be a challenge.

File Transfer

Data transfer, especially to external partners and other institutions, can be difficult as many partner companies and government agencies have strict data requirements. Currently McMaster provides some file transfer abilities through the sharing functions in OneDrive and MacDrive, and Faculty of Health Sciences IT provides access to MacDrop for FHS researchers. When these options fail, researchers may be left at an impasse, unable to transfer data to collaborators in Hospitals or private companies.

Applications

Data storage doesn't always mean file storage. Many researchers rely on applications to collect and store research data. These range from simple desktop apps that researchers can run themselves to complex server-client software that often requires IT support. This includes SQL servers and other relational databases, mobile app development, data collection and management software like REDCap and LimeSurvey, and custom software for specialized research needs. Researchers expressed support for their faculty IT departments but were often looking for more flexibility with infrastructure and software. Every system or platform has its limitations and may not fit modern complex research studies. Proprietary software changes over time and may not change in ways aligned with what researchers need.

Data Security



FIGURE 20: DATA SECURITY THEMES

Data security is a growing concern for academic institutions (Fouad, 2021) and researchers, and McMaster has not escaped notice (Bobby Hristova, 2021). Researchers are not and cannot be expected to become IT Security professionals, so they need support from the University and qualified IT Security specialists. Researchers do not always understand the risks and why they make good targets for cyber attacks. Security measures aren't just to protect research data from theft and research participants from privacy breaches but to protect researchers and the institution from ransomware attacks, phishing attempts, and other scams. There is always a trade-off between security and accessibility and convincing researchers to make their systems more restrictive (for example by implementing MFA) will always require strong rationale and evidence. Researchers need to be brought to understand why security measures are important or they will not implement them.

Sensitive Data

When working with sensitive data, especially data from human research participants or patients, maintaining data security becomes even more important. Researchers working with this data must comply with privacy legislation including Ontario's Personal Health Information Protection Act (PHIPA) for health-related data. Researchers working on international projects must comply with legislation from each local authority.

Data Ethics & Privacy

The complexity of the compliance requirements means that researchers don't always feel comfortable managing this kind of data, even with guidance from the Privacy Office or the Ethics Boards. They may not be sure when they need a Privacy Impact Assessment (PIA) or Risk Assessment of software, and software for data and project management that complies with requirements is often more limited and expensive. Researchers also worry that the new hybrid working environment makes it more difficult to secure data. Staff and trainees no longer work with their data exclusively on campus, and increasingly use remote computers to logon to campus networks or data platforms, increasing the risk vectors for cyber-attacks.

Encryption & Cyber Hygiene

Researchers are not IT experts, and they should not be expected to be. While there are researchers who are technically adept and capable, many of them find understanding and implementing encryption and other security measures confusing and difficult. While the rise in cyber-attacks gives rise to a need for increased security, we need to make sure the recommendations we give to researchers for securing their data and systems are clear, understandable, and justified.

Data Sharing

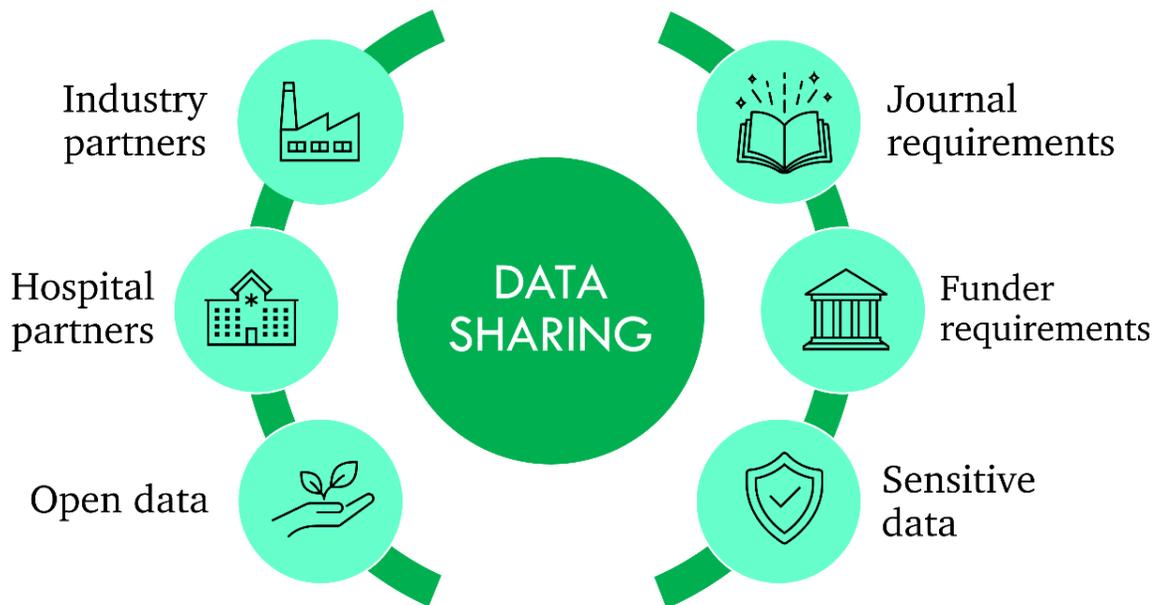


FIGURE 21: DATA SHARING THEMES

Data sharing is increasingly expected after the completion of a research project. Increasing demands from journals and funders for data sharing have put new and complex challenges on researchers' shoulders. Journal expectations have changed rapidly over the last decade and have left some researchers scrambling to keep up. While most of the researchers we talked to support open data and data sharing in principle, they have concerns about sharing their own research data. These concerns can be broadly categorized as: concerns about sharing data that they may want to use themselves in the future, concerns about protecting sensitive data—especially personal information or personal health information from research participants, concerns about the quality of their data and metadata, concerns about who will use their data and for what purposes, and concerns about the time required to meet data sharing requests and requirements. Some of the researchers we talked to now select journals in part based on their data sharing requirements.

Expectations about data sharing vary between disciplines, with some researchers reporting that data sharing is simply part of how research is done now, and others commenting that conversations about data sharing are starting to emerge in the discipline, but practices are not standard (yet).

