Research Institute of the McGill University Health Centre Research Data Management Strategy

Background

The present document was formulated by the Research Institute of the McGill University Health Centre (RI-MUHC) in response to the March 2021 Tri-Agency[1] launched Research Data Management (RDM) Policy, specifically the requirement to create and publicly post an institutional RDM strategy. As a hospital-based McGill-affiliated institute, we have adopted the key elements of McGill University Strategy.

As such, our researchers draw on the tools and guidance available through the he McGill Digital Research Services Hub (DRS). This was created In 2021, as a joint effort from the office of the Vice President - Research and Innovation, the McGill Library and IT Services. The DRS currently offers a broad array of support and guidance services. In addition, based at the Centre for Outcomes Research and Evaluation (CORE) of the RI-MUHC, we have launched the Capture and Optimization of Research Data (CORD) unit and have a fully operational Biostatistical Consulting Unit (BCU). These units form core components of our local execution at the RI-MUHC of the broader McGill University RDM strategy.

The McGill RDM Strategy and our RI-MUHC adaptation of this are living documents that will be revised and modified as needs, policies, and best practices in RDM continue to evolve. The current document should be considered in combination with the McGill University RDM. The key changes relate to the conceptualization of the research cycle (Figure 1), emphasizing the types of research that occur at the RI-

[1] The Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council (NSERC) and the Social Sciences and Humanities Research Council (SSHRC)

MUHC, and the particular services offered through the RI-MUHC that complement the consultation and tools available at McGill University main campus.

What is Research Data Management (RDM)?

RDM tackles the organization of research data through the life cycle of a research project or program (please see **Appendix A of the McGill University RDM Strategy**). It involves the planning, organization, description, storage, and sharing of research data in a secure fashion. The goal is to ensure the generation of high quality data that can then be analyzed more efficiently to generate findings that can then be disseminated and, as appropriate, translated into tools, programs, policies, and other outputs.

Vision

As part of its strong commitment to research excellence, McGill and the RI-MUHC will lead in the development of tools, support, and guidance to enable researchers to manage their research data to the highest standards across the research data lifecycle. McGill and the RI-MUHC will support researchers in incorporating meaningful RDM practices and stewardship by leveraging relationships with stakeholders at the institutional, provincial, national, and international levels.

Guiding Principles of the McGill RDM Strategy adopted by the RI-MUHC

- **Research Excellence**
    - Advance impactful RDM practices as an integral part of cultivating research excellence.
- **Researcher-oriented**
    - Support all researchers towards the adoption of RDM practices by leveraging the best possible services and tools.

- Focus on reducing barriers throughout the research data lifecycle.

- **Context-based Approach**
  - Recognize that different domains have different needs.
  - Promote a flexible RDM model that is adaptable to all research domains.
  - Align the institutional approach with recognized frameworks such as the FAIR Principles[2] (Findable, Accessible, Interoperable, and Reusable).
  - Ensure that the unique rights, interests and circumstances of First Nations, the Métis Nation and Inuit are respected by adopting a distinction based RDM approach for research involving First Nations, the Métis Nation and Inuit communities and their data, such as OCAP[3] (Ownership, Control, Access, and Possession) and CARE[4] (Collective benefit, Authority to control, Responsibility, and Ethics).

Scope

The RDM Strategy is relevant to all McGill researchers, including those based at the RI-MUHC, as defined in the McGill Regulation On The Conduct Of Research. McGill University and the RI-MUHC will take reasonable measures to ensure that researchers are made aware of this strategy and kept informed of changes to it.

---

[2] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, *3*(1), 1-9. https://doi.org/10.1038/sdata.2016.18.

[3] OCAP® is a registered trademark of the First Nations Information Governance Centre (FNIGC). https://fnigc.ca/ocap-training/

[4] Research Data Alliance International Indigenous Data Sovereignty Interest Group. (September 2019). *CARE Principles for Indigenous Data Governance*. The Global Indigenous Data Alliance. https://www.gida-global.org/care

Figure 1. The research cycle at the RI-MUHC (conceptualized by Isabel Fortier and Kaberi Dasgupta)

Oversight and Review

- The Office of the Vice Principal - Research and Innovation (VP-RI) is responsible for overseeing the creation of the McGill Institutional RDM Strategy.
- The management committee of the RI-MUHC, the leadership of the CORE RI-MUHC and its CORD and BCU units are responsible for developing the RI-MUHC adaptations of the McGill Institutional RDM Strategy.
- Spearheaded by the DRS Team, the RDM Working Group is leading the writing and consultation efforts for the McGill Institutional RDM Strategy.

