



Institutional Data Management Strategy

March 1st, 2023

Table of Contents

1. Preamble.....	3
2. Objectives.....	4
3. Definitions.....	5
4. Principles.....	7
5. Scope.....	8
6. Roles and Responsibilities.....	8
7. Approach.....	9
7.1 Awareness and support	10
7.2 Research data management practices.....	10
7.3 Governance and management	11
7.4 Infrastructure	11
7.5 Dataverse	12
8. Timeline.....	12
9. Additional resources	13
10. Contact	14

1. Preamble

This strategy is attached to the Institutional Research Policy¹ of the National Circus School. This strategy responds to the Tri-Agency Research Data Management Policy² put forth by the Canadian Institutes of Health Research³ (CIHR), the Natural Sciences and Engineering Research Council⁴ (NSERC), and the Social Sciences and Humanities Research Council⁵ (SSHRC) (hereafter, The Agencies). In particular, the Tri-Agency Policy aims to cultivate research data management to:

- support research excellence through the promotion of strong research data management practices;
- increase the citation of Canadian datasets;
- increase the recognition of Canadian researchers for their production and sharing of data;
- support researchers in data management practices that are consistent with ethical, legal, and commercial obligations and those from The Agencies;
- recognize Indigenous data sovereignty and that research by and with Indigenous peoples will be managed according to their principles and in partnership with them;
- align Canadian data management practices to international standards to enable international research collaboration; and
- increase data archiving, discoverability, and reuse to increase research efficiency, discovery, and innovation whenever possible.

The Tri-Agency Policy has three primary requirements:

Development of an institutional research data management strategy: This institutional strategy outlines how the institution will work to provide research data management support and guidance to its researchers. This strategy must be posted publicly on the institution's website by March 1, 2023.

*Development of data management plans:*⁶ A data management plan is a document that is specific to a particular research project that outlines how data will be handled through all stages of its lifecycle. As of the publishing of this strategy, the inclusion of data management plans within funding applications for The Agencies will not be required for all opportunities

¹ https://ecolenationaledecirque.ca/wp-content/uploads/2022/06/politique_institutionnelle_de_la_recherche_12.12.2018_approuvee_par_ca_lg.pdf (Last accessed: Nov 28, 2022)

² <https://science.gc.ca/site/science/fr/financement-interorganismes-recherche/politiques-lignes-directrices/gestion-donnees-recherche/politique-trois-organismes-gestion-donnees-recherche> (Last accessed: Nov 28, 2022)

³ <https://cihr-irsc.gc.ca/f/193.html> (Last accessed: Nov 28, 2022)

⁴ https://www.nserc-crsng.gc.ca/index_fra.asp (Last accessed: Nov 28, 2022)

⁵ <https://www.sshrc-crsh.gc.ca/home-accueil-fra.aspx> (Last accessed: Nov 28, 2022)

⁶ <https://assistant.portagenetwork.ca/> (Last accessed: Feb 1, 2023)

until the completion and assessment of a pilot phase for specific funding opportunities in 2023.

Data deposit: Following an assessment of institutional data management strategies, The Agencies will require grant recipients to deposit data, metadata, and code that supports their research and findings that are presented in journal publications and pre-prints, and such deposits should be made available for reuse by others where ethical, legal, and commercial obligations allow. It should be acknowledged that this is not an open data policy that requires researchers to allow universal access to all research data. This data deposit initiative aligns with the Government of Canada's commitment to open science⁷ that supports increased access to the results (including publications and public contributions) and scientific data (where appropriate) of publicly funded research.

This document serves two primary purposes. First, to bring a focus to research data management at the National Circus School (hereafter, the School) by defining the objectives, principles, and a vision for research data management at the School. Second, to define an approach for increasing the capacity for research data management at the School. The strategy and approach presented herein can be viewed as high-level direction that will guide the development of research data management policies and practices at the School in the coming years. Although lower-level details are not flushed out, such details will be determined in time as part of our approach defined in Section 7.

This document is a living document that may evolve in time as the institutional research data management efforts change depending on the capacity and needs of the National Circus School.

