# Institutional Strategy – Mohawk College Research Data Management

## Background & Purpose

In response to the Tri-Agency Research Data Management (RDM) Policy, released by the Canadian Tri-Agencies (NSERC, SSHRC and CIHR) in 2021, all postsecondary institutions and research hospitals eligible to receive Tri-Agency funds were asked to develop and post an Institutional Strategy on Research Data Management.

The Tri-Agency RDM Policy was established to support research excellence in Canada by promoting sound data management practices. The policy specifically requires the following:

1. The College post a Research Data Management Institutional Strategy on their website;
2. The College work to develop Data Management Plans for research projects (certain research grants will require a data management plan to be in place); and
3. The final phase would be the development of infrastructure for data deposit.

Mohawk College recognizes that this Institutional Strategy for RDM is a living document and is committed to a process of continuous improvement and refinement of the Strategy and associated activities. Complete development and implementation will continue over the next 5 years as awareness increases and processes are put in place to support the infrastructure and capacity required to enable this work.

## Importance of Research Data Management

The Tri-Agency RDM Policy is intended to ensure that its grant recipients are effectively and responsibly managing the data involved during the lifecycle of a research project. It is intended to guide the collection, documentation, storage, sharing, and preservation of research data.

Effective RDM is a key element of the research lifecycle that allows researchers to better access, reuse, preserve, and share data. In recent years, an increased emphasis on concepts such as reproducibility, scientific rigour, and collaboration has made proper RDM of the utmost importance.

Good Research Data Management can be summed up as making research data FAIR (*Findable, Accessible, Interoperable, and* Reusable) a set of principles designed to ensure that data are stewarded in a way that enables and enhances reuse by humans and machines. Wilkinson, et al. The Fair Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

## Vision/Guiding Principles

This strategy document is intended to be used as a roadmap to support and enable good research data management practices at Mohawk College. It is not intended to be a policy document that must be adhered to for all research data collection. As such, this is a living document which will be updated and revised over the years to best implement a sound RDM strategy at the College.

In defining our Institutional Strategy for Research Data Management, Mohawk College intends to support research activities and encourage sound research data management stewardship practices. This includes providing an environment which ensures that research data is accessible, usable, safe, and trusted. This applies to research data pertaining to Tri-Council funded grants and those that require a research data management plan. Excluded from this RDM institutional strategy is data collected as part of an applied research project with a community partner in which the intellectual property resides with either the partner solely or is shared between the partner and the College.

## Scope

This Institutional RDM Strategy will apply to all Mohawk College researchers, including staff, students, faculty, and administrators from across all campuses. Our initial focus will be on the

development of RDM governance and practices for all Tri-Agency funded research requiring an RDM plan.  Areas of the College such as Institutional Research and data collection for student services, program data, salary information do not form part of this strategy. Since most of the research conducted at Mohawk College is applied research that is industry and/or community partner driven, much of the data developed at Mohawk College would be exempt from this strategy due to legal and intellectual property constraints and practices. However, we recognize that data is an important research output, and we are committed to developing best practices for all data developed through our research.

It is important to note that researchers under the Tri-Agency data management framework are not required to share their data and the policy specifically states that it is "not an open data policy," nonetheless, there is an expectation that researchers will "provide appropriate access to the data where ethical, cultural, legal, and commercial requirements allow."

## Governance

The Institutional Strategy for Research Data Management will fall under the Research Department at Mohawk College 'Ideaworks' led by the Dean, Applied Research but with input from the key stakeholders identified below to inform the continued development of the RDM plans. A subset of the list of stakeholders below forms part of the Research Data Management Steering Committee.

## Stakeholders

There are multiple stakeholders involved in developing the Research Data Management Plan for this phase (Institutional Strategy development) and the future phases outlined in the Background section above. At present the list of stakeholders includes:

- Ideaworks – Applied Research Department of Mohawk College
- Mohawk College Library
- Centre for Teaching & Learning Innovation
- Mohawk College Information Technology (IT), including Information Governance & Information Security

- General Counsel's Office (risk management, legal and privacy)
- Mohawk College Research Ethics Board
- Centre for Indigenous Research, Knowledge & Education (CIRKL)

The list of stakeholders is expected to evolve and grow as the RDM strategy evolves.