The biggest concern that we heard is about sharing sensitive data, and in particular data from health sciences, social sciences, and biomedical research with human participants. Researchers working with human participants must go through ethics review and follow the TCPS 2 Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. They must also abide by privacy legislation in the jurisdiction of the research. Data sharing requirements from journals often conflict with REB requirements to keep data secure. Researchers have a few options for data sharing here, including de-identification/anonymization, sharing through a secure platform, or direct sharing through a Data Transfer Agreement or Data Sharing Agreement. However, researchers are often not aware of their options or not comfortable performing the necessary steps to get there (such as de-identification). In addition, sharing data from human participants requires getting consent from the participants. This step may not have occurred in older studies, meaning that researchers must engage their Research Ethics Board to seek approval of an exception. Depending on existing consent, this could mean re-contacting participants to seek consent retrospectively, or justifying the need to share anonymized non-sensitive data from a group known to support data sharing, before depositing data. (I. A. P. on R. E. Government of Canada, 2021)

McMaster RDM Services

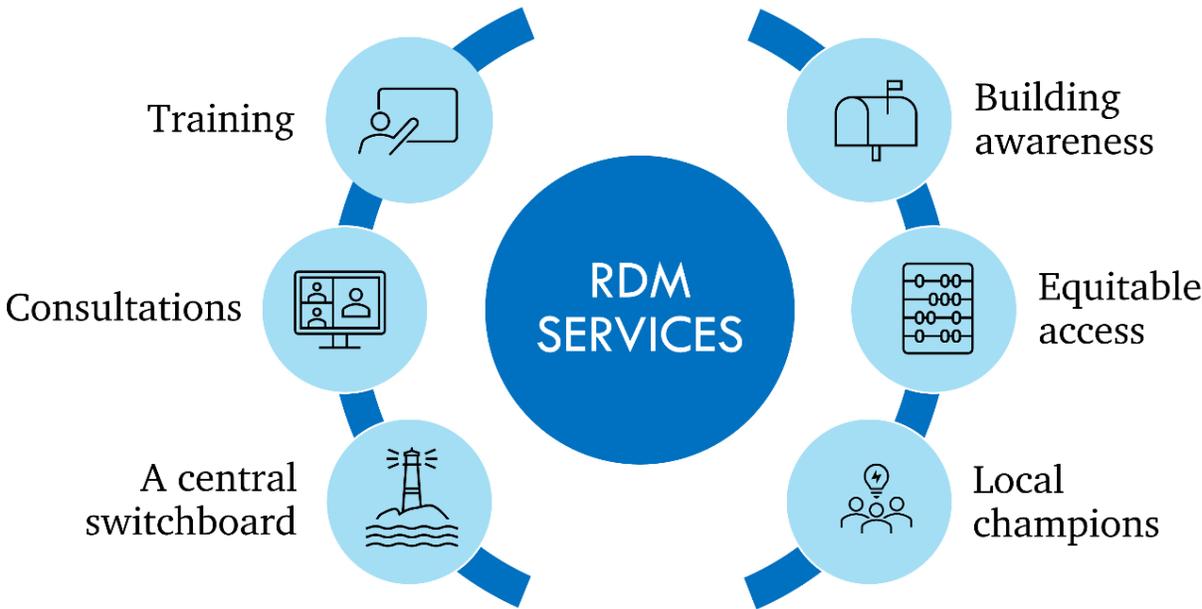


FIGURE 22: McMASTER RDM SERVICES

RDM services are really a constellation of related services like data storage and backup; data archival and sharing systems and platforms; research support for grant applications, data sharing/access/transfer agreements, and journal publications; secure platforms and applications for sensitive data, research software, and support, training, and consultations for all the above. As such, an important step to improving RDM services is collaborating and building connections and partnerships between the different campus units that provide support.

Researchers in the focus groups brought up several important concerns around the future of RDM services on campus.

Awareness

RDM services as an independent service at McMaster is fairly new and really got off the ground in January 2021. As such, building awareness among researchers has been an important task for us. However, we still have a long way to go in spreading awareness of the service into all the many departments and research units on campus. Many of the researchers we talked with were not aware of the supports and training they could access through the RDM service.

Equitable Access

Researchers made it clear that they expected RDM services to be centrally provisioned and funded, and that a fee-for-service model provides barriers for researchers without or with less funding, including graduate students, postdoctoral fellows, research staff, and researchers in less well funded disciplines such as humanities and social sciences. RDM services should be provided to all research groups on campus and not limited by ability to pay.

Integrating Local and Central Services

Researchers rely on their personal networks to gather information, find help, and access supports. They rely on staff in their departments and faculties to provide research support and information. However, that information needs to be consistent between departments and faculties. In addition, researchers pointed out that information is often siloed. For example, if a researcher in one unit licenses a particular program, it may be possible for them to share that access with others. Being able to gain access then relies on other researchers being aware of the existence of that license, which often doesn't happen. It's clear that there are advantages to providing centralized services, such as consistent support for all units and less duplication of effort. However, we need to find a way to

tap into existing local expertise and build expertise and awareness. Many researchers suggested some kind of ‘data champions’ program, where researchers could self identify as a data champion that others could contact for help. Programs like this have been run at the University of Cambridge, Curtin University, and TU Delft.

Training

The need for training in data management was strongly expressed, especially for early career researchers, graduate students, and postdoctoral fellows. Researchers suggested we take advantage of existing training in related areas such as research methods and data analysis courses. They also were strongly in favour of some form of asynchronous basic training course that was mandatory for all graduate students on campus and could be integrated into Avenue to Learn.

Consultations

Research consultations with RDM services are increasingly popular as awareness grows across campus and will likely continue to increase with new and increased data requirements from funding agencies and journals. There is interest from researchers in consultations, especially focused on data security and sensitive data. When a research project has more complex data management requirements there is more of a need for support from staff.

Conclusions and Next Steps

Like best Research Data Management practices, developing an Institutional RDM Strategy is not a “box-checking” requirement to be completed and forgotten. Rather, both are important ongoing processes championed by those responsible in the pursuit of quality research. The forthcoming [Tri-Agency Institutional Strategy Requirement](#) is a timely prompt to affirm our plan to support researchers through a moment of expansion in RDM needs, driven both by developments in internal research processes and external requirements.

While the McMaster Institutional RDM Strategy Working Group may have an intuitive sense of the current state of RDM at our institution, it is essential that we develop our strategy using the same robust data-driven process as the research it will support.

Through the [MAMIC Assessment](#), we now understand better the Governance and Policy frameworks that need to be in place to successfully integrate RDM in McMaster’s research enterprise. These facets grow out of RDM best practices outlined by DRAC. This exercise outlined gaps and needs at the university for policies and procedures, IT infrastructure, service and support, and financial support.

Our [McMaster RDM Needs Survey](#) revealed areas for growth in department engagement. This includes a clear need to support researchers with data management planning, data storage and backup, and data sharing. We confirmed that we must include all involved in the RDM process: faculty, graduate students and postdoctoral fellows, and research staff. We also learned that we need to expand promotion of RDM support services and develop RDM strategies that respond to a lack of time, knowledge, and resources.