2. Objectives

The following outlines the objectives of the research data management initiative at the School:

- Recognize and establish the importance of research data management at the School by defining overarching objectives and principles, and by outlining an approach for increasing the capacity of research data management.
- Raise awareness for research data management to all those involved in research at and with the School.
- Develop a culture where research data management best practices are integrated into common research practices for all researchers of the School.
- Recognize that data generated from research at the School is unique and a valuable research output that can impact circus arts, performing arts, and human performance as well as other domains. This data should be properly managed throughout its lifecycle to enable increased research efficiency and expedite discovery both within and outside of the School.

⁷ <https://science.gc.ca/site/science/sites/default/files/attachments/2022/Roadmap-for-Open-Science.pdf> (Last accessed: Feb 1, 2023)

- Develop a working environment at the School that supports researchers in working with and managing all types of research data. Such an environment should align with ethical, legal, and commercial requirements as well as those put forth by The Agencies.⁸
- Recognize and support Indigenous data sovereignty for research done with and about Indigenous peoples, and that related data should be managed according to principles developed and approved by these groups.

3. Definitions

Research: An undertaking intended to extend knowledge through a disciplined inquiry and/or systematic investigation.⁹ Key areas of research at the School include:

- Human Performance;
- Experiential and Environmental Design;
- Social Innovation;
- Digital Technologies.

Such activities may or may not take place at the National Circus School. Research may also constitute activities that are aimed at supporting industry partners.

Researcher: An individual performing research activities at the School who is affiliated with the School in some form, either through full- or part-time employment, contractual work, research association, long- or short-term student involvement, or volunteering.

Research data: A collection of information to be used for research purposes, including human biological materials.

Data lifecycle: The entire period of time whereby a particular dataset, or a derivative of a dataset, is in existence. The lifecycle often includes a number of sequential, and potentially interconnected, stages: planning, creation, documentation, processing, analysis, storage, archiving, and destruction.

Metadata: Details and documentation that describes, and are specific to, a dataset and that provide a sufficient overview for a general understanding of said dataset. Common information documented in metadata include: data size, file format, type of data (textual, numeric, photos/images, videos, etc.), context and details of origin, filenaming convention, guidance of internal data details, database relation or relation to other data/datasets (if applicable), owner or contact individual, etc.

Research data management (RDM): The processes and practices applied by researchers and teams to appropriately manage data through all stages of its lifecycle with the intention of

⁸ <https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management/tri-agency-research-data-management-policy> (Last accessed: Nov 30, 2022)

⁹ <https://ethics.gc.ca/eng/documents/tcps2-2022-en.pdf> (Last accessed: Jan 31, 2023)

increasing research efficiency, clarity, and data reusability. Specific procedures and practices may include (but are not limited to): planning, metadata creation, filenaming, file storage, backup, preservation, security, quality control and assurance, documentation, licensing, sharing, ethics, discoverability, and reuse, amongst others as appropriate.

Data management strategy: A roadmap that defines how an institution will develop and grow their research data management capacity to meet their research needs and align with ethical, legal, and commercial obligations as well as those put forth by The Agencies. A strategy may outline major components such as institutional readiness assessment, raising awareness, development of policies and working practices, and improving infrastructure. Low level details need not be specified nor may they be particularly clear at the time of strategy development, though all relevant details of working practices will be determined upon the execution of the strategy.

Data management plan: A formal document outlining details and practices that guide the management of research data for a specific research project. Data management plans often include (but are not limited to) details related to: data collection, metadata creation and documentation, analysis and processing, storage, access, privacy and security, sharing, reuse, preservation, and destruction. Resources¹⁰ to assist with the creation of data management plans have been developed and made available by the Digital Research Alliance of Canada.¹¹

Dataverse: A repository, either centralized or distributed, internal or external, that contains a sum of either digital or analog research datasets and associated materials and information (metadata, documentation, code, etc) that may or may not be specific to a single institution or research entity. Such a repository can enable and facilitate the preservation, discoverability, and reuse (when appropriate) of research data assets.

The Agencies: The term used throughout this document that collectively refers to the agencies behind the Tri-Agency Research Data Management Policy: the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council (NSERC), and the Social Sciences and Humanities Research Council (SSHRC).

Types of information: The following definitions detail the different types of information that may be collected during research involving participants, each of which requires important considerations with respect to data management. These definitions are from the Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans.¹²

Directly identifying information: The information identifies a specific individual through direct identifiers (e.g., name, social insurance number, personal health number).

