Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

Full protocol template for CERC HECW projects

Purpose of this CERC template:

This full proposal template is a guide for researchers to ensure that all required protocol headings and detailed components are included in your project proposal. This is important to adhere to research standards by providing a unified, customized template used by all CERC researchers for all projects. It is meant to provide flexibility for a variety of project types and yet it is based on Tri-Council Policy Statement and Canadian human research standards best practices and requirements. Please use the headings as you develop your full protocol.

Project title:

Principle Investigators and affiliations:

Protocol sections:

Project summary/Lay abstract

Background

Project aims and research questions

Methods:

- Community engagement plan
- Indigenous community plan
- Data management plan

Results:

- Knowledge mobilization

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

Project Summary / Lay Abstract (1/2 page):

Overview (1/2 page):

Please complete following table.

Project # assigned by CERC	TMU Project cost centre #	Is this project fully funded? By whom?	Date approved by CERC	Date that this proposal submitted to CERC for full review (before the REB)

- What HECW program theme does this project fit into?
    - Social determinants and Community Wellbeing
    - Health Equity and Accessibility
    - Indigenous Health and Disability
    - Digital Health and Technology
- How does this project align with CERC HECW program goals?

Project aims and research questions (5 pages):

- Background – issues; literature review supporting aims and objectives
- What is(are) the overarching research question(s)?
- What are the goals and objectives?
- What is the significance of this project?
- Why is this project important and how does it benefit the community/organization partner(s)?
- Who are the proposed partners for this project? Describe why and how they will be involved
- Why is this important to the Canadian context?
- Describe community partner engagement in project development
- Describe EDIA process, if applicable

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

Team expertise (1-2 pages):

- Most responsible PI Name and affiliation
- Co-PI Name(s) and affiliation(s)
- Names of HECW project team members and roles
- Name(s) of community/organization partner(s), if applicable
- Name(s) of people responsible for community engagement
- Primary contact name(s) and role(s) at the partner organization(s), if applicable
- Name of person primarily responsible for data collection, use and disclosure of project data
- Name of person responsible for data analysis
- Name of person responsible for creation of metadata

Name	Affiliation or Community Partner	Role in project	Expertise

Methods (5-10 pages):

- Type of project methodology
- Data sources and data types (briefly)
- Type of data analysis
- Types of data collected
- Who are the research participants?
- From where and how will participants be recruited?
- Will consent be needed? How will consent be collected and stored?
- Describe process for community partner engagement in data collection, use and disclosure?
- Who will be the REB of record?

Community partnership and engagement plan (1 page max) :

For each community partner, please describe the following (could be in a table format):

- Partner/organization name

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

- Purpose of their engagement
- Describe how you will engage with the partner throughout the project lifecycle:
  - Project Planning
  - Data collection
  - Analysis
  - Knowledge outputs
  - Knowledge sharing
  - Plan for building continued resource capacity
  - Potential for future engagement beyond this project
- What will be the demonstrated benefit(s) to the research partner(s)?
- Data re-use: has partner agreed to this and under what conditions?
- Will an MOU or agreement be in place with each partner? Describe the tenets of the partnership

Indigenous community partner engagement plan (1 – 1 1/2 pages)

- Review CARE, OCAP®, National Inuit Strategy on Research, the Metis Centre at the National Aboriginal Health Organization (NAHO), and TCPS 2, Chapter 9.
- Review CERC SOP RP-002-v1: Working with Indigenous Community Partners: Co-creation and research project approvals prior to and during research implementation [LINK to our SOPs]
- Please determine how the First Nation, Metis or Inuit community partner will be engaged, during each stage of the project/data management lifecycle for this project
- Please describe how the outputs will be used/disseminated, according to the wishes of the community partner.
- As part of your protocol and your **data management plan (next section)**, please describe the process for research self-governance and data sovereignty/data stewardship according to the principles specific to each community, using the reviewed links above.

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

EDIA Plan (1/2 - 1 page)

- How does the project align with EDIA principles (equity, diversity, inclusion and accessibility)?
- How will these principles be applied throughout the project lifecycle?
- Please use Tri-Council Policy Statement as a guide and SSHRC EDI best practice guidelines

Project Data Management Plan (DMP) – please see google form:

1. Data Storage & Infrastructure Compliance:

- This project uses the Digital Research Alliance of Canada (Alliance) infrastructure to meet strict data sovereignty standards. This infrastructure ensures compliance with the Tri-Council Policy Statement (TCPS 2), and when partnering with Indigenous data participants and communities, the principles of CARE, OCAP®, National Inuit Strategy on Research, the Metis Centre at the National Aboriginal Health Organization (NAHO), and TCPS 2, Chapter 9.

This infrastructure operates on secure servers physically hosted at Canadian universities, ensuring that all data remains exclusively under Canadian legal jurisdiction.