Our [RDM Focus Groups](#) added essential complexity to our understanding of researcher needs for data storage, data security, data sharing, and RDM services. This highlighted supporting researchers in sharing sensitive data, the need for equitable access for smaller research groups and graduate students, and the need for local champions. On a meta-level, we had 105 completed surveys and focus groups with 91 participants. There is a desire from research teams to engage actively and share challenges and needs.

We come away from this research with a better understanding of the current state of RDM at McMaster and the potential institutional frameworks and RDM training and services to support best practices:

Frameworks

- Governance and Policy
- Funding and Support Models
- Culture, Community, and Collaboration
- RDM Services and Training

RDM Practices, Tools, and Infrastructure

- Data Management Planning
- Data Storage and Retention
- Data Ethics and Sensitive Data
- Data Sharing and Access
- Data Security
- Indigenous Data Management
- Software and Applications
- Data Curation

Now with a detailed overview of the current state, we will work with university stakeholders to envision RDM futures through our Stage 3: The **Ideal State** of RDM at McMaster document. Our final **McMaster Institutional RDM Strategy** document will draw lines between this current state and our desired state, noticing gaps, finding parallels, imagining support systems.

Although this document is prepared by a smaller group, the responsibility of caring for research data is shared by all of us who intersect with research at McMaster. Our work will not end with our first draft. We are planning multiple rounds of review, engagement, revision, focus groups, faculty town halls as we move forward in this process. Through this, we will ensure our strategy reflects our research community.

Just as we are planning for adjustments in response to feedback, we are also planning systems to ensure our Institutional Strategy can adapt to changes in McMaster's research environment. Research systems must be responsive to inputs, workflows, shifting priorities, and new information. Each of us is accountable for data produced at McMaster. This document is one part of an ongoing process—a living system of research data management—which we will keep alive, together.

References

- Bobby Hristova. (2021, May 18). Hamilton police tech crime unit investigates cyber security incident at McMaster University. *CBC News*. <https://www.cbc.ca/news/canada/hamilton/mcmaster-cyber-security-incident-1.6031072>
- Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council. (2018). *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS2)*. Government of Canada. http://publications.gc.ca/collections/collection_2019/irsc-cihr/RR4-2-2019-eng.pdf
- Cooper, A., Perry, C., Szwajcer, A., Wang, M., & Khair, S. (2020). *Institutional Research Data Management Services Capacity Survey: Executive Summary*. <https://doi.org/10.14288/1.0388722>
- Dennie, D., & Guindon, A. (2017). *Report on the 2016 survey of research data management at Concordia* [Monograph]. Concordia University. <https://doi.org/10.11573/spectrum.library.concordia.ca.00983908>
- Felipe Pérez-Jvostov, Karey Iron, Shahira Khair, Seppo Sahrakorpi, & Qian Zhang. (2021). *Researcher Needs Assessment: Summary of What We Heard*. Digital Research Alliance of Canada. https://alliancecan.ca/sites/default/files/2022-03/needsassessment_alliance_20220126.pdf

- Fouad, N. S. (2021). Securing higher education against cyberthreats: From an institutional risk to a national policy challenge. *Journal of Cyber Policy*, 6(2), 137–154. <https://doi.org/10.1080/23738871.2021.1973526>
- Fry, J., Dearborn, D., Farrell, A., Khair, S., & Ripp, C. (2021). *RDM Maturity Assessment Model in Canada (MAMIC)*. <https://doi.org/10.5281/zenodo.5745493>
- Government of Canada, I. (n.d.). *Tri-Agency Research Data Management Policy—Science.gc.ca*. Innovation, Science and Economic Development Canada. Retrieved February 17, 2022, from https://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html
- Government of Canada, I. (2016, December 21). *Tri-Agency Statement of Principles on Digital Data Management—Science.gc.ca*. Innovation, Science and Economic Development Canada. http://www.science.gc.ca/eic/site/063.nsf/eng/h_83F7624E.html?OpenDocument
- Government of Canada, I. A. P. on R. E. (2021, May 17). *Guidance on Depositing Existing Data in Public Repositories: Interagency Advisory Panel on Research Ethics*. https://ethics.gc.ca/eng/depositing_depots.html
- Group, I. R. S. T. R. W. (2021). *Institutional Research Data Management Strategy Development Template (3.0)*. <https://doi.org/10.5281/zenodo.5745906>
- Group, P. R. I. E. (2019). *Portage Research Intelligence Expert Group Institutional RDM Strategy Survey—Summary of Results*. Zenodo. <https://doi.org/10.5281/zenodo.3962831>
- Health Sciences Library. (2019, September). *Health Sciences Library Plan 2017-2020*. About Us | Health Sciences Library. <https://hsl.mcmaster.ca/about-us>
- J. Rans & A. Whyte. (n.d.). 'Using RISE, the Research Infrastructure Self-Evaluation Framework' v.1.1. Digital Curation Centre. Retrieved July 5, 2022, from <https://www.dcc.ac.uk/guidance/how-guides>
- Popkin, G. (2019). Data sharing and how it can benefit your scientific career. *Nature*, 569(7756), 445–447. <https://doi.org/10.1038/d41586-019-01506-x>
- Shahira Khair, Rozita Dara, Susan Haigh, Mark Leggott, Ian Milligan, Jeff Moon, Karen Payne, Elodie Portales-Casamar, Ghilaine Roquet, & Lee Wilson. (2020). *The Current State of Research Data Management in Canada: An Update to the LCDRI Data Management Position Paper*. Digital Research Alliance of Canada. https://alliancecan.ca/sites/default/files/2022-04/2020-11_RDM_Current_State_Report_1.pdf

Appendices

Appendix 1: Survey design

The [survey design](#) was deposited in MacSphere, the McMaster Institutional Repository.

Appendix 2: Focus Group script (researcher)

Introductions & consent:

The focus groups were led by an RDM Specialist. Occasionally other members of the RDM services team would attend. At the beginning of the focus group the RDM Specialist would introduce themselves and obtain consent from participants to record the transcript of the focus group. The interviewer used the following script for obtaining consent:

The transcripts will be de-identified, and what you say during our conversation will be kept confidential. Quotes appearing in any report or publication based on this interview will be anonymized unless you give me your express consent to use your name. Nothing identifying you will be shared without your consent.

RDM Presentation:

The RDM Specialist used a short presentation on RDM to define RDM and listed a number of RDM activities to prime researchers and help them connect their research practices with RDM principles.

Questions:

Each focus group used the following questions:

Q1: Could you tell us a bit about your research and what kinds of data you work with?

- **Q1A:** Do you work with sensitive data? This could be patient or human participant data, or data that has other legal, ethical, or contractual obligations.

Q2: What are the biggest challenges you face in managing your data?

- **Q2A:** Do you face any additional challenges when working with sensitive data?

