DRAFT: University of Calgary Research Data Management Strategy

Version Dated September 14, 2022
Prepared by the RDM Strategy Working Committee

Contents

Introduction

Research data management (RDM) is the range of processes and procedures "applied through the lifecycle of a research project to guide the collection, documentation, storage, sharing and preservation of research data"[1]. All academic researchers who work with research data apply these processes and procedures in their work in a variety of ways and with varying levels of formality. Given recent developments in the research landscape – including the open science movement, drives for transparency and reproducibility, concerns about privacy and security, and increasing research costs – academic research is now at a point where it is crucial to recognize the importance of research data and commit to implementing and supporting best practices in RDM.

In March 2021, the Tri-Agencies released their [Research Data Management Policy](#), which applies to all post-secondary institutions and research hospitals which are eligible to administer SSHRC, NSERC, or CIHR funds, and to all researchers holding grants from those agencies. This Policy rests on three pillars:

1. Institutional Research Data Management (RDM) Strategies: "Each postsecondary institution and research hospital eligible to administer CIHR, NSERC or SSHRC funds is required to create an institutional RDM strategy and notify the agencies when it has been completed. The strategy must be made publicly available on the institution's website, with contact information to which inquiries about the strategy can be directed" (section 3.1 of the Policy). These strategies must be completed by March 1, 2023.
2. Data Management Plans: "All grant proposals submitted to the agencies should include methodologies that reflect best practices in RDM. For certain funding opportunities, the agencies will require data management plans (DMPs) to be submitted to the appropriate agency at the time of application, as outlined in the call for proposals; in these cases, the DMPs will be considered in the adjudication process" (section 3.2 of the Policy). The initial funding opportunities requiring DMPs will be launched in the fall of 2022, and additional programs with this requirement will roll out over the following weeks and months.
3. Data Deposit: "Grant recipients are required to deposit into a digital repository all digital research data, metadata and code that directly support the research conclusions in journal publications and pre-prints that arise from agency-supported research.... The deposit must be made by time of publication" (section 3.3 of the Policy). This requirement will be implemented after the Tri-Agencies have reviewed the published institutional strategies and "in line with the readiness of the Canadian research community" (section 4 of the Policy).

This Research Data Management Strategy has four purposes:

1. To embody the University of Calgary's commitment to meeting the institution-facing requirements of the Tri-Agency RDM Policy, and to enable our research community to meet the researcher-facing requirements of the Policy.
2. To allow the University and our research community to address the RDM requirements and obligations being implemented by other funders – such as the Canada Foundation for Innovation, the U.S. National Institutes of Health, and the U.S. National Science Foundation – as well as those being driven by changes in academic publishing and in academia more broadly.

---

[1] "1d: What is research data management?", *Frequently Asked Questions: Tri-Agency Research Data Management Policy*, last accessed July 13, 2022 at https://www.science.gc.ca/eic/site/063.nsf/eng/h_97609.html#1d

3. To establish a foundation on which we will build a culture of good research data management practices for current and future generations of researchers.
4. To establish a foundation for discussions about research data management with our research participants and partners, including partners from industry, community organizations, and Indigenous communities.

Guiding Principles

The creation and implementation of this Strategy have been informed by four guiding principles:

Research Excellence and Impact

The University of Calgary's RDM Strategy has a key role to play in sharpening our focus on research and scholarship and allowing us to drive innovation, which are key priorities of our *Eyes High* Strategy and Research Plan.

Support for our Researchers and our Partners in Research

The University of Calgary is committed to supporting our researchers and research communities in meeting RDM requirements and in incorporating best practices in RDM into their work. This will allow us to increase research capacity and drive innovation, as outlined in the Research Plan. The University further acknowledges that we have a role to play in supporting the RDM rights of all parties involved in research.