¹⁰ <https://assistant.portagenetwork.ca/> (Last accessed: Feb 1, 2023)

¹¹ <https://alliancecan.ca/en> (Last accessed: Feb 1, 2023)

¹² <https://ethics.gc.ca/eng/documents/tcps2-2022-en.pdf> (Last accessed: Jan 31, 2023)

Indirectly identifying information: The information can reasonably be expected to identify an individual through a combination of indirect identifiers (e.g., date of birth, place of residence, or unique personal characteristic).

Coded information: Direct identifiers are removed from the information and replaced with a code. Depending on access to the code, it may be possible to re-identify specific participants (e.g., the principal investigator retains a list that links the participants' code names with their actual names so data can be re-linked if necessary).

Anonymized information: The information is irrevocably stripped of direct identifiers, a code is not kept to allow future re-linkage, and the risk of reidentification of individuals from remaining indirect identifiers is low or very low.

Anonymous information: The information never had identifiers associated with it (e.g., anonymous surveys) and risk of identification of individuals is low or very low.

Publicly available information: Any existing stored documentary material, records or publications, which may or may not include identifiable information, and that has no restrictions on its use or distribution, or that may be released under certain legal conditions.

4. Principles

The National Circus School commits to its researchers to develop and implement appropriate resources and infrastructure that enables effective research data management to continue to foster high quality research in circus arts, performing arts, and human performance.

The School recognizes that its data is an asset and an important research output. Given the unique nature of the work at the School, it is recognized that quality data management procedures, preservation, and sharing (where appropriate) of data generated by the School can expedite knowledge generation and the advancement of circus arts, performing arts, and human performance both within and beyond the research of the School. All uses of data, whether internal or external, should follow FAIR¹³ principles, enabling findability, accessibility, interoperability, and reusability of data.

Data must be treated with reverence in all stages of its lifecycle, and this includes in common and daily working practices as well as in cases of collaboration, sharing, and data reuse. Data security and privacy of information are of utmost priorities, and the commitment to quality treatment of data aligns with the Institutional Policy on the Ethics of Research Involving Humans¹⁴ and the

¹³ <https://www.go-fair.org/fair-principles/> (Last accessed: Nov 28, 2022)

¹⁴ https://ecolenationaledecirque.ca/wp-content/uploads/2022/06/politique_institutionnelle_sur_lethique_de_la_recherche_avec_des_etres_humains_12.12.2018_approuvee_par_ca_lg.pdf (Approved: Dec 2018; last accessed: Dec 2, 2022)

Institutional Policy on Research Integrity.¹⁵ Engagement in collaborations, data sharing, and data reuse will be determined based on thorough evaluation of working relationships or the parties involved and will consider aspects including integrity, intellectual property, and ethics (among others) on a case-by-case basis.

The protection of Indigenous peoples and communities is recognized whereby these groups have the right to data ownership and governance for any data, including research data, from or about them. Indigenous data management and data sovereignty are supported by subscribing to and following First Nations principles of OCAP®¹⁶ and CARE¹⁷ principles.

5. Scope

The guidelines, policies, and practices developed upon the execution of this data management strategy apply to all researchers of the School and all research data at the School regardless of its status with the data lifecycle.

The policies and practices developed from this strategy do not pertain to data that are directly related to, or that support, normal operating procedures of the scholarly and educational environment at the National Circus School. Should non-research data become research data through appropriate mechanisms, such data will be managed and treated as research data thereafter.

Research data management is closely related to other topics including research that involves human participants, ethical data usage (e.g., artificial intelligence uses), and others. All researchers of the School are required to follow the guidelines regarding research involving human participants outlined in Institutional Policy on Ethics of Research Involving Humans.¹⁸ The topic of ethical data usage extends beyond research data management and will not be addressed herein.

6. Roles and Responsibilities

The following details the responsibilities of individuals, groups, and entities that have or will have a role within data management at the School.

Executive management: The executive management is responsible for the application of this strategy.