Accordingly, active research data will be stored and processed securely within this private cloud environment. Data will not be stored on local hard drives, personal computers, or commercial cloud platforms.

For the full technical security architecture, detailed infrastructure justification, and risk assessment, please refer to *Appendix A* attached to this protocol.

2. DMP

The documents below are resources to help you with the DMP. You may complete the DMP over time – the form will keep your responses and allow you to edit it as your think through your data collection, use and disclosure activities for your project. Once it is complete and fully submitted, you will receive a pdf of the data management plan. Please ensure that the information within the DMP aligns with the information in this protocol template.

Please append (or separately submit it) a pdf of your DMP as part of your full project protocol to the REB.

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

- Statistics Canada FAIR principles
- Chief Privacy Commissioner of Canada,
- Privacy by Design
- Terms relating to sensitive research data;
- Information about privacy preserving methods and de-identification;
- Global Training Centre courses on social research and data sharing;
- CIHR Research Data Management Module
- Tri-Council Policy Statement on Conduct of Human Research
- National Standards of Canada Conduct of Human Research
- Research Involving First Nations, Inuit, and Métis Peoples of Canada
- CERC SOP RP-002-v1: Research projects involving Indigenous Peoples of Canada [link to be added]

Results (2 pages):

- What are the expected data outputs (e.g., aggregated tables/graphs, narratives, qualitative analysis, artwork, academic paper, report, digital mediums, etc)
- Expected outcomes/results
- Describe the process for community/organizational input into interpretation, (if applicable):

Knowledge mobilization plan and output (1 page):

- What will be the modality of knowledge sharing [please use all that apply and describe]:
    - Academic journal
    - Open science publication
    - Social media
    - Investigative paper
    - Creative outputs
    - To be determined by community/organization partner
    - Podcast
    - Presentations

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

  - Other _____
- Describe process of capacity building (if applicable)
- How will your partners be involved in knowledge mobilization and sharing?

Feasibility (1 page):

Budget and resources (please complete table as below)

Please include costs for all resources required, include human resources (students, post-doc, analyst, data storage, honoraria etc).

Resource	Cost	Cost/FTE	Total cost	Source of funding

In addition, please write the justification for each resource and their costs.

Timeline

Please determine the estimated start and end dates for the project as a whole and its key phases/components. Please use the table below as a guide, or a Gantt chart.

Activity - output	Estimated time to complete	Estimated start date	Notes

Version History

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

Version number	Name of person completing this version	Date	Purpose (draft/edit/approval)

Appendices and Attachments (as needed)

Please attach other documents as needed, such as:

Team member certifications

REB letters of approval

Partner letters

Grant letters

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

Appendix A: Technical Infrastructure & Security

[INSTRUCTIONS: This appendix contains standardized language for you to use in your written protocol, as above. This refers to CERC data storage, encryption, and security. Please retain only the sections relevant to your specific methodology (e.g., Surveys vs. Focus Groups) and delete the others.]

This appendix describes how research data are collected, stored, secured, processed, and protected within the CERC HECW infrastructure.

1. Applicable Policies

Since this project involves capturing research data from human participants, the following policies, etc. apply:

- *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2 (2022)*
  https://ethics.gc.ca/eng/tcps2-eptc2_2022_chapter5-chapitre5.html
- *Tri-Agency Statement of Principles on Digital Management*
  https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management/tri-agency-statement-principles-digital-data-management

2. Data Collection, Use, and Disclosure Processes

2.1 DATA COLLECTION (Choose the chunk(s) required (one or more))

2.1.1 SURVEYS (REDCap)

This project will collect information from human participants using a virtual server using online survey software from Vanderbilt University named REDCap. The research information will be collected via survey. Only needed information from and about participants will be requested. The information will be collected directly from the individuals who are the subject of the study unless the individual expressly requests support from a research team member or a personal support partner to help input their information. Even then, only information permitted by the participant should be included. The information collected by this project will be stored in the HECW active data store, a

private, secure, AES encrypted storage volume hosted on Arbutus Cloud that stores the active research data, as described in the HECW data management plan.

2.1.2 FOCUS GROUPS / INTERVIEWS — IDENTIFIERS RETAINED

This project will collect information from human participants using Focus Group interviews hosted on Zoom video conferencing platform. The recordings of the meeting will be saved locally and will be deidentified upon transcription. Pseudonymized transcripts obtained from Focus groups / interviews will be stored in the HECW active data store described in the HECW data management plan.

The identifiers retained for reference purposes will be recorded in a separate linking sheet that connects participant identifiers to assigned pseudonyms. This linking sheet will be encrypted by the research team using GNU Privacy Guard (GPG) with asymmetric encryption (public/private key pairs) and the encrypted file will be stored in the /project directory of the Narval HPC cluster, operated by Calcul Québec under the Digital Research Alliance of Canada at École de technologie supérieure in Montréal. Storing the encrypted linking sheet on Narval, distinct from the pseudonymized transcripts housed in the HECW Active Data Store, provides both physical and logical separation. The linking sheet (if used) will be retained only for the minimum period required for data verification or participant withdrawal requests, after which it will be securely destroyed.

