CERC Health Equity & Community Wellbeing Training

January 29-30, 2025

Led by: Meghan Landry, ACENET

A **regional partner** of the
Digital Research Alliance of Canada

Learning outcomes

- To clarify the **role of Digital Research Alliance of Canada and the greater Alliance Federation** in relation to HECW's research
- To define access to **digital research infrastructure, services, expertise, and resources** for all researchers
- To breakdown and explore the research data management lifecycle for HECW and the **various processes for Data Collection, Analysis, Access, Use and Storage**
- To provide **best practices for managing sensitive data in Canada**
- To demonstrate use of open-source qualitative research tools, **Taguette and REDCap**

The Digital Research Alliance of Canada is a not-for-profit organization funded by the Canadian Government serves Canadian researchers, "by integrating, championing and funding the infrastructure and activities required for advanced research computing (ARC), research data management (RDM) and research software (RS), [providing] the platform for the research community to access tools and services faster than ever before."

3

Federation Consortia
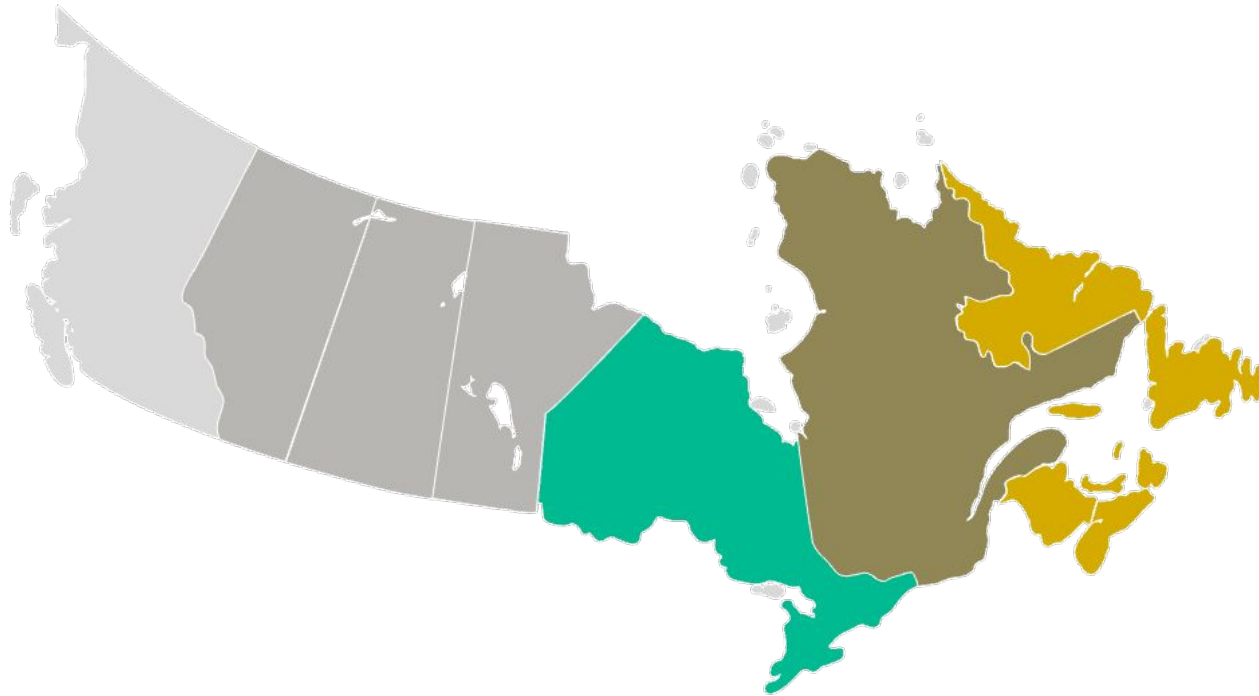
- ACENET
- Calcul Québec
- Compute/Calcul Ontario
- Prairies DRI
- BCDRI

Supporting Canadian Researchers - all disciplines

Three operational pillars:

1. Advanced research computing
2. Research data management
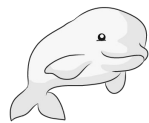3. Research software

1. Advanced research computing

Resources

- High performance, big data and GPU [computing](#)

- [Cloud](#) development space and storage that includes an outward facing IP address

- Petabytes of [active data storage](#) with secure backup systems

- Secure large file [transfer portal](#) (Globus)

Advanced Research Computing

National Systems (as of 2024)

Béluga - CPU, GPU, Storage, Cloud

Narval - CPU, GPU, Storage

Graham - CPU, GPU, Storage, Cloud

Niagara - CPU, Storage

Cedar - CPU, GPU, Storage, Cloud

Arbutus - Cloud

Major upgrade of our Advanced Research Computing infrastructure (2025 & onwards)

Our Advanced Research Computing infrastructure is undergoing major changes in the winter of 2024-2025 and spring of 2025 to provide better High Performance Computing (HPC) and Cloud services for Canadian researchers. This page will be regularly updated to keep you informed of the activities concerning the transition to the new equipment.