Q3: What services and training options would be the most useful to help you meet those challenges?

Q4: Where do you go to find information and guidance for managing your data?

- This could be a website, an organization, a person, or a unit at McMaster.

Q5: For you, what would be the best way for us to advertise RDM services, workshops, and resources?

If there was time remaining the focus group would go into open discussion.

Appendix 3: Focus Group script (research support staff)

The script for interviews with research support staff was very similar to the script for researchers except the questions would ask about “the researchers you work with” instead of “your research”.

Appendix 4: Environmental Scan - McMaster RDM Stakeholders and Service Providers

RDM connects across all disciplines and multiple services to an expansive and diverse stakeholder group. This is a non-exhaustive list of RDM stakeholders that has been compiled as an initial environmental scan. Please contact rdm@mcmaster.ca to help us add additional groups to this list or correct inaccuracies in the list.

Unit	Sub-Unit / Service	Resources Provided	Leadership and Contacts
University Library	Research Data Management Services (.5 with RHPCS)	RDM website, RDM training, RDM guidance, LibGuide	Vivian Lewis (University Librarian) Dr. Jason Brodeur (Associate Director, Digital Scholarship Services; Administrative Director, Sherman Centre for Digital Scholarship) Angela Di Nello (Operations Director) Dr. Isaac Pratt (Research Data Management Specialist) Subhanya Sivajothy (Data Analysis and Visualization Librarian) Danica Evinger (Research Data Management Specialist)
	McMaster Dataverse	Institutional data repository (federated with Borealis Dataverse)	Dr. Isaac Pratt (Research Data Management Specialist) Danica Evinger (Research Data Management Specialist)
	MacSphere	Institutional repository for scholarly documents (graduate thesis)	Gabriela Mircea (Digital Repository Librarian)
	DASH - Data Analysis Support Hub	Data analysis & coding support, Maps, Data, GIS	Vivek Jadon (Data Specialist) Subhanya Sivajothy (Data Analysis and Visualization Librarian) Christine Homuth (Spatial Information Specialist)
Academic	Data Science Research Lab	Graduate course/academic research on data cleaning and data quality	Dr. Fei Chiang (Faculty Lead, Computing and Software)
	McArthur Lab	Infectious disease databases, computing resources, GitLab, Linux Workhorses, Server Provision	Dr. Andrew McArthur (Lead Researcher) Amos Raphyena (Lead Software Developer) Brian Alcock (Lead Biocurator)
Research	OVPR - Office of the Vice-President Research	Strategic Plan for Research , Associate Deans of Research (ADRs)	Dr. Karen Mossman (VP, Research) Dr. Ranil Sonnadara (Special Advisor to the VP, Research) Associate Deans of Research
	ROADS – Research Office for Administration, Development, and Support	Open access/RDM Webpage	Dr. Sherisse Webb (Director)

	<u>RHPCS – Research and High-Performance Computing Support</u>	RDM technology, Data storage, RDM guidance	Angela Di Nello (Operations Director) Todd Pfaff (Technical Director)
	<u>MILO – McMaster Industry Liaison Office</u>	Data sharing agreements, data security, IP & licensing	Dr. Gay Yuyitung (Executive Director) Amber Metham (Associate Director)
	<u>MREB – McMaster Research Ethics Board</u> <u>HiREB - Hamilton Integrated Research Ethics Board</u> <u>AREB - Animal Research Ethics Board</u>	Ethics, data security, data storage, de-identification	Nick Caric (Assistant Director, Research Ethics, MREB) Erin Bell (Manager, HiREB) Debbie Prince (Ethics Officer, Veterinary Services and Compliance)
Research Centres and Institutes	<u>SCDS - Sherman Centre for Digital Scholarship</u>	Digital Scholarship Services	Dr. Jason Brodeur (Associate Director, Digital Scholarship Services; Administrative Director, Sherman Centre for Digital Scholarship) Dr. Andrea Zeffiro (Assistant Professor, Department of Communication Studies and Media Arts; Academic Director, Sherman Centre for Digital Scholarship)
	<u>MIRI - McMaster Indigenous Research Institute</u>	Training for Indigenous and Non-Indigenous Researchers, connection to OCAP and other best data practices, development of research resources	Dr. Savage Bear (Director & Assistant Professor) Dr. Jennifer Walker (Associate Professor, Department of Health Research Methods, Evidence & Impact)
	<u>MacData</u>	Fellowship program	Dr. Paul McNicholas (Director) Dr. Antoine Deza (Associate Director) Tina Macala (Admin Assistant)
	<u>Spark</u>	RDM guidance, RDM training	Dr. Michelle Dion (Academic Director) Allison Van (Executive Director) Dr. Tara Marshall (Acting Academic Director – July-Dec 2022)
	<u>SEAL - Secure Empirical Analysis Lab</u>	Secure data storage and access	Lily Wang (SEAL Manager)
	<u>CRDCN</u>	Canadian Research Data Centre Network -> National research and training platform for quantitative social sciences. Hosted at McMaster. McMaster RDC is an affiliated site.	Dr. S. Martin Taylor (Executive Director)
	<u>RDC – Statistics Canada Research Data Centre at McMaster</u>	Secure data access to Stats Canada data sources.	Dr. Michael Veall (Academic Director)

UTS - University Technology Services		MacDrive, OneDrive	Gayleen Gray (Assistant Vice President & Chief Technology Officer)
	IT Security	IT Security for Researchers, Information Security and RDM training modules.	Tracy Dallaire (Director, Information Security Services) Greg Atkinson (Senior Manager, Cyber Security Systems & Information Security Officer) Miroslav Cika (Research & IT Security Integration Analyst)
Health Sciences	CSU – Computer Services Unit	REDCap, MacDrop, Departmental storage systems	Tracy Mestdagh (Director, FHS IT) Graig Kelly (Technology Manager, Infrastructure)
	HRS - Health Research Services	Connection to RDM services for Health Sciences researchers	Tracy Arabski (Director)
	IHLL - Indigenous Health Learning Lodge	Training for Indigenous and Non-Indigenous Health Sciences Researchers, connection to OCAP and other best data practices, development of research resources	Dr. Bernice Downey (Associate Dean) Lori Davis Hill (Acting Executive Director) Alexandra Trottier (Executive Director)
	Health Sciences Library	RDM guidance, Research Toolkit LibGuide	Neera Bhatnagar (Head of Systems)
Affiliate	The Research Institute of St. Joe's Hamilton	Academic and research community	Sarah Howe (Executive Director Research)
Affiliate	Hamilton Health Sciences' Research Administration	Five specialized research institutes operating in partnership with McMaster University	Katie Porter (Director of Research Administration)
Affiliate	PHRI -- Population Health Research Institute	Guidance on clinical trial data and PHI storage and management , Institutional Strategy Working Group	Matthew Smart (Project Manager, ICT)
	Departmental Research Staff	RDM guidance, RDM training	

McMaster University Library

Research Data Management – MUL + RHPCS

McMaster's two RDM Specialists are available to researchers to consult on and assist with a variety of research data management, including [data management planning](#), [data storage and backup](#), [creating documentation and metadata](#), and [facilitating data sharing](#). Curators of [McMaster Dataverse](#). RDM services at McMaster are a shared responsibility of McMaster University Library (Sherman Centre) and RHPCS.

