Title: Research Data Management Institutional Strategy	
Date First Created: March 2023	Date Reviewed:
Approved By: (Terminal Committee) Baycrest Academy for Research and Education Executive Team	Editor: Jean Lazarus

1.0   Policy Statement

The commitment to research excellence is a principle that is embodied in all research conducted across the Baycrest entities. The Institutional Research Data Management Strategy is a requirement of the Canadian Tri-Agencies. The strategy describes the environment that will enable and support Research Data Management practices and follows the FAIR Principles to make research data **f**indable, **a**ccessible, **i**nteroperable and **r**eusable (https://www.go-fair.org/fair-principles/).

As a member of the Toronto Academic Health Sciences Network (TAHSN) community, the Academy's Institutional Research Data Management Strategy is in alignment with the Research Data Management Principles shared by the University of Toronto and its affiliated institutions (https://research.utoronto.ca/engaging-research/university-toronto-institutional-research-data-management-strategy).

These Principles are:

- Promote Research Integrity and Excellence
- Recognize the Value of Data
- Encourage the Implementation of Data Management Plans
- Facilitate Long-Term Access Through Data Deposit
- Reflect Institutional Practices and Standards
- Honour Indigenous Community-Driven Principles
- Foster a Culture of Inclusive Representation and Public Trust
- Observe Jurisdiction and Legalities
- Strengthen Partnership and Collaboration
- Mitigate Risk Related to Confidential Data
- Safeguard Sensitive Data
- Integrate Excellence in all Disciplinary Approaches
- Connect through Communication and Engagement Opportunities

- Provide Infrastructure that Supports Diverse and Complex Programs of Research
- Ensure Support Services are Available
- Commit to Advocacy and Support for Researchers' Needs

Baycrest Academy for Research and Education ("Academy") will develop tools, and provide the support and resources to enable all staff who are conducting research to manage their research data to the highest levels of excellence throughout the research data lifecycle. The Academy will incorporate resources provided by partners, such as the University of Toronto, and prepared by other institutions, such as the Digital Research Alliance of Canada.

2.0 Definitions and Abbreviations

"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 (Government of Canada, FAQ Tri-Agency Research Data Management Policy).

"Data deposits" – Refers to when research data collected as part of a research project are transferred to a research data repository

"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 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. (Government of Canada, FAQ Tri-Agency Research Data Management Policy).

"Data sharing" – includes deposit of data in external repositories or sharing of data with external researchers.

"External" – refers to entities (researchers, repositories) outside of Baycrest entities.

"Metadata" - "Metadata" are data about data—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 greatly across disciplines. However, metadata generally state who created the data and when, and include information on how the data were created, their quality, accuracy and precision, and other features necessary to enable discovery, understanding and reuse. (Government of Canada, FAQ Tri-Agency Research Data Management Policy).

"Project" – refers to any research project, study, experiment, clinical trial, etc. It can be primary or secondary research.

"Publications" – refers to dissemination of results via means including (but not limited to) preprints (Author's original/submitted manuscript), postprints (Author's accepted manuscript), and journal publications.

"Repositories" – refers to any online digital platform designed to store research data. The repository enables the sharing, when appropriate or requested, of the data.

"Researcher" – are scientists, research staff, students, and staff in Baycrest entities, as well as collaborators working on projects approved by Baycrest Academy for Research and Education's Research Ethics Board.

"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. (Government of Canada, FAQ Tri-Agency Research Data Management Policy).

"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—from data creation, processing, analysis, preservation, storage and access, to sharing and reuse (where appropriate), at which point the cycle begins again. 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. (Government of Canada, FAQ Tri-Agency Research Data Management Policy).

3.0 Background and Scope

This policy addresses the Institutional Strategy for RDM. The RDM institutional strategy complements policies such as, the Research Data and Record Retention policy. It is also informed by policies such as, Principles and Responsibilities Regarding the Conduct of Research. This policy provides researchers with the structure to develop their data management plans for their research projects and program and outlines expectations regarding securely preserving and using their research data throughout the research data lifecycle, reusing data over the course of their careers, and when appropriate, sharing their data.

The RDM Strategy is relevant to researchers in all Baycrest entities. The Academy will take reasonable measures to ensure that researchers are made aware of this strategy and kept informed of changes.

4.0 Strategy Goals and Objectives

The goals and objectives were developed using the RDM Maturity Assessment Model in Canada (MAMIC) Version 1.0 (https://zenodo.org/record/5745493#.Y_3hWnZOIPY). The MAMIC tool was used to determine the Academy's state of readiness for developing and implementing RDM plans. The MAMIC focuses on four key areas: Institutional Policies and Processes, IT Infrastructure, Support Services and Financial Support. The goals and objectives reflect the areas of strength and gaps that will be addressed through the implementation of the Institutional Strategy.

Goal #1: Increase Awareness

- **Objective 1.1**: Promote the Academy's Institutional RDM Strategy to research community members
- **Objective 1.2**: Provide research community members with DMP resources.
- **Objective 1.3**: Survey researchers to assess data management practices. Periodic follow-up surveys will be conducted to assess whether baseline practices are changing.