The infrastructure renewal will replace the nearly 80% of our current equipment that is approaching end-of-life. The new equipment will offer faster processing speeds, greater storage capacity, and improved reliability.

Taken from Infrastructure Renewal (2025)

New system details [edit]

New System	Old System to be Replaced
Arbutus	Cloud (as a virtual infrastructure there is no change to the cloud interface.)
Rorqual	Béluga
Fir	Cedar
Trillium	Niagara & Mist
Nibi	Graham

Taken from Infrastructure Renewal (2025)

Data storage options for HECW: active in-use data

Institutional options - TMU

1. GoogleDrive - 1TB of storage
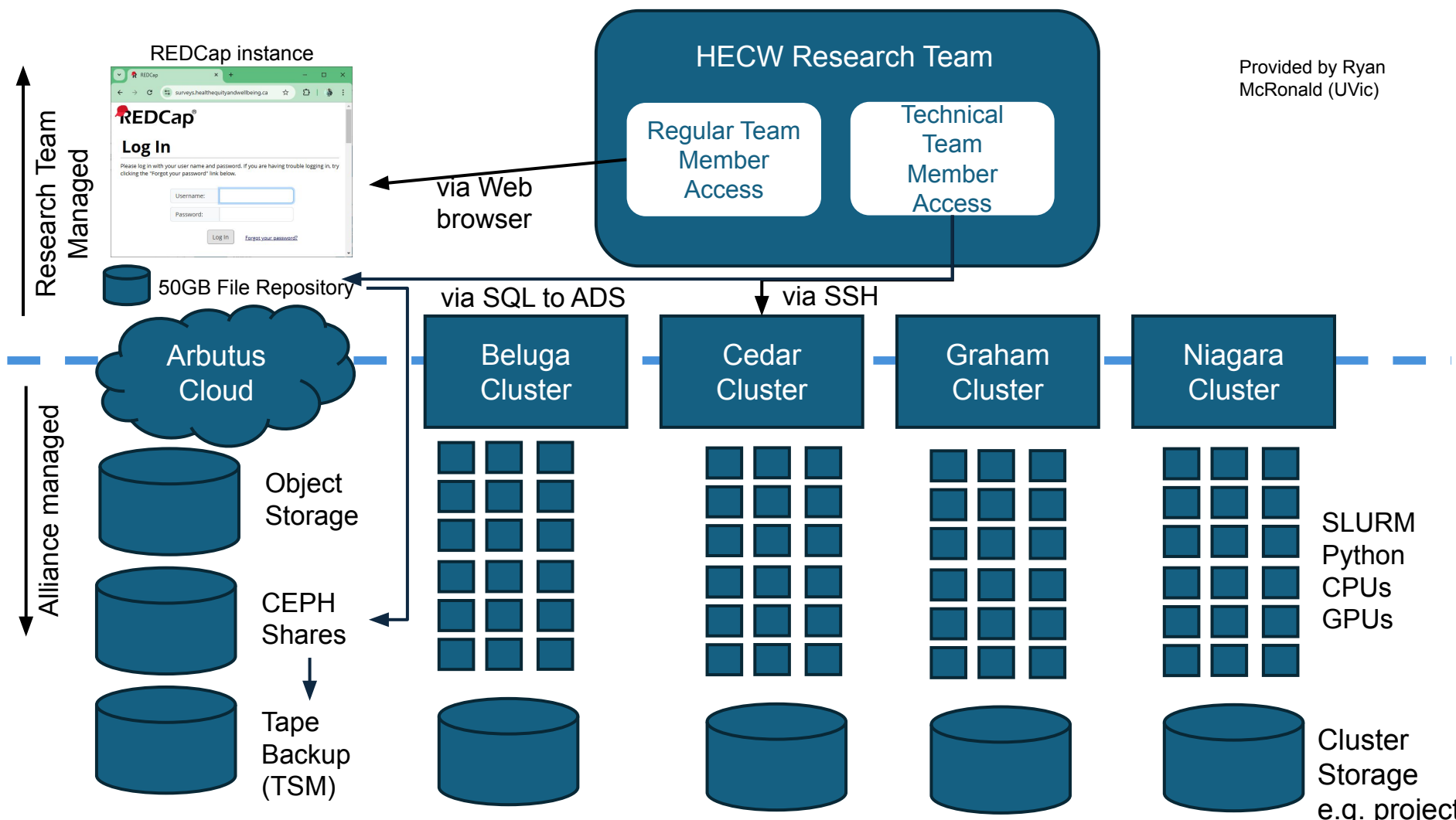   a. Would be suitable as a secondary backup for <u>low-risk data</u>