Collaboration

The University of Calgary recognizes that collaboration across RDM-supporting units and portfolios, and with external stakeholders and community partners, is crucial to ensuring we foster an environment conducive to best practices in research data management, and to providing the support our researchers need to meet RDM requirements. Collaboration with other institutions and national organizations is also essential in order to support RDM in the broader research environment. The University further acknowledges that RDM best practices can support efforts to move towards interdisciplinary and transdisciplinary research, as outlined in *Eyes High* and *Growth Through Focus*.

Context-Based Approach

The University of Calgary acknowledges that there can be no "one-size-fits-all" approach to RDM, given the range of research questions, data types, methods of data collection and analysis, disciplinary practices, ethical obligations, legal and regulatory frameworks, and partnership environments with which and within which our researchers work. The University will support our researchers in implementing the RDM best practices appropriate to their research context, and will ensure that our RDM support is at a consistently high standard regardless of the research context.

The Importance of Research Data and Research Data Management

Research data, as defined by the Tri-Agencies, "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"[2]. Research data result from the work conducted on research materials, be they books, ecosystems, subatomic particles, genes, individual humans, or entire communities[3]; they can even be derived from other research data. At a fundamental level, the University of Calgary recognizes that data, in all of their varied forms, are at the heart of the research enterprise.

In its 2016 [Statement of Principles on Digital Data Management](#), the Tri-Agencies stated that "The ability to store, access, reuse and build upon digital research data has become critical to the advancement of science and scholarship, supports innovative solutions to economic and social challenges, and holds tremendous potential for Canada's productivity, competitiveness and quality of life"[4]. Regardless of the types and quantities of research data our researchers work with, or their potential significance beyond a particular project, the University of Calgary believes that there are real and tangible benefits of RDM at the level of the individual research project, including increasing the rigour, transparency, and efficiency of research, as well as reducing costs. Employing best practices in RDM will also allow researchers to more effectively share their research data, where such sharing is permitted and appropriate, to allow future research and the incremental growth of knowledge. In short, good RDM practices make research better and accelerate the expansion of knowledge.

The University of Calgary acknowledges that there are many factors that must be considered in implementing best practices in RDM: these include different types and scales of data; different data collection and analysis methodologies; evolving research practices; technological changes; costs; human and infrastructure resource needs; and ethical and legal considerations. We will work to support our research communities in navigating these factors, and to create an institutional framework which provides comprehensive support for RDM.

---

[2] "1b: What are research data?", *Frequently Asked Questions: Tri-Agency Research Data Management Policy*, last accessed July 13, 2022 at https://science.gc.ca/eic/site/063.nsf/eng/h_97609.html#1b

[3] "1c: How are research materials related to research data?", *Frequently Asked Questions: Tri-Agency Research Data Management Policy*, last accessed July 13, 2022 at https://science.gc.ca/eic/site/063.nsf/eng/h_97609.html#1c

[4] "Preamble", *Tri-Agency Statement of Principles on Digital Data Management*, last accessed July 13, 2022 at https://www.science.gc.ca/eic/site/063.nsf/eng/h_83F7624E.html

Indigenous Data Sovereignty

In *ii' taa'poh'to'p*, the University of Calgary's Indigenous Strategy, the University made seven commitments to truth, reconciliation, and transformation with regard to its relationship to Indigenous peoples. This Research Data Management Strategy will uphold these commitments and work to further their implementation. We recognize that the commitments made in this strategy are only the beginning of the discussions around Indigenous data sovereignty at the University of Calgary, and commit to the work of co-developing with our partners a framework to support Indigenous data sovereignty.

In this Strategy, we commit to respecting, supporting, and enabling Indigenous peoples' inherent sovereignty over the research data generated by and with First Nations, Métis, and Inuit peoples and communities within the land currently known as Canada, including the rights to own, control, collect, access, possess, protect, use, and share these data. We will work with our Indigenous partners in research to ensure that Indigenous ways of knowing, doing, connecting and being are respected in all aspects of research data management. We will work with our research community to ensure that all researchers working with Indigenous individuals, communities, collectives, and/or organizations treat the data generated by this work in a good way. We acknowledge that this may mean parallel paths are necessary for the management of research data relating to Indigenous peoples and communities. We also acknowledge that the University of Calgary has a role to play in helping to build capacity for RDM for Indigenous researchers, communities, collectives, and organizations.