¹⁵ https://ecolenationaledecirque.ca/wp-content/uploads/2022/06/politique_institutionnelle_sur_lintegrite_en_recherche_12.12.2018_approuvee_par_ca_lg.pdf (Approved: Dec 2018; last accessed: Dec 2, 2022)

¹⁶ <https://fnigc.ca/ocap-training/> (Last accessed: Nov 30, 2022)

¹⁷ <https://www.gida-global.org/care> (Last accessed: Nov 30, 2022)

¹⁸ https://ecolenationaledecirque.ca/wp-content/uploads/2022/06/politique_institutionnelle_sur_lethique_de_la_recherche_avec_des_etres_humains_12.12.2018_approuvee_par_ca_lg.pdf (Approved: Dec 2018; last accessed: Nov 28, 2022)

The Research Center Director: The director of the Research Center is responsible for ensuring the application of this strategy and for allocating a budget to allow for the development of enabling research data management at the School.

Data steward: The data steward is responsible for general oversight of and support for research data management at the School. The individual interacts with and serves as a resource to researchers within, affiliated with, or collaborating with the School to advise on data management practices. The individual also interacts with tangential resources as necessary to support research data management at the School.

Researchers: Individuals that perform research and produce data that lies at the foundation of the data lifecycle. Researchers have a critical role with respect to research data management and they must comply with the practices and guidelines developed through this strategy.

Research data management working group: The group of individuals that develops and implements research data management practices at the School. This group is expected to include (at a minimum) the Data Steward, Research Ethics Board Chair, Research Center Director, and Library Manager. This group is aware of the many areas that research data management touches and will seek out appropriate consultation and discussion to ensure research data management practices align well with all tangential entities to the School. This group may evolve into the governance panel as specified in Section 7.3.

Support services: The collective group of services that provide research data management support to researchers. This may include both expertise and guidance from knowledgeable individuals (i.e., Data Steward) as well as tools, documents, and information that are created or curated to assist researchers with research data management.

Research Ethics Board (REB) of the National Circus School: The mandate of the REB is to ensure respect for the dignity of human beings who participate in research activities carried out under the authority or under the auspices of the National Circus School. The Board of Directors of the School has entrusted the REB with the responsibility of evaluating the ethics of these activities in accordance with the applicable ethical requirements and laws.

7. Approach

Achieving the objectives for developing research data management at the School will require efforts in multiple directions including:

- awareness and support,
- development of research data management practices,
- governance and management,
- infrastructure development, and
- the creation of a dataverse.

Additional details and envisioned tasks for each of these five components are defined below. These components and the related tasks have been identified and selected based on the intersection between the necessity for research data management, the research environment the School, and the capacity and available resources at the School. The implementation of these tasks will require frequent collaboration and communication with the Research Ethics Board, information technology support, and library services, in addition to other resources as necessary, to ensure consistency and capacity among all entities that have a relationship to research data. In consideration of the capacity and resources available at the School, implementation of the items in the approach below may leverage resources outside of the School that, for example, are aimed to support Canadian researchers in general.

7.1 Awareness and support

Raising awareness within the School for research data management and providing support services in various forms is essential to developing a culture that exercises best practices in handling research data.

- Raise awareness about the benefits of research data management to researchers of the School and tangential research entities and partners.
- Identify and address educational needs of researchers related to research data management and research computing in general.
- Promote research data management among researchers of the School by providing training opportunities for researchers to understand existing and new research data management guidelines and practices.
- Develop resources in the form of documentation and presentations that enable researchers to exercise research data management best practices for all data types. Services and resources include both individuals with technical and operational knowledge as well as tools, documents, and curated information that researchers can reference and that speak to adopted best practices as well as tasks that will eventually be mandated by The Agencies (e.g., data management plans and data deposit).
- Provide onboarding procedures for new researchers (e.g., full-time, part-time, students, collaborators, etc.) to ensure all individuals have the knowledge and tools to appropriately work with and manage research data.
- Provide assistance to researchers of the School for research data management and related practices via access to a knowledgeable individual.

7.2 Research data management practices

The development of the working practices for research data management that researchers may use on a daily basis will require addressing many topics in data management. Given the research environment of the School, input from researchers will be invaluable to developing practices and procedures that align well with all interests.

- Through collaborative discussions with researchers of the School, develop specific research data management practices that fit the research environment of the School and that support the research directions of all researchers. Topics of focus will include (but are not limited to): working and archived data and file usage, storage, data transfer/transport, backup, security and privacy, metadata and documentation, archiving and long-term preservation, geographical implications, scope and jurisdiction of data management procedures and support, consideration of and procedures for analog data, Indigenous data sovereignty, management of Indigenous data for communities outside of Canada, data sharing and collaboration, secondary data use, intellectual property, copyright, and ownership.
- Streamline research data management practices to align with and parallel the production of other research-related work, especially including ethics applications for research involving human participants.
- Develop and integrate appropriate procedures for engaging in research with and about Indigenous people to support Indigenous data sovereignty.