National options via DRAC

GUI-options (suitable for sensitive data)
1. NextCloud - only 100GB of data
2. REDCap file repository - set at 50GB limit right now (can be increased to accommodate TBs)
3. *SQL server (ActiveDataStore): available soon*

Requires familiarity with the command line
1. Clusters
   a. Provides access to data analytics tools
   b. *Not suitable for sensitive data because it is a shared platform*
2. Cloud computing & virtual machines

support@ace-net.ca

REDCap instance

Provided by Ryan McRonald (UVic)

Research Team Managed

Alliance managed

50GB File Repository

via Web browser

via SQL to ADS

via SSH

HECW Research Team

Regular Team Member Access

Technical Team Member Access

Arbutus Cloud

Object Storage

CEPH Shares

Tape Backup (TSM)

Beluga Cluster

Cedar Cluster

Graham Cluster

Niagara Cluster

SLURM Python CPUs GPUs

Cluster Storage e.g. project

Getting Access

https://alliancecan.ca/en/services/advanced-research-computing/accessing-resources

Resources

- RAS - On-Demand
- RAC - yearly
- Resources for Research Groups **(RRG)**
- Research Platforms and Portals **(RPP)**

Services

- CCF Wiki
- Disciplinary Support
- Software
- Special support, i.e. COVID

Expertise

- Visit Consortia websites
- Visit disciplinary support wiki
- Attend training courses

Services

- ‣ Technical support

- ‣ Training

- ‣ Centralized software stack

- ‣ Specialized Data Portals - i.e., Jupyter Hub

- ‣ Federated Research Data Repository (FRDR) Collaboration

- ‣ Globus File Transfer

- ‣ Database as a service (SQL)

Expertise

‣ Consultation - Helping to determine the resources needed

‣ Designing, optimizing and troubleshooting computer code

‣ Customizing tools

‣ Specialized support is available for a range of disciplines

‣ Cybersecurity

2. Research data management

Research data management

RDM Services We Provide

- Federated Research Data Repository (FRDR)
- Lunaris
- DMP Assistant
- Learning & Training
- Network of Experts
- Publications (Zenodo)

RDM Services We Support

- Borealis (Dataverse)

ACENET
accelerate discovery

support@ace-net.ca

Alliance-managed data repositories

FRDR: The Federated Research Data Repository (FRDR) is a bilingual publishing platform for sharing and preserving Canadian research data. It is a curated, general-purpose repository, custom built for large datasets.

- DOIs provide a persistent identifier for data citation
- Default allocation of 1TB of repository storage
- Dataset review and ongoing support provided by the FRDR curation team
- Optional embargoes on data and metadata records

Data repositories

[Lunaris](): Canada's national discovery service for multidisciplinary data from research institutions across the country. Datasets from over 90 academic, government, and research repositories are made discoverable in a bilingual interface that provides keyword and map-based searching.

- Information about datasets in the form of metadata records are harvested from repositories and made available for discovery.
- There are over 80,000 datasets from over 100 Canadian repositories and data collections currently indexed by Lunaris.

Controlled Access Management Initiative

The Controlled Access Management for Research Data Initiative aims to enable collaboration between Canadian data repositories and research institutions to enhance research data management and research security. This initiative is a pilot project that builds on the work of the Sensitive Data Repository Project.

Goals:

1) collaboratively innovate workflows, processes, training, and resources for depositing restricted access research data within data repositories and for granting access to other researchers; and

2) test technology for managing controlled access to research data within data repositories.

3. Research software

Research software

The research platforms or services that make up the Canadian DRI include a range of capabilities across the research lifecycle, including data acquisition and management, processing and visualization, storage and preservation, sharing and discovery.