<https://rdm.mcmaster.ca/research-data-management>

McMaster Dataverse

[McMaster Dataverse](#) is part of the [Borealis Dataverse](#) network. This research data repository for our faculty, students, and staff provides secure access controls, licensing, version management, data citation, DOI minting, and more. A secure environment houses these files on Canadian servers in Toronto, ON. Does NOT accept content that contains confidential or sensitive information.

<https://borealisdata.ca/dataverse/mcmaster>

DASH - Data Analysis Support Hub

DASH assists researchers with data analysis and software applications (Excel, SPSS, R, Python, etc.). Available to students, staff, and/or faculty members from any discipline. This team has data visualization, GIS, and statistics experts.

<https://library.mcmaster.ca/services/dash>

MacSphere

MacSphere is McMaster University's Institutional Repository (IR). This space preserves and provides access to documents deposited by university departments and centres on campus. This is also the main repository for PhD and Masters' thesis projects. PIDs for MacSphere are through Handle. Repository framework is through DSpace.

<https://macsphere.mcmaster.ca/>

Academic

Data Science Research Lab

Faculty member Dr. Fei Chiang helms the Data Science Research Lab in the Faculty of Engineering at McMaster University. As a research centre, its focus is data management, systems, and machine learning. Their team investigates data quality, profiling, fairness, integration, and privacy through seminars, courses, projects, and cross-disciplinary partnerships (technology, healthcare, financial).

<https://db.cas.mcmaster.ca/>

<https://datasci.cas.mcmaster.ca/>

McArthur Lab

Part of the Michael G. DeGroot Institute for Infectious Disease Research, the McArthur Lab is steered by Dr. Andrew McArthur. This lab uses complex informatics + databases on research relating to genomic surveillance of microbial drug resistance, biological databases, genome assembly and molecular epidemiology sequencing, automated literature curation, and vocabularies for knowledge integration. Resources, tools, and services include infectious disease databases, computing resources, GitLab, Linux Workhorses, Server Provision. Dr. McArthur leads a monthly meeting with the CSU and any FHS labs requiring computational work.

<http://mcarthurbioinformatics.ca>

Research

Office of the Vice-President Research

McMaster's OVPR spearheads research activities at the university and pursues national and international partnerships. They oversee an extensive research portfolio and all affiliate policies, practices, infrastructure, and platforms. McMaster's VPR is Dr. Karen Mossman. Associate Deans of Research (ADRs) in each of our six faculties steer translation of this vision into their respective departments.

<https://research.mcmaster.ca/>

ROADS - Research Office for Administration, Development and Support

ROADS provides researchers with support in navigating funding, ethics, and more. It introduces researchers to Open Access and RDM through links to the library and RDM webpages. ROADS also hosts the data storage and security tools guide for MREB.

<https://research.mcmaster.ca/support-for-researchers/open-access-data-management/>

RHPCS - Research & High-Performance Computing Support

RHPCS is the second connection point for RDM services at McMaster (with MUL). RHPCS supports researchers across all disciplines in accessing and developing complex digital research infrastructure. Currently, it exists partially on a cost recovery basis. Services related to RDM include data storage and backups, connections to High-Performance computing resources ([Compute Canada/Compute Ontario](#), [SHARCNET](#)), laboratory and research infrastructure design, web survey implementation ([LimeSurvey](#)), and more.

<https://research.mcmaster.ca/research-resources/local-it/rhpcs/services/>

<https://researchcomputing.mcmaster.ca/>

MILO - McMaster Industry Liaison Office

MILO connects researchers with industry collaborators and supports the dissemination of research results through commercialization. RDM-related connections are working with researchers on data sharing agreements, data security, and licensing.

<https://research.mcmaster.ca/mcmaster-industry-liaison-office-milo/>

MREB - McMaster Research Ethics Board

MREB is the Research Ethics Board for projects that do not fall under the scope of HiREB or AREB. Their online system is McMaster Research Ethics Manager (MacREM), where researchers can submit, review, and update ethics applications. In addition to an online application portal, they provide sample documents and guides, resources, tutorials, and more. Human ethics boards are important collaborators on data storage, security, privacy, sharing, preservation, re-use. Their site includes links for further learning on research data security, research data management plans, and data encryption. It does not yet link to the McMaster RDM site.

<https://research.mcmaster.ca/ethics/mcmaster-research-ethics-board-mreb/>

HiREB - Hamilton Integrated Research Ethics Board

HiREB exists as a bridge between McMaster University's Faculty of Health Sciences and our city's hospital network, St. Joseph's Healthcare Hamilton, and Hamilton Health Sciences. This integrated group works to protect research participants in research projects involving physicians, staff, students, volunteers, visitors, or patients. This ensures projects conducted within all three institutional frameworks align with current ethical standards. HiREB advises researchers on all ethical, scientific, and technical aspects of planning research projects, including RDM guidance.

<https://hireb.ca/>

AREB - Animal Research Ethics Board

AREB is the ethics board for all McMaster research involving living animals. Research, teaching, and testing activities require completing an Animal Utilization Protocol (AUP). AREB reviews AUPs for classroom and research activities and approves those aligned with legal and ethical responsibilities.

<https://healthresearch.mcmaster.ca/areb/>

Research Centres and Institutes

SCDS - Sherman Centre for Digital Scholarship

The Sherman Centre is one of the connection points for RDM services at McMaster (with RHPCS). It connects to pedagogy, research, training, publishing and more through digital technology. Sherman Centre offers researcher consultation + support services. In addition to the RDM and DASH teams, it hosts a graduate residency, an invited speaker series, an undergraduate course, and four annual workshop series. All events are free and open to the public, with 40+ past workshops available as online learning modules.

<https://scds.ca/what-we-do/technical-services/>

MIRI – McMaster Indigenous Research Institute

MIRI is a leader in the field of Indigenous research. They undertake and facilitate research partnerships and support Indigenous Knowledge and methodologies. MIRI also provides training and guidance for Indigenous and Non-Indigenous researchers, ethics boards, and decision-makers. As such, they are an important partner in developing an Institutional Strategy that aligns with Indigenous Data Sovereignty principles.