Existing McGill University Support for RDM

Recognizing the significance and value of research data, the **McGill Digital Research Services (DRS) Hub** was formed in 2021 as a unified effort to streamline RDM support for researchers. The mission of the McGill DRS Hub is to provide a "one-stop shop" for RDM services. The McGill DRS Hub is a collaboration between VP-RI, McGill Library, and IT Services. A portfolio of training and consultations on a variety of RDM-related topics, including data curation, RDM-related research software, software development and sharing, and Advanced Research Computing (ARC) use, is offered and organized by McGill Library, DRS Hub and by the National ARC Platform in collaboration with Calcul Québec (CQ).

New events are added regularly based on researchers' needs (see **Appendix B** for a full list of current RDM services and support provided by DRS, Library, and IT services). In alignment with the Tri-Agency RDM Policy priorities, Data Management Plan (DMP) guidance is a key focus of current institutional training and support. In addition, the McGill Library partners with national RDM service providers to offer researchers an institutional data repository (McGill University Dataverse), which facilitates making research data FAIR.

Existing RI-MUHC Support for RDM

At this point in time, clinical data from many RI-MUHC led projects cannot be held within the McGill University Dataverse as they include individual level participant health data that is subject to strict privacy and confidentiality legislation. These data are stored on the secure servers of the RI-MUHC within the MUHC digital environment. The RI-MUHC's IT Services and Solutions department manages these servers in collaboration with IT Services at the MUHC. Storage is also available at the RI-MUHC for other types of research and the RI-MUHC IT Services is collaborating with McGill University to allow researchers to access Compute

Canada/Calcul Quebec storage through McGill main campus, as applicable. RI-MUHC scientists are strongly encouraged to consult the IT Services and Solutions Department to, on a project-by-project basis, identify secure storage and computational space needs, secure methods of data access and sharing, and reviewing security of digital apps and tools used to collect data from study participants.

Scientists are also able to encouraged to consult with the CORD unit (Figure 2) for guidance on developing and structuring questionnaire and case report forms, for programming these into electronic data capture systems like RedCap, and for monitoring quality and completeness of data entry. They may additionally consult the Maelstrom team for data cataloguing and data harmonization services. Biostatistical consultation is available, as discussed, through the BCU, and may be relevant at the data organization planning stage with an eye to future analyses.