The details of the Canadian research software landscape are described in detail in the [2022 Research Software (English/French) Current State Paper.](#)

Available software

A current list of the software available on our national systems is linked here: https://docs.alliancecan.ca/wiki/Available_software.

- This list changes frequently as new software is added. You can request the installation or updating of a particular program or library by contacting technical support.
- If you wish to use our software environment on your own system, please see Accessing CERN Virtual Machine File System.

It is worth noting that Humanities & Social Sciences-specific software is very limited, and there is a preference to support open-source software.

HSS-specific support

HSS-specific Support

- Humanities and Social Sciences DRI experts in **BC, Alberta, Quebec, and Atlantic Canada** (as of May 2025)

- **A National Humanities and Social Sciences team** with domain and technical experts located at sites across the country.

- **HSS targeted training courses** – including Software, Library, & Data Carpentry

- Support for **Digital Humanities Summer Institute** (DHSI), **DHSI-East**, and the **Canadian Certificate of Digital Humanities**

The HSS National Team (as of May 2025)

Team Lead, Humanities & Social Sciences Research Specialist, ACENET

Megan Meredith-Lobay, Associate Director, Research Computing Group at Simon Fraser University

Sarah Cameron-Pesant, HSS specialist (Calcul Québec)

Lina Marie Harper, HSS specialist (Calcul Québec)

support@ace-net.ca

Our dedication to HSS Support is also measured by:

Awareness	● Raising awareness with researchers within the range of target disciplines about what the Alliance offers them.
Internal Advocacy	● Working with other Alliance technical teams to ensure that infrastructure aligns with the Humanities and Social Sciences community's needs.
Training	● Training researchers on how to use National systems. ● Training Alliance staff on the methods and tools used by these researchers.
Support	● Supporting the community use of Alliance infrastructure and responding to their unique challenges.

HSS Training opportunities

1. **DRI organizations (to name a few):**
   a. SciNet
   b. Simon Fraser University Research Computing Group
   c. Sharcnet
   d. ACENET
2. **Canadian Certificate for Digital Humanities**
   a. Would benefit social scientists as well - many workshops are introductory and use humanities and/or social sciences examples & data
3. **Library & Data Carpentry workshops**
   a. Data Carpentry has subsection for social sciences

support@ace-net.ca

Training opportunities

4. **The HSS national team** offers virtual training events 2-3 times per year

    a.   Upcoming training in Python and web scraping in February

5. Digital Research Alliance of Canada **webinars** **& asynchronous** **training resources**

6. **Custom group training** and **one-on-one consultations**

7. Digital Research Alliance of Canada's new training discovery platform Explora (new in May 2025)

   -   Access to Explora is open, free and bilingual—find training in research computing, data management, software, cybersecurity and more.

DRI Usage Examples

Humanities & Social Sciences Researchers that incorporate supercomputers and/or cloud computing into their research

Cloud: Ian Milligan, Department of History, University of Waterloo

Project: [Web Archives for Longitudinal Knowledge (WALK)](#)

Project facilitates researcher engagement with the web archive collections being built by Canadian universities in partnership with the Internet Archive and their Archive-It subscription service.

ACENET
accelerate discovery

support@ace-net.ca

Cloud: Susan Brown, Department of English, University of Guelph

Project: [Canadian Writing Research Collaboratory (CWRC)](#)

Supports a number of humanities research projects through a hosting environment featuring tools that allow researchers to work collaboratively with different types of materials and resources, such as XML documents, bibliographic records, multimedia objects, or linked open data (LOD).

[Image source](#)

Cloud: Jon Saklofske, Department of English, Acadia University

Image source

Project: NewRadial

A web-based digital environment for humanities research and collaboration that encourages users to occupy, search, sort, and annotate database objects in a visual field.

Designed to function as a workspace in which primary objects from remote and locally-hosted humanities-related databases can be browsed, collected, curated, correlated, and augmented by multiple users in a dynamic visual environment.

support@ace-net.ca

Clusters: Marcia English, Human Nutrition, StFX

Project: [Food Matrix Interactions with Small Molecules](#)

Marcia was also a recipient of the Embedded Technical Support Program for her project on food matrix interactions with small molecules that influence flavour.

The support provided was to assist with simulation design, data analysis and data visualization, as well as preparation of a manuscript.

[Image source](#)

Clusters: Erin Mazerolle, Psychology, StFX

Project: Building Better Brain Imaging Analysis Tools

Erin uses ACENET's high-performance computing clusters to analyze data on her work with brain imaging data.

https://ace-net.ca/building-better-brain-imaging-analysis-tools.html