7.3 Governance and management

Governance and management responsibilities are necessary to provide continual oversight of research data management at the School, both during the development phase as well as during continued implementation.

- Develop a research data management governance panel that consists of a small group of individuals, both from within and outside of the School, to guide current and future research data management.
- Perform ongoing and periodic assessments of research data management practices at the School to determine necessary adjustments or additions to resources for research data management.
- Assign data steward responsibilities of oversight, procedure and practice assessment, and general management for research data management within the School to a researcher of the School.
- Identify and address financial needs to support research data management efforts at the School to ensure the availability of future infrastructure and support services.

7.4 Infrastructure

Improving the research data management capacity and practices at the School will require addressing infrastructure related to data management and computing in general.

- Understand the current data and computing landscape at the School to provide insight on existing data practices and supporting infrastructure.
- Identify and implement appropriate infrastructure and related procedural changes and/or additions to support the research data management needs of the School and its researchers. This includes access to necessary hardware and software that may allow for

integrating new research data management practices within the normal operating procedures of researchers.

- Communicate and educate researchers on infrastructure and related procedural changes and/or additions to ensure all researchers can effectively and efficiently work via these new mechanisms.
- Provide the necessary resources for researchers to collaborate with external researchers and partners through tools, computing infrastructure, and protections that provide data security and that support the development of research and partner collaborations.
- Perform ongoing assessment of the computing landscape at the School to guide future efforts to ensure the capacity, practices, and infrastructure related to research data management can meet the needs of all research directions.

7.5 Dataverse

Aligning with the notion that data generated through research at the School should be regarded as an asset, the creation of a dataverse will support data preservation and data reuse.

- Develop a dataverse specific to the School that contains existing and new data sets, associated documentation, metadata, and code to serve as a centralized repository to facilitate data preservation and data reuse where appropriate. Access to and reuse of data will be evaluated based on the data or dataset of interest, requesting individual(s), and intended research aims, though the development of such a dataverse and the ability for individuals to request access to data does not imply open data access. The implementation of a dataverse may leverage tools and resources that are made available to researchers of the School. It is acknowledged that analog data will require special consideration and resources.
- Assign oversight and management responsibilities of the School dataverse.
- Maintain an awareness of external dataverses and datasets, including those within and outside Canada, that may be relevant to the work of the School and make this information available to researchers through documentation and/or training opportunities.

8. Timeline

The major components above will be the focus of research data management capacity development at the School through 2023-2024; note that the calendar and fiscal year at the National Circus School begin (Q1) on July 1. Table 1 briefly outlines major tasks related to the approach at large as well as the anticipated completion timeline. Some tasks, including assessing the School's data management needs and computing landscape, as well as the development of this strategy, have commenced since Q1 of 2022-2023.

Table 1. Timeline for major tasks for research data management development at the National Circus School through 2023-2024.

Task	Related component(s)	2022-2023				2023-2024			
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Identify of needs related to research data management and infrastructure; understand computing and data landscape	Awareness and support, Infrastructure								
Develop institutional data management strategy; obtain board approval; publicly post and notify The Agencies by March 1, 2023	Institutional strategy								
Develop governance panel	Governance								
Implement infrastructure additions/changes; inform/educate researchers	Infrastructure								
Develop data management practices; produce supporting documentation and resources	RDM practices, Awareness and support								
Develop dataverse including data deposit and directory development	Dataverse								

Following the development phase, ongoing tasks include:

- providing continued awareness and support,
- periodically assessing and adapting infrastructure,
- assessing and adapting research data management practices as necessary,
- providing general governance and management, and
- providing continued oversight and management of the School dataverse.

9. Additional resources

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>

Tri-agency statement on principles of data management:

<https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management/tri-agency-research-data-management-policy>

Tri-agency research data management FAQ: <https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management/tri-agency-research-data-management-policy-frequently-asked-questions>

Data management plan assistant: <https://assistant.portagenetwork.ca/>

10. Contact

Please direct all comments and questions to critac@enc.gc.ca.