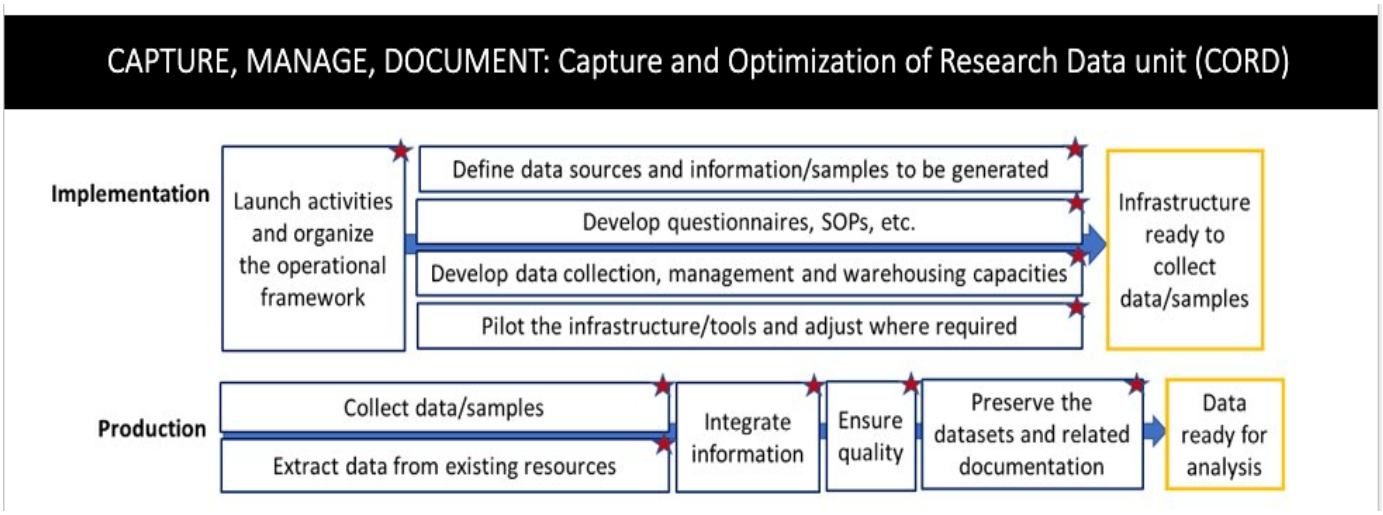

Figure 2. Support available through the CORD Unit

Goals and Objectives of the McGill University RDM

Goal #1: Increase Awareness

- **Obj. 1.1**: Promote McGill's Institutional RDM Strategy to research community members
- **Obj. 1.2**: Provide forums for feedback to advance RDM efforts at McGill
- **Obj. 1.3**: Empower researchers to adopt robust RDM practices by leveraging existing institutional resources
- **Obj. 1.4**: Foster domain-specific communities of practice through symposia, workshops, targeted outreach, and partnerships

Goal #2: Expand RDM Support and Training

- **Obj. 2.1**: Provide RDM training programs for researchers and students aligning with collaborative efforts both nationally and internationally
- **Obj. 2.2**: Work with the research community and key stakeholders (e.g., ethics, IT services, procurement, and legal) to improve the following core RDM service domains:
    - **Obj. 2.2.1**: Data management plans (DMPs)
    - **Obj. 2.2.2**: Data sharing, deposit, and preservation
- **Obj. 2.3**: Leverage existing networks of research support professionals, including the McGill DRS Hub, liaison librarians, and the Digital Research Alliance of Canada (the Alliance), to expand frontline RDM consultation capacity
- **Obj. 2.4**: Expand and improve support and training services using evidence-based evaluations and reviews
- **Obj. 2.5:** Increase numbers of staff with expertise to provide RDM-related support and services to advise researchers in RDM best practices
- **Obj. 2.6**: Increase RDM skills and competencies training across the academic curriculum

Goal #3: Establish a formal governance structure for research data at McGill

- **Obj. 3.1**: Form an advisory research data committee to the VP-RI, comprised primarily of faculty members with representatives from key stakeholders (e.g., IT, Library, VP-RI, Provost, Research Ethics Boards, etc.) to:
    - **Obj. 3.1.1**: Complete an analysis of relevant policies to ensure our institutional RDM framework is both coherent and in compliance with applicable laws and regulations
    - **Obj. 3.1.2**: Establish roles and responsibilities for compliance with the Tri-Agency RDM policy
    - **Obj. 3.1.3**: Propose revisions to, and updates of, existing research data-related policies
    - **Obj. 3.1.4**: Support departments and faculties in recognizing and rewarding efforts to value good RDM practices and ethical data sharing when possible
- **Obj. 3.2**: Strengthen communication and coordination with Affiliated Hospitals and research institutes in establishing more streamlined RDM workflows and processes
- **Obj. 3.3**: Pursue a research data stewardship model that positions McGill as a leader within the larger ecosystem of national and international RDM organizations
    - **Obj. 3.3.1**: Recommend the hiring of a research data privacy officer to oversee institutional capacity in accordance with applicable privacy laws and regulations

Goal #4: Develop RDM Services Through Partnerships

- **Obj. 4.1:** Seek commitments and investments from RDM service providers, funders, and governmental agencies, to ensure the long-term sustainability of institutional RDM support
- **Obj. 4.2**: Establish long-term partnerships

    - **Obj. 4.2.1**: Partner with provincial and national RDM infrastructure providers (e.g., Calcul Québec, the Alliance and its FRDR repository, Scholars Portal, etc.) and existing research community platforms to address gaps in repository technologies for managing, sharing, depositing, and archiving sensitive data and large data
- **Obj. 4.3**: Grow institutional RDM capacity
    - **Obj. 4.3.1**: Develop expertise in RDM amongst research support staff (e.g., grant officers and IT support staff), REB staff and members, and librarians
    - **Obj. 4.3.2**: Promote integrated interoperable systems for research-related records (e.g., DMPs, REB protocols, and institutional grant management)
    - **Obj. 4.3.3**: Focus on ensuring equitable, diverse, and inclusive representation in RDM-related roles

Additional Goals and Objectives of the RI-MUHC adaptation of the McGill University RDM strategy

Goal 1: Develop and offer services to support RDM

The CORD unit operated in a pilot phase in 2021 with important hirings and reorganization in 2022. The unit continues to grow and expand its operations. Our goal is to guide RDM development and execution for all studies led through the RI-MUHC, to ensure best practices and the highest quality data possible.