<https://miri.mcmaster.ca/>

MacData

MacData is a research institute connecting researchers, students, industry, government, and community through data leadership. Objective is to foster collaboration between the many institutes, centres, and researchers who are working with data. They also run seminars, workshops, and graduate fellowships on big data. Their website has long lists structured around interest points, including pages on RDM and Data Harmonization and Integration.

<https://macdata.mcmaster.ca/>

Spark - A centre for social research innovation

Spark is a centre supporting researchers, especially those working in social sciences. They host events (ex. Transparency and Positionality in Coding, Social Observation), training (managing impact), resources for research (toolkits, software, data), services (assessments, research support), and CoLab (a shared working space). They connect to our RDM site on their resources page. Spark is also a collaborator with RDM services through a new Data Champions Grant through the Alliance. Spark houses the SEAL Lab and the RDC.

<https://spark.mcmaster.ca/>

SEAL - Secure Empirical Analysis Lab

SEAL is a high security lab within Spark providing researchers with a safe place to store and access confidential data from companies, governments, researchers, and community organizations. SEAL is a trusted intermediary between data owners and research users. In addition to this resource, services relating to RDM include secure storage, data sharing agreements, and a selection of datasets available in SEAL. This operates on a cost recovery model with 1200\$/year starting costs - some of which can be waived. No data quota, backed up to local drives and encrypted off site backups. McMaster Dataverse hosts the CRA T3010 Charity Returns dataset.

<https://seal.mcmaster.ca/>

CRDCN - Canadian Research Data Centre Network

Hosted at McMaster, the CRDCN is a national research and training platform for quantitative social and health sciences. Funded by SSHRC, CIHR, SSHRC, CIHR, CFI, the FRQ, Statistics Canada and forty-two partner universities. This network provides access to Statistics Canada data via affiliated sites at 33 campuses across Canada. Through its infrastructure, services, and training, it helps inform and advance public policy and knowledge. The McMaster Research Data Centre (RDC) is an affiliated site of the CRDCN.

<https://crdcn.org>

RDC - Statistics Canada Research Data Centre at McMaster

McMaster's member entity of the CRDCN. A secure facility/computer lab on campus. To use the centre researchers with approved projects must become "deemed employees" of Statistics Canada. This gives access to Statistics Canada's population, social, and health surveys, as well as the Census and administrative databases. Access is free of charge to faculty, students, and staff. Researchers cannot use datasets outside of the facility; only vetted output is released. The McMaster RDC is administered by the Faculty of Social Sciences.

<https://rdc.mcmaster.ca/>

<https://socialsciences.mcmaster.ca/statistics-canada-research-data-centre-rdc-at-mcmaster>

UTS – University Technology Services

UTS provides IT support for nearly everyone in the university. They administer [MacDrive](#) and the University Office 365 including OneDrive, Teams, Sharepoint. Students have access to Google drive as part of GSuite, but not MacDrive. OneDrive gives all users, 1 TB and up to 5 TB on request.

<https://uts.mcmaster.ca/services/communication-collaboration-and-storage/>

IT Security – Information Technology Security

IT Security is a division of UTS. They provide resources and guidance for students, instructors, researchers, employees, technical, and leaders. They have lead McMaster researchers through several security incidents, including ransomware attacks. IT security overlaps with RDM by providing information on data classification, data lifecycle, IT security hygiene, and how to report an incident.

<https://informationsecurity.mcmaster.ca/people/researchers/>

MacDrive

MacDrive is a customized McMaster version of [Seafile](#), which is an open source EFSS (Enterprise File Synchronization and Sharing) solution, similar to DropBox. It allows people and teams to store files on a central server and access them via a web interface or client on a device. Folder encryption possible. Triple copies of all data are distributed across a cluster of computers which are distributed across two separate datacenters. This service is available to employees but not students. 300 GB storage is accessible through MacID login. UTS supports MacDrive.

<https://macdrivehelp.mcmaster.ca/>

Faculty of Health Sciences

Health Research Services

Health Research Services is a research office specific to FHS supporting health sciences researchers. This includes secure and manage funding, as well as connecting researchers with laboratory equipment and negotiating the allocation of research space. Through a focus on ethics and biohazards, they link out to HiREB, AREB, and MREB, as well as Clinical Trials Ontario. They are the on-campus collaborator for the Research Institute of St. Joe's Hamilton and Hamilton Health Sciences' Research Administration.

<https://healthresearch.mcmaster.ca/>

Indigenous Health Learning Lodge (IHLL)

IHLL works alongside the FHS to support education and research that intersects with Indigenous health and well-being. This initiative developed out of a comprehensive Indigenous Health Initiative to integrate Indigenous cultural knowledge. They foster Indigenous health through faculty leadership and support, administrative connectivity, student support and services, research, education and curriculum, and Indigenous ways of knowing.

<https://ihll.mcmaster.ca/>

CSU - FHS Computer Services Unit

CSU is the Computer Services Unit of the Faculty of Health Sciences. They provide IT hardware and software support to FHS and its constituent administrative, clinical, educational, and research stakeholders. RDM-related services include MacDrop file servers. The Cluster—a group of 6 FHS labs—in CSU jointly funds an \$8M cloud computing environment with significant storage capacity.

<https://www.csu.mcmaster.ca/>

REDCap

REDCap (Research Electronic Data Capture) is a secure software used to build and manage research surveys and databases. Developed by Vanderbilt University, it has grown to a network of 5900+ institutional partners internationally. Although REDCap is not open-source software, it is available at no charge to consortium members who have successfully validated an end-user license agreement. As such, it is available to McMaster researchers at no cost. It is administered by the CSU.

<https://mctr.mcmaster.ca/>

Health Sciences Library

The HSL provides researchers, faculty, students, and staff with Research, eResource, Service Desk, Guides and Tutorials, Research Consultation services. One of HSL's planning objectives is to "Learn more about health researchers' data management need with a view to defining future services" and to "Partner in the administration of campus-wide data management needs assessment survey" (Health Sciences Library, 2019). Researchers can book a consultation with an HSL Librarian. There is also a Researcher's Toolkit LibGuide with information about RDM.

<https://hsl.mcmaster.ca/>

<https://hslmcmaster.libguides.com/research-toolkit>

The Research Institute of St. Joe's Hamilton

This academic and research community serves as the research arm of St. Joseph's Hospital, Hamilton. Their research efforts balance across five pillars: mental health and addiction; lungs + chest; Father Sean O'Sullivan Research Centre (biostatistics + imaging research); kidney; genitourinary.

