

# UBC Research Data Management Strategy

*Draft released September 1, 2022*

*UBC RDM Strategy Steering Committee*

DRAFT



# UBC Research Data Management Strategy - DRAFT

## Contents

UBC Research Data Management Strategy	1
<i>Draft released September 1, 2022</i>	1
<i>UBC RDM Strategy Steering Committee</i>	1
UBC Research Data Management Strategy - DRAFT	2
Background	3
Strategy overview	5
Goal 1: Develop institutional research data management governance, policies, and standards	6
Goal 2: Improve access to data repositories or other platforms	7
Goal 3: Provide tools and expertise for data management planning	8
Goal 4: Enhance data stewardship and research data management practices	9
Goal 5: Enhance UBC research data management expertise through strategic hiring and training	10
Goal 6: Evaluate strategic partnership opportunities	11
Appendix A – RISE model	12
Appendix B – RISE Assessment Consultation Process	13

# Background

In 2017, a Research Data Management (RDM) Task Force was convened at UBC to establish a roadmap and begin to address the Tri-Agency Statement of Principles on Digital Data Management. In January 2018 the Task Force was joined by representatives of the Tri-Agencies and Portage to discuss progress on the Tri-Agency Policy on Research Data Management; a draft of that policy was released for consultation in May of that same year. In anticipation of the release of the final version of the Tri-Agency Policy on Research Data Management, UBC commenced a project to create an institutional strategy for Research Data Management. The project was co-sponsored by the office of the Vice-President, Research and Innovation (VPRI); UBC Library; and Office of the Chief Information Officer (CIO). Recognizing that elements of an institutional RDM strategy are distributed across the institution and include both technical and human infrastructure guided by effective engagement with relevant stakeholders, the project steering committee includes members from across the institution.

## Methodology & Process

To assist in evaluating and planning, UBC selected to adapt the Research Infrastructure Self-Evaluation (RISE) tool, developed by the Digital Curation Centre in the United Kingdom. The RISE framework assists institutions to benchmark themselves in order to facilitate RDM service planning and development. The tool provides a framework for discussion, enabling relevant stakeholders to contribute their experience to all aspects of a holistically envisioned service. to reach a shared vision of where the RDM service should be. This model provides a framework to assess the current level of RDM maturity across the institution and breaks down different goals that can be prioritized for the future. Based on this model, the UBC RDM strategy attempts to document both our current and desired state in the broad scope of RDM. A working group was struck from members of the steering committee to apply the RISE model, adapt it to the UBC context, and consult with the research community. Extensive consultation was, and remains, critical to this process.

Extensive consultation with the UBC research community formed an integral component of the development of this strategy. First, when implementing the RISE model, establishing the maturity baseline and setting priorities for the future, and then once again during the drafting of this document.

The multifaceted approach to engaging with the community included town halls, focus groups, as well as sessions integrated into regular meetings of existing groups of stakeholders. Throughout the consultation period a web-based form remained available to collect community feedback.

See Appendix B for more details on the application of the RISE model at UBC.

## Key definitions

### 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.

### Research Data Management

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.

### FAIR Principles

The FAIR Data Principles (FAIR being an acronym for Findable, Accessible, Interoperable, and Reusable), are a set of guiding principles proposed by a consortium of scientists and organizations and proposed in 2016 to support the reusability of digital assets. These principles have since been adopted by research institutions, funders, and journals worldwide. The guidelines are timely as we see the unprecedented volume, complexity, and speed in creation of data.

# Strategy overview

UBC recognizes the importance of data throughout the research lifecycle and is committed to fostering excellence in Research Data Management (RDM). Research data is integral to all manner of research across all disciplines. This strategy looks at the future of research data management for all research data collected, managed, stored and accessed during any stage of the data lifecycle, by all disciplines, system wide. While it has been developed to ensure compliance with the Tri-Agency Research Data Management Policy, this strategy aims to define and prioritize goals as articulated by the UBC community with a 5-year outlook following the RISE framework.

## Scope

The purpose of the UBC Institutional RDM Strategy is to foster a culture where consideration for research data and its management is incorporated throughout research. The strategy will also support UBC researchers in adopting responsible and sustainable RDM practices and following the FAIR Principles to make research data findable, accessible, interoperable and reusable. This strategy does not propose the creation of new or altering of existing UBC policies nor is it a strategic plan. Rather, this strategy outlines areas of focus that will allow the institution to mature its practices and support for RDM in a coordinated manner.