Professor Isabel Fortier, scientific lead of the CORD unit, also leads the Maelstrom unit within her own research program. Maelstrom provides consultative services for data cataloguing and harmonization, key activities in RDM. We aim to ultimately

integrate these activities into CORD operations, as we grow the unit, its staffing, and its capacity.

Goal 2: Update the MUHC/RI-MUHC Data Governance Framework

While RI-MUHC scientists both collect primary data and use secondary data sources. MUHC clinical and administrative data are potential secondary data sources, available through the MUHC Data Warehouse. The MUHC and its RI are in the process of updating a Data Governance Framework that will allow development of policy that ensures that MUHC data are widely available for research purposes in conformity with security and privacy legislation. They are also jointly engaged in a cataloguing exercise of the wide range of data available.

Looking Ahead

Digital technologies are profoundly transforming academic research across all disciplines. The McGill RDM Strategy will remain a living document subject to recurring evaluation and reviews. With the involvement of key stakeholders, the Digital Research Services Hub (DRS) will spearhead new initiatives and efforts to enhance McGill researchers' knowledge and skills in RDM, while adapting to the evolving requirements and legislations in Quebec, Canada and internationally.

Building on existing initiatives, McGill will continue to lead in developing and leveraging national and international RDM resources to broaden the global impact of its research. We will strengthen our collaborative efforts with recognized RDM organizations and service providers through the DRS to expand McGill's portfolio of RDM resources. The DRS will also lead the efforts to develop an implementation plan for this institutional strategy with the goal of augmenting RDM support resources and staff at McGill over the next 3 to 5 years. We will focus our initial efforts in raising awareness of RDM best practices through Library workshops, the McGill RDM Online Learning Program and the regular "DRS Drop-in Sessions" to

provide training, resources and forums for researchers to connect with subject matter experts from Library, IT and Ethics.

McGill is committed to working in partnership with the Tri-Agency, the Digital Research Alliance of Canada, Canadian Universities, and other stakeholders to ensure the success in adopting innovative RDM practices at the institutional level and beyond. The RI-MUHC is committed to working in collaboration with McGill University and enhancing its growing portfolio of fee for service consultation and support that are critical to the research operations of its community.

Appendix A: Definitions

- **CARE Principles** were developed by the Global Indigenouse Data Alliance. CARE refers to Collective Benefit, Authority to Control, Responsibility and Ethics. "The CARE Principles for Indigenous Data Governance are people and purpose oriented, reflecting the crucial role of data in advancing Indigenous innovation and self-determination." (Based on CARE Principles for Indigenous Data Governance).
- **Community of Practice (CoP)** refers to a group of people who share a common concern, a set of problems, or an interest in a topic and who come together to fulfill both individual and group goals. CoP often focuses on sharing best practices and creating new knowledge with ongoing interactions in meetings or collaborative platforms to communicate, connect and conduct community activities. (Adapted from: Community of Practice)
- **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. (Adapted from: Tri-Agency RDM Policy FAQ)
- **Data Lifecycle** refers to all the stages in the existence of data from creation to destruction. The data lifecycle provides a high-level overview of the stages involved in successful management and preservation of data for use and reuse. This broadly includes the following stages: Plan, Create, Process, Analyze, Disseminate, Preserve and Reuse. (Adapted from: CASRAI Definition of Data Lifecycle, DataOne, & Alliance-Portage)