Specific areas of institutional support for RDM in regards to Indigenous data sovereignty are described in the following sections.

Scope

This Strategy is relevant to all University of Calgary researchers, as defined in the University's Research Integrity Policy.

This Strategy applies to all digital research data (born-digital or digitized) generated by researchers at the University of Calgary. It does not apply to physical materials.

The support services described in this Strategy will be provided to all research activities, regardless of whether they are led by faculty, postdoctoral fellows, graduate or undergraduate students; and whether they are funded by the Tri-Agency or other funders, or unfunded. RDM support services will be provided to all those involved in research at the University of Calgary, including faculty, postdoctoral fellows, staff, and students.

Oversight and Review

This Strategy comes under the purview of the Vice President (Research). A steering committee, comprising key institutional stakeholders in RDM, is responsible for implementation and post-implementation review of the Strategy. These stakeholders include representatives of the Vice President (Research) Office; the Research Services Office, including the Research Ethics Boards; Libraries and Cultural Resources, Information Technology, including Research Computing Services; the Office of Indigenous Engagement; the Cumming School of Medicine; Graduate Studies; the Graduate Students

Association; the Postdoctoral Association; and legal counsel. The steering committee will consult with external stakeholders and community partners as needed.

The Strategy will be implemented on a five-year timeline, with a review in the fourth year; it will be revised as appropriate at that time, to ensure it continues to meet the needs of the University of Calgary and its diverse research community. As this is the initial RDM Strategy, there will be an additional review after the first year of implementation to ensure that any major issues which may arise are addressed appropriately.

Institutional Support for Research Data Management

Institutional Policies and Processes

The University of Calgary currently has several institutional policies, systems, schedules and processes which address different aspects of RDM to varying degrees (see Appendix A for a full list). Creating and maintaining a robust policy, procedure, and process framework for RDM will be essential to allowing us to support our research communities and meet our institutional obligations.

In the first five-year term of this strategy, the University of Calgary will

- Review our existing policies, procedures, standards, systems and schedules (see the full list in Appendix A) to determine if revisions are needed to adequately address RDM concerns;
- Implement any necessary revisions identified during the review process, such as creating new policies, making amendments to existing documents, and/or establishing common definitions;
- Educate researchers, staff, and other stakeholders as to which elements of the policy framework apply to research data, and which ones do not;
- Ensure that all RDM-related policies, practices and procedures are supportive and respectful of Indigenous ways of knowing, doing, connecting and being, including respecting and supporting Indigenous data sovereignty, governance, and management;
- Develop and implement a robust Libraries and Cultural Resources policy for Collection Development which encompasses the collection and preservation of research data; and
- Ensure that the development of our RDM-related policy framework keeps up with the implementation of RDM policies and requirements by funders, publishers, and legislative bodies.

IT Infrastructure

Information technology infrastructure which supports RDM is currently primarily provided by two units on campus: Research Computing Services and Libraries and Cultural Resources. We are working from a solid base that provides many services for our research community, including active storage, secure data storage, a range of High-Performance Computing (HPC) services, and an institutional data repository. As both technology and the needs of our researchers evolve, we must continue to provide both internal resources and access to external resources to accommodate this evolution.

In the first five-year term of this strategy, the University of Calgary will

- Continue institutional support for, and strong policy in, areas such as identity and access management;
- Undertake hardware refreshes and/or replacements in a timely fashion;
- Ensure necessary software licenses and service subscriptions are acquired and/or maintained;
- Pursue CoreTrustSeal certification for our institutional data repository;
- Ensure researchers have access to long-term data preservation and secure storage for sensitive data;
- Explore options for ensuring researchers using RCS infrastructure are engaging in best practices in RDM before they begin using the infrastructure (e.g., through the submission of data management plans);
- Ensure that RDM-related IT infrastructure is appropriately resourced, in terms of both funding and personnel; and
- Ensure that the development and maintenance of our RDM-related IT infrastructure keep up with the implementation of RDM policies and requirements by funders, publishers, and legislative bodies.