## Research Ethics Considerations

Mohawk College is dedicated to upholding all legal and ethical responsibilities while conducting research involving human participants. Our Mohawk College Research Ethics Board (REB) is comprised of faculty, staff and community partners, and conducts independent reviews of research protocols. All research proposals involving human participants are approved, conditionally endorsed, or rejected based on potential risks to participants' well-being. The Research Ethics Board operates within the guidelines set forth by the Tri-Agency Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2 (2018) and the Ethical Conduct of Research Involving Humans Policy: https://www.mohawkcollege.ca/about-mohawk/leadership-and-administration/policies-and-procedures/corporate-policies-and/ethical-0.

Mohawk College will commit to ensuring that REB is aware of RDM requirements and can ensure the continuance of the REB's high standards

## Indigenous Data Considerations

Mohawk College will support its researchers throughout the RDM process, placing a significant emphasis on respecting Indigenous data sovereignty whenever applicable. This commitment involves leveraging existing policies of Mohawk College and working closely with CIRKL. Mohawk College will act in accordance with the Fair Principles, the First Nations principles of Ownership, Control, Access, and Possession (OCAP), and the UN Declaration on the Rights of Indigenous Peoples. Ideaworks, the Applied Research Division of the College ensures that no research involving Indigenous communities is undertaken without the knowledge and collaborative support of the CIRKL group. CIRKL offers guidance to researchers on applications

focused on Indigenous-related research and they will form part of the ongoing group of stakeholders supporting and working on Research Data Management processes.

## Priorities / Action Plans

1. Enhance awareness of RDM across Mohawk College, by:

- Implementing communication plans; and
- Providing training, resources, and education on research data management throughout the stages of data-management lifecycle.

2. Identify the ideal state for RDM and work to implement the necessary processes for compliance, by:

- Identifying current practices in use for data storage and data management;
- Identifying gaps and develop processes to address these gaps;
- Creating a timeline to identify milestones and necessary next steps in implementation; and
- Working to identify additional resources required for data collection and storage.

3. Provide support to researchers at Mohawk College to encourage best practices, by

- Providing training resources to increase the knowledge of RDM as it relates to ongoing research initiatives;
- Ensuring researchers have access and are aware of services and resources that may be required at various stages of the research project; and
- Enabling REB and OCAP training support to researchers.

4. Committing to continuous improvement and best practices, by:

- Investigating best industry resources available to support RDM;
- Working with stakeholders to identify any technical supports, tools, storage options required; and
- Continuing to reassess progress in RDM implementation over the next 5 years.

## Definitions

The definitions below are as referenced in the Tri-Agency RDM Policy FAQ

**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.

**Data Management Plan**: 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 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.

**Metadata**: "Metadata" are data that define and describe the characteristics of other data. Accurate and relevant metadata are essential for making research data findable. A principle to help determine what information should be included in metadata is the open archival information system model criterion that the information be "independently understandable." "Independently understandable" means enough information has been provided in the metadata for someone else to be able to understand the data set without needing its creator explain it.

There are many metadata standards (often referred to as "schemas") prescribing how to treat metadata, and they vary across disciplines. However, metadata states who created the data and when, and includes information on how the data were created, their quality, accuracy and precision, and other features necessary to enable discovery, understanding and reuse.

**Research:** Research is an undertaking intended to extend knowledge through a disciplined inquiry and/or systematic investigation. ([Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS2](#))

**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.

**Research Data Lifecycle:** The data lifecycle includes phases from data creation, processing, analysis, preservation, storage, and access, to sharing and reuse (where appropriate), at which point the cycle begins again.

**Research Data Management:** Research data management (RDM) refers to the processes applied through the lifecycle of a research project to guide the collection, documentation, storage, sharing and preservation of research data.

RDM is essential throughout the data lifecycle. Data management should be practiced over the entire lifecycle of the data, including planning the investigation, conducting the research, backing up data as it is created and used, disseminating data, and preserving data for the long term after the research investigation has concluded.

**Sensitive data:** Classified, usually private, information that must be protected and is inaccessible to outside parties unless specifically granted permission.

**Applied research:** Any original investigation, undertaken in conjunction with a community partner, to acquire new knowledge, or to apply existing knowledge in a novel way, to solve a practical problem directed primarily towards a specific practical aim or objective.

## Relevant Policies

Academic Freedom in Research Policy

Acceptable Employee Use of IT Resources Policy

Administration of Research Grants and Contracts Policy

Ethical Conduct of Research Involving Animals Policy (FOR NOW)

Ethical Conduct of Research Involving Humans Policy

Information Governance and Security Policy

Integrity in Research and Scholarship Policy

Research Involving Biohazardous and Radioactive Materials Policy

Mohawk College Privacy and Legal Statements