- **Objective 1.4**: Encourage researchers to start preparing robust RDM practices by leveraging existing resources provided in Appendix I

Goal #2: Expand RDM Support and Training

- **Objective 2.1**: Monitor RDM training programs for researchers and students provided by local or national organizations and encourage researchers to attend.
- **Objective 2.2**: Work with key stakeholders (e.g., scientists, ethics, IT services, procurement, and legal) to improve the following core RDM service domains:
  - **Objective 2.2.1**: Develop procedures for how researchers should address Data management plans (DMPs)
  - **Objective 2.2.2**: Ensure that policies and procedures address security and risk assessment related to research data, data sharing, data deposit, and data preservation
- **Objective 2.3:** Develop a core-group of staff with knowledge and expertise to provide RDM-related support and services and to advise researchers in RDM best practices
- **Objective 2.4**: Assess current computing power to determine whether changes and updates are required. Develop and promote procedures on accessing and using computing power. Ensure that procedures focus on both internal computing power access and external resources, such as Compute Canada.

Goal #3: Incorporate the governance for research data in the Data Science Committee

- **Objective 3.1**: Include research data oversight as a key component of the newly formed Data Science Committee. The committee, reporting to the Rotman Research Institute Scientists Committee is comprised primarily of scientists members with representatives from key stakeholders (e.g., IT, Research Ethics, and Research Operations, etc.).
- **Objective 3.2**: A subset of the Data Science Committee will:
  - **Objective 3.2.1**: Complete an analysis of relevant policies to ensure the institutional RDM framework is in compliance with applicable laws and regulations

- **Objective 3.2.2**: Propose revisions to, and updates of, existing research data-related policies
- **Objective 3.3**: Strengthen communication and coordination with other research institutes in establishing more streamlined RDM workflows and processes
- **Objective 3.4**: Conduct a review of data repositories and archiving platforms that can be used to create and support data repository. The review will include location of the repositories, platforms that can be used, annual costs associated with the repositories, and staffing resource requirements to support the repositories and platforms.
  - **Objective 3.4.1:** Asses storage space requirements for behavioural data that is centralized across different labs; fosters integration; and, is in alignment with Open Science principles.
  - **Objective 3.4.2**: Develop a policy on data storage, length of time stored, archival principles, encryption, access, and multiple storage locations. The policy will be in alignment with privacy principles.
  - **Objective 3.4.3**: Assess the management of software packages needed for the different stages of the research life-cycle and the Research IT capacity to support different software versions.
  - **Objective 3.4.4**: Assess the use of the University of Toronto's resource, Dataverse, for public archival.
- **Objective 3.5**: Create an inventory and repository for project data.

Goal #4: Develop RDM Services through Resource development and Partnerships

- **Objective 4.1**: Conduct an analysis of RDM Services requirements to determine staffing resources to support the implementation of the Institutional Strategy in accordance with applicable privacy laws and regulations, and related activities, such as data management and data repository management.
- **Objective 4.2**: Develop a multi-year budget to ensure the long-term sustainability of institutional RDM support
- **Objective 4.3**: Foster and access provincial and national RDM infrastructure providers and existing research community platforms to address gaps in repository technologies for managing, sharing, depositing, and archiving sensitive data and large data

- **Objective 4.4**: Grow institutional RDM capacity:

  - **Objective 4.4.1**: Develop expertise in RDM amongst research support staff (e.g., grant officers and IT support staff), REB staff and members.

  - **Objective 4.4.2**: Promote integrated interoperable systems for research-related records (e.g., DMPs, REB protocols, and institutional grant management)

  - **Objective 4.4.3**: Focus on ensuring equitable, diverse, and inclusive representation in RDM-related roles

Appendix I

https://cris.utoronto.ca/dri_portal/home/

The Centre for Research & Innovation Support portal provides access to the University of Toronto's resources related to data management planning, research data storage and sharing, working with research data, and privacy and security considerations.

https://onesearch.library.utoronto.ca/sites/default/files/rdm/dmp_planning_question_guide_-_u_of_t_libraries_-_apr2021.pdf

A guide to facilitate writing the DMP using questions and guidance that are focused on each section of the plan.

https://onesearch.library.utoronto.ca/sites/default/files/rdm/dmp_planning_checklist_-_u_of_t_libraries_-_apr2021.pdf

This checklist can be used to get started with creating and DMP.

https://library.utoronto.ca/researchdata

How do I…

- Consent forms with consent to share data

- Keep my data secure

- Write a data management plan

- Encrypt my data

- Find out what funders require

- Choose an online survey tool

- Handle sensitive date

https://alliancecan.ca/en/services/research-data-management

Explains Research Data Management (RDM)

Outlines services they provide:

- FRDR (Federated Research Data Repository – national platform for researchers to discover, share and preserve Canadian research data. Helps to make data FAIR (findable, accessible, interoperable and reusable)

- DMP Assistant (national online data management planning to assist researchers in preparing data management plans (DMPs). A free service

    o You can also access DMP Assistant through this link, https://assistant.portagenetwork.ca/

- Learning & Training (tools and resources to help with data management planning, access, preservation, and discovery. Training resources include brief guides, primers, DMP templates, exemplars, institutional strategies guidance, and many more)

    o The DMP templates include discipline focused. For example, Neuroimaging in the Neurosciences, Open Science Workflows, Mixed methods (surveys and qualitative research)

- Network of Experts (critical in developing and providing resources, expert advice and practical help to assist with the management of research data at every stage of the data lifecycle)

- Publications (Zenode) (reports related to RDM)