Where appropriate, we will work with partners such as the Digital Research Alliance of Canada and the Borealis Dataverse Repository service to maximize the availability of digital research infrastructure to which our research communities have access.

Support Services

Support services for RDM are currently primarily provided by three units: Libraries and Cultural Resources, IT/Research Computing Services, and the Research Services Office. These services include basic training and consultation for key areas of RDM, such as data management plans; basic training and consultation for data management within the High-Performance Computing setting; and basic guidance on meeting funders' RDM requirements. We have a core of knowledgeable and experienced staff who can guide the development of our RDM support services as they expand to meet the needs of our research communities.

In the first five-year term of this strategy, the University of Calgary will

- Increase the number and range of RDM-related training and capacity-building opportunities for faculty, postdoctoral fellows, staff, and students;
- Increase the number of librarians and staff in Libraries and Cultural Resources who can provide timely, accurate and detailed RDM-related services, support, and training, particularly in the areas of data curation and preservation;
- Increase the number of staff in IT/Research Computing Services who can provide timely, accurate and detailed RDM-related services, support and training;
- Increase the ability of staff in the Research Services Office to provide timely, accurate and detailed RDM-related support for grant applications, ethics applications, and contracts and agreements, by providing RSO staff with appropriate resources and training opportunities;
- Increase the ability of the Indigenous Research Support Team to provide timely, accurate and detailed RDM-related support specific to Indigenous data sovereignty, by providing IRST

members with appropriate resources and training opportunities, and by ensuring that this team is funded appropriately on an ongoing basis;

- Increase the ability of research-supporting staff in the faculties – e.g., research facilitators, privacy specialists, and others – to provide RDM-related support to researchers, by providing these staff with appropriate resources and training opportunities;
- Ensure that researchers can easily locate and access usable, up-to-date RDM-related resources and services through the development of an RDM-focused virtual support 'desk';
- Ensure that researchers have resources and opportunities which allow them to increase their knowledge and skills in the area of Indigenous data sovereignty area (e.g., supporting training programs for frameworks such as OCAP®);
- Support researchers in working appropriately with Indigenous partners and communities in terms of Indigenous data sovereignty and research data management;
- Establish a proactive and reactive communications plan to keep researchers informed of changes and opportunities in the RDM landscape;
- Ensure that RDM-related support services are appropriately resourced, in terms of both funding and personnel; and
- Ensure that the development and maintenance of our RDM-related support services keep up with the implementation of RDM policies and requirements by funders, publishers, and legislative bodies.

Collaboration across the RDM-supporting units - including but not limited to Research Computing Services, Libraries and Cultural Resources, and the Research Services Office - will be essential in ensuring the success of our RDM efforts at the University of Calgary; ways to support this collaboration should be explored, such as the development of a system or platform to link RDM-related services, requirements, and compliance across administrative units.

Building a Broader Research Data Management Culture

Research data management will impact our research community's regional, national, and international colleagues and partners. For this reason, the University of Calgary will participate in efforts to build a broader culture of RDM, both within and beyond the university.

In the first five-year term of this strategy, the University of Calgary will

- Develop a network of faculty, postdoctoral fellows, students and staff who will act as 'Research Data Champions' to promote and demonstrate the importance of research data and RDM to their colleagues;
- Work with our partner institutions, both directly and through organizations such as Universities Canada, U15, and the Canadian Association of Research Libraries, to advocate for common approaches to and increased support for RDM-related needs, such as financial support for researchers and institutions for RDM-related costs, frameworks for supporting and implementing Indigenous data sovereignty at research institutions and in Indigenous communities, *inter alia*;

- Work with Indigenous stakeholders and communities to provide access to research data generated by, with or about Indigenous individuals or communities which are held at the University of Calgary;
- Work with Indigenous stakeholders and communities to build capacity for data governance and management within communities; and
- Encourage our tenure and promotion committees to explore how to incorporate data-related work into their review of research excellence and impact, in line with the University's commitment to the San Francisco Declaration on Research Assessment.