<https://research.stjoes.ca/>

Hamilton Health Sciences' Research Administration

This administrative body coordinates research at on varying scales across seven hospitals and a cancer centre. These efforts are spearheaded by five specialized research institutes: PHRI; Escarpment Cancer Research Institute;

Thrombosis and Atherosclerosis Research Institute; Offord Centre for Child Studies; GERAS Centre for Aging Research; and the Clinical Research Library and Biobank.

<https://www.hamiltonhealthsciences.ca/research-innovation/research/>

PHRI - Population Health Research Institute

PHRI is one of five joint research institutes of McMaster University and Hamilton Health Sciences. This extensive international network includes collaborators in 102. They conduct a wide range of research including large trials, studies, and registries. PHRI has several RDM professionals on their team.

<https://www.phri.ca/>

Appendix 5: RDM - National + International Context

In addition to institutional-level stakeholders and service providers, McMaster's research enterprise connects with national and international stakeholders. Research, data sharing, data security, and collaboration are not limited to our institution.

As such, our RDM services must also align with expectations and best practices within those research communities we are a part of. This list is by no means exhaustive; however, it serves as a starting point to highlight the national and international contexts of RDM services at McMaster.

National Level

Government:

- [Innovation, Science and Economic Development Canada](#)
 - [CIHR/IRSC](#)
 - [NSERC/CRSNG](#)
 - [SSHRC/CRSH](#)

Ethics:

- [FNIGC/CGIPN](#) - First Nations Information Governance Centre / Le Centre de gouvernance de l'information des Premières Nations
- [CAREB-ACCER](#) - Canadian Association of Research Ethics Boards / L'association canadienne des comités d'éthique de la recherche

Libraries:

- [CARL/ABRC](#) - Canadian Association of Research Libraries / Association des bibliothèques de recherche du Canada
- [Scholars Portal](#) – A Service of the Ontario Council of University Libraries

Repositories:

- [FRDR – DFDR](#) - Federated Research Data Repository
- [Borealis Dataverse](#)

Researchers:

- [CARA/ACAAR](#) - Canadian Association of Research Administrators/ L'Association canadienne des administratrices et des administrateurs de recherche
- [CAGS](#) - Canadian Association for Graduate Studies
- [U15](#) - Group of Canadian Research Universities

IT + Research Services:

- [CANARIE](#) - Canadian Network for the Advancement of Research, Industry and Education
- [Compute Canada/Calcul Canada](#)
- [DRAC](#) - Digital Research Alliance of Canada

International Level

Research Services:

- [CODATA](#) – Committee on Data: International Science Council
- [euroCRIS](#) – International Organization for Research Information
- [NISO](#) - National Information Standards Organization
- [NIH](#) - National institutes of Health
- [NSF](#) - National Sanitation Foundation
- [DataCite](#)
- [AAU](#) - Association of American Universities
- [ORCID](#) - Open Researcher and Contributor ID
- [DOI](#) – Digital Object Identifier System

Libraries/Researchers:

- [AAAS](#) - American Association for the Advancement of Science
- [APLU](#) - Association of Public & Land Grant Universities
- [ARDC](#) – Australian Research Data Commons
- [Data Curation Network](#)
- [DataONE](#) – Data Observation Network for Earth
- [DCC](#) - Digital Curation Centre
- [RDA](#) – Research Data Alliance
- [RDAP](#) - Research Data Access & Preservation Association

IT:

- [XSEDE](#) - Extreme Science and Engineering Discovery Environment
- [EGI](#) - Advanced Computing Services for Research
- [PRACE](#) – Partnership for Advanced Computing in Europe

Repositories:

- [The Dataverse Project](#)
- [Dryad](#)
- [FigShare](#)
- [Re3Data](#)

Appendix 6: Current RDM Requirements

Many grants and publishers have RDM requirements as a condition of funding and publishing. These are constantly in development and the scope of organizations McMaster researchers intersect with that are requesting Data Management Plans, Information Security Frameworks, and Data Deposit are extensive. However, a brief overview of the most prominent requirements is outlined here.

Data Management Plans

As of June 2022, [the Tri-Agencies have announced the first set of grants requiring Data Management Plans \(DMPs\)](#). These requirements are a step towards an anticipated phasing in DMP requirements for further competitions. Further details on expectations for DMPs and how the Tri-Agencies will consider them in the adjudication process will come once the grants are launched. In addition to national funding agencies, other grants are requesting data management strategies, such as ORF-RI and IDRC.

IT Security

Currently, IT security requirements for grants are limited. However, with the launch of the [National Security Guidelines for Research Partnerships](#) in Summer 2021, we anticipate this increasing in priority in the years to come. Grants currently requiring Information Security plans are [NSERC Alliance Grants](#), [ORF-RI](#), [CFI](#), and some Private Sector Partnerships.

Data Deposit

The [Tri-Agency Research Data Management Policy](#) outlines an aim that grant recipients will be required to deposit their project data in a digital repository by time of publication, “where ethical, cultural, legal and commercial requirements allow, and in accordance with the FAIR principles and the standards of their disciplines.” (I. Government of Canada, n.d.) Limited data deposit requirements have existed for [CIHR recipients since January 1, 2008](#). Since 1990, SSHRC’s [Research Data Archiving Policy](#) has required all research data collected using SSHRC funds should be preserved and made available to others within 2 years. However, SSHRC does not yet have a formal data deposit requirement.

Appendix 7: Definitions

Building this glossary was influenced by Queen’s University’s [Research Data Management Institutional Strategy](#) and McGill University’s [RDM Strategy Draft v1.0](#).

Archiving: “Archiving” refers to “a curation activity that ensures that data are properly selected, stored, and can be accessed, and for which logical and physical integrity are maintained over time, including security and authenticity.” RDM Terminology Working Group. (2021). Research Data Management Terminology. CODATA. Retrieved June 30, 2022, from <https://codata.org/initiatives/data-science-and-stewardship/rdm-terminology-wg/rdm-terminology/>

Data: “Data are facts, measurements, recordings, records, or observations collected by researchers and others, with a minimum of contextual interpretation. Data may be in any format or medium taking the form of text, numbers, symbols, images, films, video, sound recordings, pictorial reproductions, drawings, designs or other graphical representations, procedural manuals, forms, diagrams, workflows, equipment descriptions, data files, data processing algorithms, software, programming languages, code, or statistical records.” Innovation, Science and Economic Development Canada. (2021). Frequently Asked Questions - Tri-Agency Research Data Management Policy—Science.gc.ca. Government of Canada. Retrieved June 30, 2022, from https://science.gc.ca/eic/site/063.nsf/eng/h_97609.html#1a Tri-Agency Definition adapted from CASRAI.