While recognizing that research frequently involves collaboration outside the university, this strategy represents UBC's institutional approach to RDM. Other institutions will have different priorities and areas of focus.

As progress is made advancing the goals as outlined in this strategy, it is expected that re-evaluation of RDM maturity will be required to adapt and refocus the strategy over time. The research landscape, resources and priorities will inevitably change and this strategy must change with them. It is therefore also a principal component of this strategy that it be considered a living document and should be regularly reviewed and revised.

## Priorities for Action

Following extensive consultation with the UBC research community, a series of 20 recommendations were identified regarding the future of research data management at UBC, organized across six areas of focus, or goals. Of the 20 recommendations, the strategy working group wishes to focus on six recommendations as core to the strategy's success.

Namely, that UBC should:

- Form an implementation committee to review this strategy, prioritize and establish a plan to address these recommendations.
- Investigate support options for researchers needing to create data management plans.
- Establish a unified RDM portal that provides a trusted single source of information across all areas of the institution.
- Establish a unified communication strategy across research data management service providers that is cohesive and consistent.
- Explore centralized options to support data management plans, data security services, and research data repositories.
- Seek continued access to appropriate storage throughout the data lifecycle for researchers, considering the unique challenges of different classifications of data.

# Goal 1: Develop institutional research data management governance, policies, and standards

Existing UBC policies including other governance and standard documentation address many of the aspects of RDM across the research ecosystem. While UBC may wish to consider implementing a specific overarching RDM policy in the future, this was not identified as a priority by the community at this time. Meaningful progress could be reached through implementation of standards and guidance documentation. To add clarity and direction to the institutional governance framework, development in this area should consider the Indigenous Strategic Plan, the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP), and British Columbia's Declaration on the Rights of Indigenous Peoples Act (DRIPA) and Indigenous data as priorities.

Despite the presence of existing governance documents, awareness of them could be improved. In concert with centralized training and awareness programs outlined in subsequent sections, attention should be paid to raising awareness in the community about existing governance structures and policies. As new artifacts are developed, they should be included in these awareness programs.

## Recommended Actions

1. Develop governance documentation that supports good RDM Practices, while prioritizing research involving Indigenous data and implementation of the DRIPA.
2. Establish a program to raise awareness in the UBC research community regarding governance principles, documentation, and resources associated with RDM.

## Goal 2: Improve access to data repositories or other platforms

Increases in the size and diversity of research datasets in recent years has also raised demands for additional capacity of the continuum of data storage throughout the data lifecycle — active, repository and archival. To support the RDM needs of all researchers, adequate and appropriate storage must be available. UBC has already committed significant investment into active storage and high-performance computing (HPC) through the Advanced Research Computing unit to complement limited national resources (namely through the Digital Research Alliance of Canada, DRAC). However, UBC's one-time capital investment will reach end-of-life before the conclusion of some current research projects but there is a lack of clarity regarding continuity of resources beyond the investment's time period. In addition, storage of sensitive data, including health data, poses additional challenges since it is not currently supported by national resources.

Research that involves partnerships with Indigenous communities also carries an additional responsibility to ensure that principles such as Ownership Control Access and Possession (OCAP™), UNDRIP, DRIPA, and other place-based, community-specific protocols are considered and supported. When considering all of these aspects of RDM, a strong and clear overlap with the research ethics process was revealed. The feedback from the UBC research community continually emphasized the importance of involving and coordinating with research ethics processes.

Data deposit at the conclusion of a research project frequently requires similar storage capacity compared to active storage. Researchers face similar challenges when it comes to data sensitivity and long-term storage. UBC researchers have access to the UBC Dataverse Collection for small-to-medium size, low-risk datasets and the national Federated Research Data Repository (FRDR) platform for large data sets. However, increased demand for data management and especially data curation requires institutional resources beyond that which are currently available and only the current scale of data deposits can be supported. Moreover, neither Dataverse nor the FRDR platform currently support the storage of High or Very High-Risk information such as identifiable health data.

### Recommended Actions

3. Seek continued access to appropriate storage throughout the data lifecycle for researchers, while considering the unique challenges of different classifications of data.
4. Foster an environment of learning about RDM including processes, practices and documentation, especially to support the proactive integration of Indigenous data protocols and other sensitive data considerations into RDM.
5. Increase institutional capacity for repositories' data curation and metadata practices that are consistent and provide enhanced discovery and reuse for research datasets.
6. Continue to support and promote access to existing national resources.