- **Data Stewardship** refers to knowledge and skills required to effectively manage data assets. Data stewardship is often described as data governance in action. This includes the oversight of data to ensure fitness for use, the accessibility of the data, and compliance with polices, directives and regulations. (Adapted from: Statistics Canada Data Literacy Training)
- **Distinctions-based** [refers to] the three federally recognized Indigenous groupings in Canada: First Nations, Métis, and Inuit. A distinctions-based approach [is] intended to remedy the previous "pan-Aboriginal" or "one size fits all" approach to Indigenous policy and decision making, in which the unique rights, interests and circumstance of First Nations, the Métis Nation and Inuit are acknowledged, affirmed, and implemented. (Adapted from: The Government of Canada's terminology and linguistic bank and Principles respecting the Government of Canada' relationship with Indigenous peoples)
- **FAIR Principles** for scientific data management and stewardship are international best practice for improving findability, accessibility, interoperability and reuse of research data. (Adapted from: Tri-Agency RDM Policy FAQ**)**
- **OCAP®** is a registered trademark of the First Nations Information Governance Centre (FNIGC). It refers to Ownership, Control, Access and Possession. (Adapted from: First Nations Information Governance Centre Understanding OCAP®).
- **Research Data Management** refers to the processes applied through the lifecycle of a research project to guide the collection, documentation, storage, sharing and preservation of research data. (Adapted from: Tri-Agency RDM Policy FAQ and Alliance-Portage Definition)
- **Researcher** means any member of the University community who engages in or supervises research. (Adapted from: Regulations on Conduct of Research at McGill)

Appendix B: Currently Available Services and Support

- **General Guidance and Questions**
  - McGill Digital Research Services Hub
- **RDM Training and Support**
  - RDM guidance: Library RDM Web Guide & DRS FAQ
  - McGill RDM Learning Program – DRS & Library
  - RDM instructional videos in English and French – DRS & Library
  - RDM, RS & ARC general consultations – DRS Helpdesk via drs@mcgill.ca and Library via rdm.library@mcgill.ca
  - DRS drop-in sessions – DRS, Library, ITS, Ethics
  - OCAP® workshop training, in partnership with First Nations Information Governance Centre
- **Data Management Plan (DMP) Resources and Support**
  - DMP reviews – Library & DRS
  - DMP consultation for grants and research contracts – DRS
  - DMP Assistant Tool – Library
  - DMP workshops – Library
- **Data Deposit Resources and Support**
  - McGill University Dataverse – Library
  - Data deposit and preservation consultations – Library
  - Data deposit workshops – Library
- **Ethics**
  - Forms and Guidelines - McGill Research Ethics Board (REB) Office
  - Submission Process - McGill Research Ethics Board (REB) Office
  - Medicine REB - Medicine IRB - Research Ethics Office (Institutional Review Board)
- **Cybersecurity**
  - Cloud Directive – IT Services

    - Secure Your Journey – IT Services
    - Cybersecurity plan consultation for large grants – DRS
    - Enterprise software for secure data storage – IT Services
    - Available software at McGill – IT Services
- **Advanced Research Computing**
    - Infrastructure and Services - Calcul Québec
    - User Support – Calcul Québec
    - Training – Calcul Québec
    - RDM at The Digital Research Alliance of Canada
    - Advanced Research Computing Services – The Digital Research Alliance of Canada
- **Community Resources**
    - Canadian Research Data Centre Network
    - Computational and Data Systems Initiative – Faculty of Science
    - McGill-Concordia Laboratory of the Quebec Inter-University Centre for Social Statistics - QICSS
    - NeuroHub

Resources

European Commission, Directorate-General for Research and Innovation, (2016). *H2020 programme guidelines on FAIR data management in Horizon 2020*. Retrieved from https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

European Commission, Directorate-General for Research and Innovation, (2019). *Cost-benefit analysis for FAIR research data: cost of not having FAIR research data*, Publications Office. https://data.europa.eu/doi/10.2777/02999

Government of Canada. (2020). *Roadmap for Open Science*. Office of the Chief Science Advisor. Retrieved from https://www.ic.gc.ca/eic/site/063.nsf/eng/h_97992.html.

Government of Canada. (2021). *Principles respecting the Government of Canada's relationship with Indigenous peoples*. Retrieved from https://www.justice.gc.ca/eng/csj-sjc/principles-principes.html.

Project de loi n64 2021, c25: An Act to modernize legislative provisions as regard the protection of personal information. Retrieved from National Assembly of Quebec: http://www2.publicationsduquebec.gouv.qc.ca/dynamicSearch/telecharge.php?type=5&file=2021C25F.PDF

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC

(General Data Protection Regulation). Retrieved from https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng

Research Data Alliance International Indigenous Data Sovereignty Interest Group. (September 2019). *CARE Principles for Indigenous Data Governance*. The Global Indigenous Data Alliance. www.gida-global.org/care

Tri-Agency of Canada. (March 15, 2021). *Tri-Agency Research Data Management Policy*. https://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, *3*(1), 1-9. https://doi.org/10.1038/sdata.2016.18