Data Deposit: ““Data deposit” refers to when the research data collected as part of a research project are transferred to a research data repository. The repository should have easily accessible policies describing deposit and user licenses, access control, preservation procedures, storage and backup practices, and sustainability and succession plans. The deposit of research data into appropriate repositories supports ongoing data-retention and, where appropriate, access to the data. Ideally, data deposits will include accompanying documentation, source code, software, metadata, and any supplementary materials that provide additional information about the data, including the context in which it was collected and used to inform the research project. This additional information facilitates curation, discoverability, accessibility and reuse of the data.” Innovation, Science and Economic Development Canada. (2021). Frequently Asked Questions - Tri-Agency Research Data Management Policy—Science.gc.ca. Government of Canada. Retrieved June 30, 2022, from https://science.gc.ca/eic/site/063.nsf/eng/h_97609.html#1a

Data Management Plan (DMP): “A data management plan (DMP) is a living document, typically associated with an individual research project or program that consists of the practices, processes and strategies that pertain to a set of specified topics related to data management and curation. DMPs should be modified throughout the course of a research project to reflect changes in project design, methods, or other considerations. DMPs guide researchers in articulating their plans for managing data; they do not necessarily compel researchers to manage data differently.” Innovation, Science and Economic Development Canada. (2021). Frequently Asked Questions - Tri-Agency Research Data Management Policy—Science.gc.ca. Government of Canada. Retrieved June 30, 2022, from https://science.gc.ca/eic/site/063.nsf/eng/h_97609.html#1a

Data Stewardship: “The process of Data Stewardship involves ensuring effective control and use of data assets and can include creating and managing metadata, applying standards, managing data quality and integrity, and additional data governance activities related to data curation. It also may include creating educational materials, policies, and guidelines around data at an institution.” National Library of Medicine. (n.d.) Data Stewardship. Network of the National Library of Medicine. Retrieved June 30, 2022, from <https://nmlm.gov/guides/data-glossary/data-stewardship>.

Digital Research Infrastructure: Digital Research Infrastructure describes the suite of interrelated equipment, computer hardware and software, and data collections required to carry out scientific enquiry, research, scholarship, or creative practice, as well as the complementary expertise, services, and resources that enable their sharing, adoption, use, and reuse.

Adapted from Innovation, Science and Economic Development Canada. (2019). Digital Research Infrastructure Contribution Program: Program guide. Government of Canada. Retrieved September 9, 2022 from <https://ised-isde.canada.ca/site/digital-research-infrastructure/en/digital-research-infrastructure-contribution-program-program-guide#1>.

Indigenous Research: “Indigenous Research” is “research in any field or discipline that is conducted by, grounded in or engaged with First Nations, Inuit, Métis or other Indigenous nations, communities, societies or individuals, and their wisdom, cultures, experiences or knowledge systems, as expressed in their dynamic forms, past and present.” Social Sciences and Humanities Research Council. (2021, May 4). Definitions of Terms. Government of Canada. Retrieved June 30, 2022, from <https://www.sshrc-crsh.gc.ca/funding-financement/programmes-programmes/definitions-eng.aspx>

Industry Partner Organization: “A for-profit organization, or an organization that assists, supports, connects and/or represents the common interests of a group of for-profit, incorporated organizations, such as an industry association or a formal or informal consortium.” Social Sciences and Humanities Research Council. (2021, May 4). Definitions of Terms. Government of Canada. Retrieved June 30, 2022, from <https://www.sshrc-crsh.gc.ca/funding-financement/programmes-programmes/definitions-eng.aspx>

Intellectual Property: ““Intellectual Property” means databases, audio-visual material, electronic circuitry, biotechnology and genetic engineering products, computer software recorded in any format, inventions, discoveries and all other products of research (which inventions, discoveries or other products are capable of protection pursuant to any law of Canada or any other country or which may be otherwise licensable) where any of the foregoing are created, whether by discovery, invention or otherwise by an IP Creator as hereinafter defined.” University Secretariat. (2018). Joint Intellectual Property Policy. McMaster University. Retrieved June 30, 2022, from <https://secretariat.mcmaster.ca/app/uploads/Joint-Intellectual-Property.pdf>.

Persistent Unique Identifier: “A Persistent Unique Identifier (PID) is a string of letters and numbers used to distinguish between and locate different objects, people, or concepts. A well-known example of a PID is a Digital Object Identifier (DOI) which is used to locate specific digital objects, frequently a journal article. Another example is ORCID, a PID for researchers.” National Library of Medicine. (n.d.) Persistent Unique Identifier. Network of the National Library of Medicine. Retrieved June 30, 2022, from <https://nmlm.gov/guides/data-glossary/persistent-unique-identifier>.

Preservation: “An activity within archiving in which specific items of data are maintained over time so that they can still be accessed and understood through changes in technology.” RDM Terminology Working Group. (2021). Research Data Management Terminology. CODATA. Retrieved June 30, 2022, from <https://codata.org/initiatives/data-science-and-stewardship/rdm-terminology-wg/rdm-terminology/>

Researcher: “A Researcher is involved in an undertaking to extend knowledge through a disciplined inquiry or systematic investigation.” University Secretariat. (2017). Research Integrity Policy. McMaster University. Retrieved June 30, 2022, from <https://secretariat.mcmaster.ca/app/uploads/Research-Integrity-Policy.pdf>.

Research Data: “Research data are data that are used as primary sources to support technical or scientific enquiry, research, scholarship, or creative practice, and that are used as evidence in the research process and/or are commonly accepted in the research community as necessary to validate research findings and results. Research data may be experimental data, observational data, operational data, third party data, public sector data, monitoring data, processed data, or repurposed data. What is considered relevant research data is often highly contextual, and determining what counts as such should be guided by disciplinary norms.”

Innovation, Science and Economic Development Canada. (2021). Frequently Asked Questions - Tri-Agency Research Data Management Policy. Government of Canada. Retrieved June 30, 2022, from https://science.gc.ca/eic/site/063.nsf/eng/h_97609.html#1a

Research Data Management: “Research Data Management” is “the storage of, access to and preservation of data produced from one or more investigations, or from a program of research. Research data management practices cover the entire lifecycle of the data, from planning the investigation to conducting it, and from backing up data as it is created and used to preserving data for the long term after the research has concluded. It also includes data-sharing, where applicable.” Social Sciences and Humanities Research Council. (2021, May 4). Definitions of Terms. Government of Canada. Retrieved June 30, 2022, from <https://www.sshrc-crsh.gc.ca/funding-financement/programs-programmes/definitions-eng.aspx>

Research Institution: “An institution with a research mandate and qualified research staff and/or research facilities.” Social Sciences and Humanities Research Council. (2021, May 4). Definitions of Terms. Government of Canada. Retrieved June 30, 2022, from <https://www.sshrc-crsh.gc.ca/funding-financement/programs-programmes/definitions-eng.aspx>