Suncrest College Research Data Management Strategy

Introduction and Background

The National Research Council (NRC) and the Tri-Agency Funders (CIHR, SSHRC, NSERC) undertook public consultation to understand the state of research data management within the post-secondary landscape. This led to the NRC and Tri-Agency releasing a policy statement on research data management. Within this policy, it stipulates that each institution will have its own strategy for research data management (RDM) and that:

> "the strategy must be made publicly available on the institution's website, with contact information to which inquiries about the strategy can be directed. Having the strategies publicly available will help the agencies and the broader research community to understand institutions' current and planned RDM capacity, challenges, and needs, and will facilitate ongoing dialogue and collaboration on the advancement of RDM in Canada."

In accordance with this, Suncrest College set out to develop a subcommittee whose task was to outline a strategy for the college and its audience. This Committee is known as the

Research Data Management Advisory Committee (RDMAC).

In line with the Ownership, Control, Access and Possession standard and the Tri-agency acknowledgment of Indigenous data sovereignty, the Suncrest College Research Data Management Strategy will also acknowledge the data management needs of Indigenous peoples, communities, and nations. This strategy will encompass a multitude of facets to help the College not only bring awareness to the data management lifecycle but to the education and needs of the populations we serve within research. For this reason, the strategy is integrative and holistic in its inclusion of RDM for both Indigenous and non-Indigenous populations.

Importance of Research Data and Research Data Management

What is Research Data?

Research data is seen as output (Tri-Agency) but can comprise any of the following definitions/ideas. Note that this is not an exclusive list but merely a list to generate ideas and to help guide the researcher/employee in discerning the scope of their own work:

- Data that is generated by equipment/computations and that can be analyzed.
- Information that has been collected or observed that answers or validates a question.
- Stories that are told (images, recordings, artefacts, etc.) that are collected through the process of storytelling.
- Raw data that is not necessarily published but needs to be managed and/or organized.
- Qualitative and quantitative data points that are used to analyze and garner greater understanding of a function, process, or item, typically will be published, or shared in a scholarly manner.

What is Research Data Management?

Research data management (RDM) refers to the processes applied throughout the lifecycle of a research project to guide the collection, documentation, storage, sharing, and preservation of research data

(Digital Research Alliance of Canada). This can include the collection, documentation, storage, sharing, and preservation of research data. RDM practices are integral to conducting responsible research and can help researchers save resources by ensuring their data is complete, understandable, and secure. RDM practices also follow institutional and funding agency guidelines that protect their investments. The broader research community can derive maximum value from research data that can be accessed, shared, reused, and repurposed. In simplified terms, RDM refers to how we handle the data internally and externally to the project such as how researchers handle data when they are doing the research and when they complete and archive it.

Indigenous data sovereignty refers to the right of Indigenous peoples to control data from and about their communities and lands, articulating both individual and collective rights to data access and to privacy (from Indigenous Data Sovereignty). Indigenous research data must be repatriated, and permissions must be sought in sharing these knowledges, creative forms, and/or stories.

Why is RDM Important?

RDM practices expose researchers, faculty, students, and administrators to a range of benefits including:

- increased competitiveness in granting applications
- increased accountability and safety measures
- increased accuracy and validity of data
- ensured long-term preservation of data when applicable
- ensured consistency in data depositing and sharing requirements
- clarity in guidelines
- increased credit and impact of the data.

Part of the RDM process includes the creation of Data Management Plan (DMP). A DMP is a formal document that details the strategies and tools you will implement to effectively manage your data during the active phase of your research, and the mechanisms you will use for preserving and appropriately sharing your data at the end of the project. A DMP is a "living" document that can be modified throughout your project to reflect any changes that have occurred (Digital Research Alliance of Canada).

A DMP helps:

- Meet grant application requirements and/or adhere to institutional data mandates.
- Make it easier for all team members to document, understand, find, and use data.
- Plan the resources, tools, and expertise needed for data management.
- Identify challenges for storing, handling, and managing the types and volume of data.
- Ensure reliability, authenticity, accuracy, and reproducibility of data.
- Have a detailed account of data collection, handling, and stewardship practices.
- Plan how to make data FAIR (findable, accessible, interoperable, and reusable) to maximize the research potential and impact of data (Digital Research Alliance of Canada).

Scope and Objective

The target audience for the Suncrest College RDM Strategy are the Staff, Students, Partners, External Researchers, and Funders who operate within the auspices of applied research at the College. The overall

objective of this strategy is to develop, promote, bring awareness to, and maintain a research data management strategy for the college. Components of this strategy will highlight the key players involved in consultation and development of this key document and the key activities associated with where the college will head in the next three years. Major components of this strategy follow the data lifecycle as specified by the Digital Research Alliance of Canada. Additionally, the principles of OCAP™ are followed with respect to the Indigenous Data Sovereignty aspects of this strategy. This strategy reflects the breadth and scope of all types of research data handling and generation at the College.

The strategy targets four priority objectives over the next 3 years (April 2023 to April 2026). These are as follows:

1. Provide training, resources, and education on the stages of the research data management lifecycle.
2. Continue to expand and increase the capacity for RDM expertise at the College.
3. Use and integrate resources that exist provincially and nationally to further RDM education within the institution.
4. Support and encourage appropriate Indigenous Research Data Management processes and activities.