## Goal 3: Provide tools and expertise for data management planning

Data management plans (DMPs) are living documents that describe the practices, processes and strategies of managing data throughout the life cycle of their research. UBC Library has been supporting DMPs for the last decade with hundreds of researchers using the service. However, the community's feedback indicated a strong desire and need for guidance, templates and exemplars that are discipline and funder-specific. The community's feedback also indicated the need to reduce administrative burden by aligning the language used in UBC's Research Information System (RISe) and DMPs. With this type of central, unified and ongoing support, DMPs could assist researchers to proactively establish how they will manage their data through all stages of a research project and beyond.

### Recommended Actions

7. Focus on providing and facilitating access to discipline- and funder-appropriate data management planning software, language, templates, and exemplars.
8. Investigate the potential for unified and coordinated services to support data management planning.
9. Coordinate DMPs and ethics applications to reduce administrative burden on researchers.

## Goal 4: Enhance data stewardship and research data management practices

Existing efforts to support UBC researchers' RDM needs have been provided through many UBC departments. Some academic disciplines have well established practices in support of RDM where others have few resources and limited experience. These efforts and practices, to date, have been largely uncoordinated and independent of each other. The UBC community expressed a clear desire to strive for a more coordinated support structure, in the form of a centralized resource. The centralized resource should be sufficient to support those with limited existing resources while not interfering with the existing practices of the more established areas. The central resource should include training and documentation related to all aspects of RDM and should aim to limit how often individuals need to seek specialized information from those with differing expertise. Rather, resources that in the past might have been split between different UBC departments should be combined to reduce overlap and provide a more complete RDM picture.

Early-career researchers and trainees expressed the strongest desire for RDM support and training. This cohort specifically identified the need to better understand available resources given their direct involvement with many of the tasks associated with data management.

When considering how to raise awareness of these resources and how best to deliver training, UBC should look to existing programs and curriculum. Attempts should be made to build RDM into existing programs and practices.

### Recommended Actions

10. Establish a unified RDM hub that provides a single, trusted source of information for all areas of the institution.
11. Develop RDM training and awareness materials and programs that utilize different formats and media.
12. Enhance and develop in-person RDM training and awareness workshops and presentations.
13. Both online and in-person training and awareness programs should consist of unified material from all RDM related service providers (e.g.: Ethics, Library, IT, IRSI, Legal, and ARC) to ensure participants do not need to attend separate sessions related to each area.
14. Strive to embed RDM training and awareness into existing programs starting at the undergraduate level. This includes developing materials that speak to the nuanced requirements of different disciplines and departments.

# Goal 5: Enhance UBC research data management expertise through strategic hiring and training

UBC holds a wealth of RDM expertise and resources across its campuses and within various units and departments. This includes resources and expertise at UBC Advanced Research Computing, UBC Library, Faculties, research institutes, Research Ethics, IT, the Indigenous Research Support Initiative, and more. The community feedback indicated researchers' desire for a more coordinated, unified approach to RDM support and communications. Moreover, research data management requires constant skill upgrading to keep up with this dynamic field.

## Recommended Actions

15. Establish a unified communication strategy across research data management service providers that is cohesive and consistent.
16. Ensure appropriate staffing for centralized research data management services such as data management plans, data security services, and research data repositories.
17. Encourage the development of skills and knowledge of research data management practices and tools among staff who support and conduct research.

## Goal 6: Evaluate strategic partnership opportunities

UBC has established involvement with a variety of partners across many research contexts such as the DRAC, CANARIE, U15, BCNET, the Research Universities Council of BC, and countless discipline- and topic-specific committees, organizations, and collaborations. In many cases, UBC has taken a leadership role and contributes at a variety of levels across the research ecosystem in others. In conjunction with these groups, UBC also has regular, direct contact with the major funding organizations in BC, Canada, and internationally including the Tri-Agencies, CFI, BCKDF, NIH etc.

Partnerships such as these shape the research landscape across Canada. Key partners also provide important resources for UBC researchers that augment what is available institutionally.

### Recommended Actions

18. Advocate for changes to funding models to better support research that partners with Indigenous communities.
19. Maintain and evolve existing partnerships with the Digital Research Alliance of Canada.
20. Seek new partnership opportunities with Indigenous communities and related support organizations.

# Appendix A – RISE model

(to be populated based on the content currently published)

# **Appendix B – RISE Assessment Consultation Process**

(to be populated based on the content currently available)