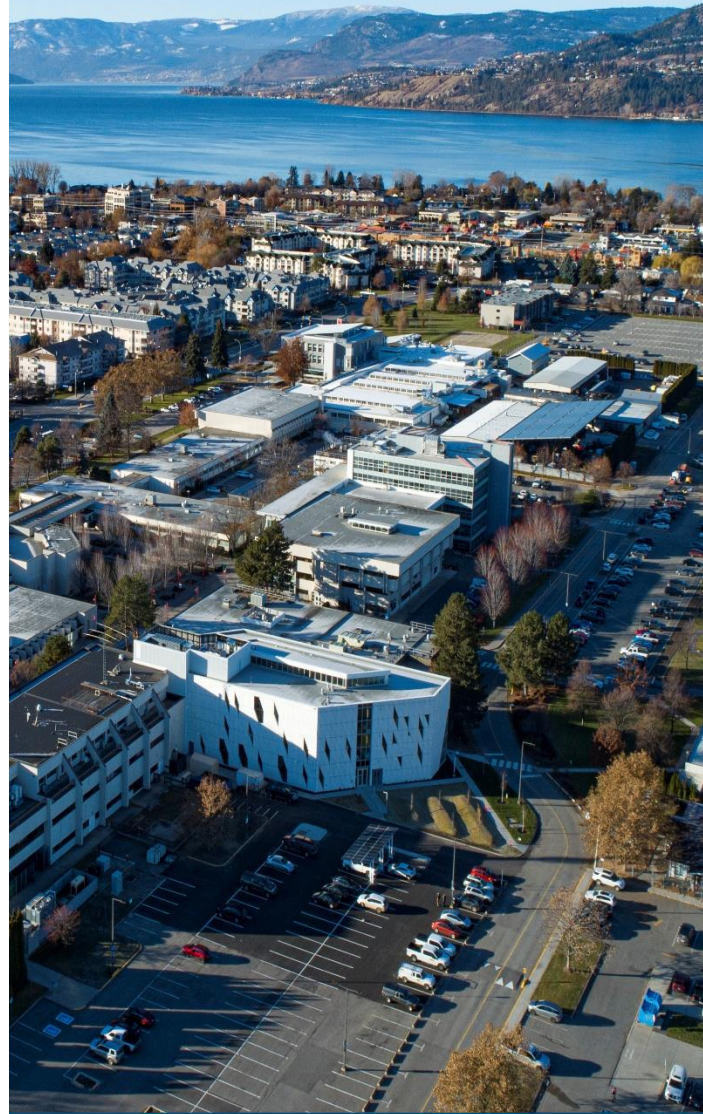


Okanagan College Research Data Management Strategy



Background

The Okanagan College Research Data Management (RDM) Strategy is developed at the convergence of two initiatives: 1. to meet the Tri-Agency Research Data Management Policy that requires all institutions seeking research funding from the federal agencies to have a research data management strategy by March 1, 2023; and 2. to fulfill the e-Okanagan College's Strategic Plan to have a research data management plan that guides RDM practices for all researchers at Okanagan College.

Our initiatives reflect the growing recognition in the broader Canadian and International research landscape that data, as the basis of research findings and claims, is an important research output. RDM concerns the proper collection, documentation, storage, sharing, and preservation of research data, while supporting research efficiency, transparency, discoverability, and collaboration. To ensure best practices at all stages of the research lifecycle, RDM strategies will gradually be promoted and adopted as direction from the Tri-Agency evolves. Given this context, Okanagan College is creating an RDM Strategy to align its priorities with such development. Furthermore, Okanagan College will prioritize supporting researchers in learning and integrating RDM practices into their current and future research endeavors.

Scope & Context

The Strategy articulates Okanagan College's commitment to RDM by identifying priority areas to raise its RDM capacity over the next five years. The Strategy and related policies and support will apply to all research activities conducted by Okanagan College staff, students, and research contractors including but not limited to Tri-Agency-funded research.

While the Strategy is not intended to create or implement policy, Okanagan College will develop RDM policies in the future to provide Okanagan College researchers and community research partners with clear guidance and specific requirements surrounding RDM practices. A separate implementation plan will embody our Strategy and include information on allocation of resources, responsibilities of different units, standard practices to be followed, and outline steps forward.