5. Data and projects that are not research related are excluded from the scope of this strategy. However, those involved are highly encouraged to participate in approved processes and education. This level of data is to follow the basic concepts of data governance and should involve levels of Senior Executive Approval as needed when it exceeds the normal scope of practice. However, resources created from our Strategy can be used as a guideline for that data which falls outside of scope.

Oversight and Review

This research strategy is to live within the Research Centre at Suncrest College but will be revisited after the three-year term allotted to the implementation plan. Oversight of this strategy resides with the RDMAC.

The RDMAC will:

- Develop a Research Data Management Strategy for Suncrest, which will include gathering the evidence and information required to assess current state against ideal state.
- Build a comprehensive workplan with timelines and roles and responsibilities to implement the Research Data Management Strategy and to address gaps between current and ideal state.
- Identify and establish series of workgroups to support this work as needed.
- Provide recommendations to the Vice President, Partnerships and Innovation and the Applied Research Lead regarding next steps emerging from Research Data Management Strategy.
- Seek input from a range of academic and administrative partners through liaising with the Suncrest community and other partners as required in investigating best practices from across Canada in RDM.

The RDMAC will be comprised of representation from areas of the College most affiliated with the oversight and maintenance of research data management. The membership shall consist of

representatives from the following areas:

- Suncrest College Applied Research ~~Centre~~
- ECRF/Suncrest College Research Farm
- Suncrest College Legal/Finance
- Suncrest College Information Technology Services
- Suncrest College RIDE Committee
- Suncrest College Biosafety/Facilities
- Suncrest College Teaching and Learning

The Vice-President, Partnerships & Innovation will receive updates on the work and progress of the Advisory Committee through the Applied Research Lead and will obtain approval for any final reports or strategy from the Vice-President, Partnerships & Innovation prior to implementation.

Timelines and Activities

To meet the four priority areas in a three-year time frame, the following activities have been identified by the RDMAC to move work forward. As we move the course, the RDMAC will put together a more fulsome work plan that aligns will all these activities and will include metrics of success, more detailed timelines, and activity owners. This will be tracked by the Applied Research Office.

Priority Area	Activities
1. Provide training, resources, and education on the stages of the data management lifecycle	Develop a RDM guideline documents and graphics Develop training workshops for faculty, staff and students that are accessible and can be delivered in a timely manner. Develop a glossary of terms that can be consistently referenced throughout the institution and updated regularly Develop guidelines and policies as relevant to the RDM and lifecycle. Offer educational opportunities around data transferring procedures and policies as developed. Reflect on what our research values are as an institution- integrate into various initiatives and documents (resources that are tangible) Begin to develop capacity that will lead to RDM policy development post 3 years of Strategy implementation Develop data management workshops that can be delivered to both internal and external partners.
2. Continue to expand and increase the capacity for RDM expertise at the College	Develop workplan that aligns with strategy overall and identifies yearly metrics to be tracked Create faculty and staff champions to represent the data lifecycle Create working groups to reflect various working group interests such as IT and infrastructure, data transfer, data depositing and curation, and others.

3. Use and integrate resources that exist provincially and nationally to further RDM education within the institution	Develop links with various external groups (provincial and national) and bringing salient points back to RDMAC and college.
4. Support and encourage appropriate Indigenous Research Data Management processes and activities	Encourage conversations with Indigenous partners Source or develop and deliver Indigenous data management related workshops to both internal and external partners Liaise with appropriate partners in further developing RDM capacity for Indigenous research within the College

Looking Ahead

Moving forward, Suncrest College and its RDMAC will work on developing a comprehensive work plan that will help to move the identified goals forward within a three-year timeframe. This will include associated work groups falling under the RDMAC to target specific topics such as data management planning, IT infrastructure and depositories, and archiving and re-use of data among other topics proposed and identified by the Committee. While this Strategy identified the beginning components of our work as a College, it is the basis of our work in moving forward and will be a living document.

Other Relevant Strategies and Policies

The following is a list of the relevant policies, procedures, and external reference documents for the RDMAC, and this strategy will be used in its progression forward:

Internal Policies:

- Integrity in Research and Scholarship
- Research Administration

External References and Documents:

- Tri-Agency RDM Policy: https://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html
- Digital Research Alliance of Canada RDM: https://alliancecan.ca/en/services/research-data-management
- Ownership, Control, Access, and Possession (OCAP): https://fnigc.ca/ocap-training/
- Research Involving the First Nations, Inuit, and Métis People of Canada: https://ethics.gc.ca/eng/tcps2-eptc2_2018_chapter9-chapitre9.html
- The CARE Principles for Indigenous Data Governance: https://datascience.codata.org/articles/10.5334/dsj-2020-043/

- Research Data Management Terminology https://codata.org/initiatives/data-science-and-stewardship/rdm-terminology-wg/rdm-terminology/

Definitions

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" (Tri-Agency Research Data Management Policy, Frequently Asked Questions, Government of Canada 2021).

Metadata: Metadata is data about data and is the information needed to make a dataset discoverable, citable, and usable by others.

Research: Research is creative and systematic work that is undertaken to increase knowledge in a particular area or discipline. It involves the collection, organization, and analysis of information to increase understanding of a topic or issue.

Research Data: "Data that are used as primary sources to support technical or scientific enquiry, research, scholarship, or artistic activity, 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. 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" (Committee on Data, International Science Council).

Research Data Management: "Research Data Management refers to the storage, access and preservation of data produced from a given investigation. Data management practices cover the entire lifecycle of the data, from planning the investigation to conducting it, and from backing up data as it is created and used to 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)" (Committee on Data, International Science Council).