Looking Ahead

This Strategy is a living document: it will necessarily evolve as research data management requirements, practices, and understanding evolve. Moving forward, the University of Calgary will ensure that our institutional structures support our research communities as they respond and adapt to this evolution, and that best practices in RDM are part of the fundamental fabric of research as we continue our journey of research excellence.

Definitions

Research Data

Data that are used 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 a given research community as necessary to validate research findings and results.

(See "1b: What are research data?", *Frequently Asked Questions: Tri-Agency Research Data Management Policy*, last accessed July 13, 2022 at https://science.gc.ca/eic/site/063.nsf/eng/h_97609.html)

Research Data Management

The range of processes and procedures "applied through the lifecycle of a research project to guide the collection, documentation, storage, sharing and preservation of research data".

(See "1d: What is research data management?", *Frequently Asked Questions: Tri-Agency Research Data Management Policy*, last accessed July 13, 2022 at https://www.science.gc.ca/eic/site/063.nsf/eng/h_97609.html#1d)

Sensitive Data

Data that must be safeguarded against unwarranted access or disclosure. From a legal/administrative perspective, this can include:

- Personal information
- Personal health information
- Research data about humans (i.e., data that are subject to TCPS2)
- Educational records
- Customer records
- Financial information
- Criminal information
- Confidential personnel information
- Information that is deemed to be confidential
- Information entrusted to a person, organization or entity with the intent that it be kept private and access be controlled or restricted
- Information that is protected by institutional policy from unauthorized access

Sensitive data may also include Indigenous data and/or Traditional Knowledge, and certain types of geographic information (e.g., detailed locations of endangered ecosystems or species).

(Adapted from Sensitive Data Expert Group. (2020). Sensitive Data Toolkit for Researchers Part 1: Glossary of Terms for Sensitive Data used for Research Purposes. Zenodo.

https://doi.org/10.5281/zenodo.4088946 ; and from presentations by Victoria Smith (Privacy, Policy and Sensitive Data Coordinator, Digital Research Alliance of Canada))[5]

Appendix A: Current policies, standards, and schedules with a relation to RDM

The following University of Calgary policies, standards, and schedules have some relation to research data management:

- Acceptable Use of Electronic Resources and Information Policy
- Information Asset Management Policy
- Intellectual Property Policy
- Information Security Classification Standard
- Master Records Retention Schedule (MaRRS): Research Involving Human Subjects
- Master Records Retention Schedule (MaRRS): Research Not Involving Human Subjects
- Research Integrity Policy
- Storage of Inactive Clinical Research Records Policy

There are other policies and systems which might be expected to relate to RDM in some way, but in fact do not as they stand currently: the Privacy Policy and the University Classification System (UCLASS). Research data is excluded from the Freedom of Information and Protection of Privacy Act (FOIPP; see section 4(1)(i)) of the act), and thus does not fall under the Privacy Policy. UCLASS applies only to Business Information Assets, as indicated in section 4.9 of the Information Asset Management Policy; Scholarly Information Assets are instead identified and classified "in accordance with discipline-specific taxonomies and the Information Security Classification Standard".

---

[5] Within the University of Calgary policy framework that exists at the time of drafting of this strategy, there is no specific definition of "sensitive data". These types of data would generally be classed as Level 3 (Confidential) or Level 4 (Restricted) information assets, according to the Information Security Classification Standard. The broad RDM community in Canada generally understands the widely-used term "sensitive data" in the manner indicated in the proposed definition.