Okanagan College recognizes that its approach to RDM resides within a broader research and academic culture, specific to our institution. This Strategy therefore reflects and respect Okanagan College's mission, vision, values, responsibilities, commitments, and priorities, including, but not limited to:

Applied Research: RDM practices must be compatible with contractual commitments with industry partners that may restrict the public dissemination of data.

Indigenous Data Sovereignty: For research conducted by and with First Nations, Métis and Inuit communities, data management practices must be co-developed with these communities to ensure the practices are in accordance with the communities rights to control their own data, as reflected by the principles of Ownership, Control, Access and Possession (OCAP) and Collective Benefit, Authority to Control, Responsibility, and Ethics (CARE).

Sensitive Data: The collection, analysis, use and reuse of sensitive data, especially data pertaining to individuals, must take place within a framework that is consistent with institutional research ethics principles, as well as with Okanagan College's commitment to Equity, Diversity, Inclusion, and Social Justice.



Oversight and Review

This Strategy (Version 1.0) was developed by the Research Data Management Strategy Project Team led by the Director of Learning and Applied Research (LAR) and included representatives from LAR, IT Services, and Library Services. The Project Team assessed existing RDM capacity using the Maturity Assessment Model in Canada (MAMIC). MAMIC is a Canadian-specific benchmark tool that was developed by researchers at Carleton University, the University of Toronto, Memorial University, and the University of Victoria, based on the Research Infrastructure Self-Evaluation (RISE). RISE is a benchmarking tool developed for higher education institutions in Europe to facilitate RDM service planning and development at the institutional level. The MAMIC tool supports institutions in evaluating the current state of institutional RDM services and to generate discussion amongst various stakeholders to better understand the current state of RDM support and to inform future strategy development. MAMIC focuses on four areas of service and support: Institutional Policies and Processes, IT Infrastructure, Support Services, and Financial Support. This assessed the maturity and scale of current services.

The results of the MAMIC indicated that Okanagan College is at the beginning phase of understanding and socializing RDM practices, policies, and procedures. Okanagan College's RDM strengths from the MAMIC show that there is availability of infrastructure and associated services covering identity management and verification, with mediated access to research data according to security needs. Areas identified requiring further development include devising adequate policies and procedures,

offering archival storage, and creating general support services. Okanagan College will continue to work with stakeholders to develop an RDM plan that reflects the needs of our researchers and community.

The MAMIC assessment was followed with a preliminary survey with Okanagan College researchers and stakeholders. Some participants from the survey group also participated in a focus group led by the Project Team. Core questions were posed to participants that elicited helpful information on topics such as defining RDM, scope of Okanagan College researchers, types of data requiring support, barriers to implementing RDM practices, Indigenous research considerations, that has informed Okanagan College's Strategy.

This Strategy is intended to be a living and evolving document with ongoing oversight and revision. Under the guidance of Learning and Applied Research, the Strategy will be reviewed and updated annually for the first three years and every two years thereafter.

Strategy Areas

1. Developing and maintaining RDM policies, guidelines, and procedures
 - Elements of RDM practices currently exist in the Research Ethics Board (REB) Application for Ethical Review process (https://www.okanagan.bc.ca/sites/default/files/2020-06/form_1_-_reb_application_form.pdf, Data mentioning: 31, 46 (p), 47(o), Data section: 36 – 42).
 - In addition to promoting the Tri-Agency's RDM Policy, Okanagan Collage is responsible for developing its internal data management policies, guidelines, and procedures in accordance with RDM principles and practices.
 - Tentative timeline: March to December 2023.
2. As a priority focus, Okanagan College will begin to conceptualize and formulate RDM policies to address RDM standards, data storage options, data management plans (DMP) and data deposit requirements in the next two years, aiming to launch these policies by 2025.
 - Establish policies and guiding documents in support of researchers working with Indigenous partners to conduct ethical research that respects Indigenous data sovereignty and benefits the community.
 - Develop policies and procedures regarding data security measures and protection of proprietary or sensitive data.
 - Tentative timeline: March 2023 to 2025.

3. Providing active storage and access to repositories

- Provide related training for researchers and their students.
- Support obtaining and developing data depositories.
- Provide guidance on RDM and data depository security concerns.
- Tentative timeline: January 2024.

4. Raising awareness

- Promote RDM principles such as FAIR, CARE, and First Nation Principles of OCAP.
- Raise awareness of the three pillars of Tri-Agency's RDM requirement (Institutional RDM strategies, data management plans, and data deposits).
- Ensure that researchers can easily locate and access usable, relevant, and current RDM-related resources.
- Tentative timeline: Ongoing, beginning May 2023.

5. Providing support services

- Support researchers to develop RDM knowledge and skills.
- Provide RDM training and consultation.
- Tentative timeline: Ongoing, beginning May 2023.



Definitions

The following definitions and others are taken from the Committee on Data of the International Science Council.

Research data

Data are used as primary sources to support technical or scientific enquiry, research, scholarship, or artistic activity, and 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. All other digital and non-digital content have the potential of becoming research data. Research data may be experimental data, observational data, operational data, third party data, public sector data, monitoring data, processed data, or repurposed data.

Research data management

Data Management refers to the storage, access and preservation of data produced from defining the research question to the given investigation. Data management practices cover the entire lifecycle of the data, from formulating the research question to planning the investigation to conducting it, and from backing up data as it is created and used for long term preservation of data deliverables after the research investigation has concluded. Specific activities and issues that fall within the category of data management include: file naming (the proper way to name computer files); data quality control and quality assurance; data access; data documentation (including levels of uncertainty); metadata creation and controlled vocabularies; data storage; data archiving and preservation; data sharing and reuse; data integrity; data security; data privacy; data rights; notebook protocols (lab or field).

Data management plan

A data management plan is a formal statement describing how research data will be managed and documented throughout a research project, including the terms regarding the subsequent deposit of the data and a data repository for long-term management and preservation.

Repository

Repositories preserve, manage, and provide access to many types of digital materials in a variety of formats. Materials in online repositories are curated to enable search, discovery, and reuse. It is integral that there be sufficient control in order for the digital material to be authentic, reliable, accessible and usable on a continuing basis.

Open data

Open data encompasses structured data that are accessible, machine-readable, usable, intelligible, and freely shared. Open data can be freely used, re-used, built upon, and redistributed by all – subject only, at most, to the requirement to attribute and share.

(Access more Research Data Management Terminology curated by the International Science Council's Committee on Data at <https://codata.org/initiatives/data-science-and-stewardship/rdm-terminology-wg/rdm-terminology/>)



The Research Data Lifecycle



Plan



Create



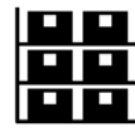
Process



Analyze



Disseminate



Preserve



Reuse

Plan

Identify the data that will be collected and plan for data management throughout the lifecycle. Create a data management plan, which may be required by funding agencies.

Create

Carry out your research. Document the data collection methods and instruments necessary to interpret and use the data.

Process

Process the data that has been collected. This could involve cleaning data, combining data collected from multiple sources, converting data from one format to another, and validating data. These processes need to be documented so results could be replicated from the raw data.

Analyze

Interrogate the data to produce the findings of your research. Document the instruments and methods used for analysis. Any code written to analyze or visualize data may need to be preserved and made available in support of results.

Disseminate

Indicate where and by what terms data can be accessed. Consider using a data repository to enable discovery of your data. Funders may require that data be deposited in a publicly accessible repository. Restrictions on access may be imposed where data are of a sensitive or confidential nature.

Preserve

Prepare data for archiving in a suitable location, locally and/or in a secure repository. This stage may involve quality assurance, file format conversion, creation of metadata records, and licensing datasets for re-use.

Reuse

Data that are available for discovery and access may be re-used by other researchers. Your data may become the raw material for someone else's research, or may be used for other purposes, e.g. policy-making, development of commercial products, and teaching.

Content adapted from Portage Network, *Primer – Research Data Management* (CC BY-NC 4.0) <https://zenodo.org/record/4000999#.Y20IE3bMKUI>, from Portage Network, *Brief Guide – Data Management Plan* (CC BY-NC 4.0) <https://zenodo.org/record/4495482#.Y21PEHbMKUk>, and from University of Reading, *The Research Data Lifecycle* (Used with permission) <https://www.reading.ac.uk/research-services/research-data-management/about-research-data-management/the-research-data-lifecycle>



Resources

1. Government of Canada (2021). Tri-Agency Research Data Management Policy. <https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management/tri-agency-research-data-management-policy>
2. Government of Canada (2016). Tri-Agency Statement of Principles on Digital Data Management <https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management/tri-agency-statement-principles-digital-data-management>
3. Research Data Management Terminology. <https://codata.org/initiatives/data-science-and-stewardship/rdm-terminology-wg/rdm-terminology/>