2.1.3 FOCUS GROUPS / INTERVIEWS — NO IDENTIFIERS RETAINED

This project will collect information from human participants using Focus Group interviews hosted on Zoom video conferencing platform. The recordings of the meeting will be saved locally and will be deidentified upon transcription using [DESCRIBE METHOD USED FOR TRANSCRIPTION]. All recordings, and all documents and transcripts containing identifiers will be destroyed after transcription and de-identification.

The deidentified transcripts will be stored in the HECW active data store described in the HECW data management plan.

2.2. DATA USE

The information collected by this project will be used for analysis purposes, based on the approved project protocol, and participant consent, and for no other purposes. The information will be used for the purposes of [PRIMARY PURPOSE] (e.g., "evaluating the experiences of [POPULATION] in accessing [SERVICE/SYSTEM]"). It will also be used to

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

[SECONDARY PURPOSES - if applicable] (e.g., "inform policy recommendations, develop best practice guidelines").

2.3. DISCLOSURE OF INFORMATION

Aggregated and summary information collected by this project will be disclosed in our research reports and articles. De-identified information will be made available to third parties who request access to our data for the purposes of analyzing our research methods and relating our conclusions. Personal information will not be disclosed to any third party.

If these data collection, use, and disclosure processes are adapted between when the surveys and other research tools have been designed, reviewed, and finalized then this section will be updated and reviewed. That is, any changes to the data management of the research data will need to be re-approved.

3. SAFEGUARDING INFORMATION

The following diagram illustrates the architecture which will be used to collect and manage research data:

[SELECT APPROPRIATE DIAGRAM BASED ON YOUR PROJECT TYPE]

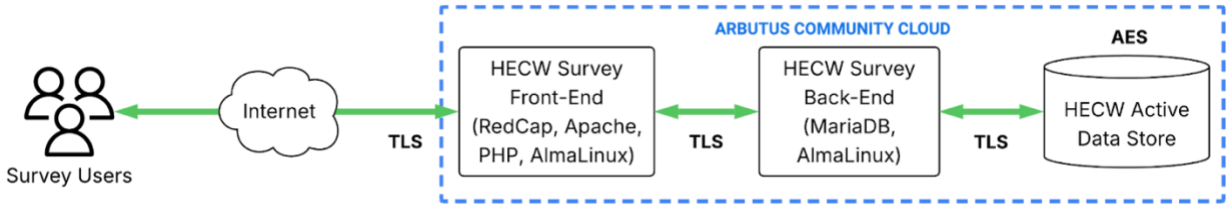

Fig. Project Architecture (Surveys Only using Redcap)

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

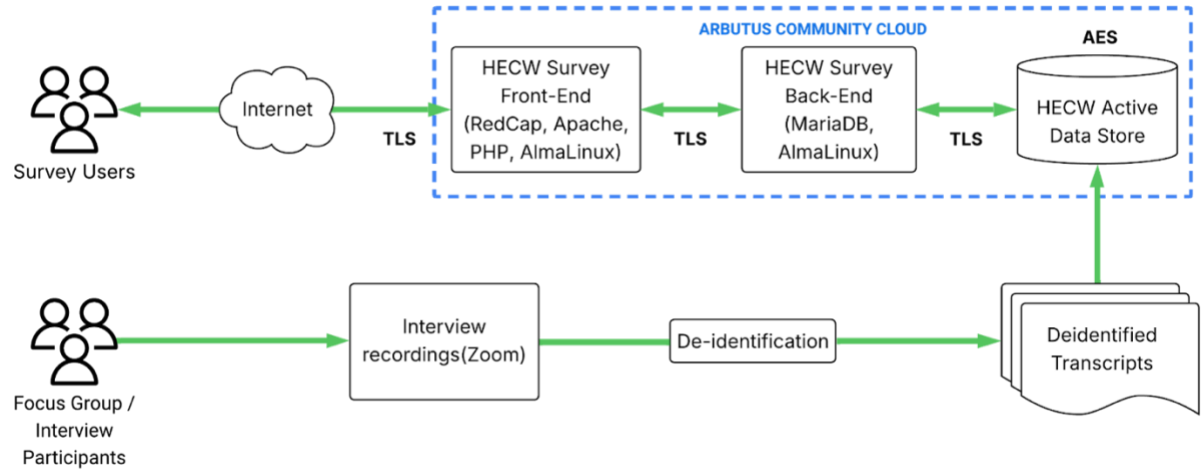

Fig. Project Architecture (Surveys + Focus Groups/Interviews - Identifiers Destroyed)

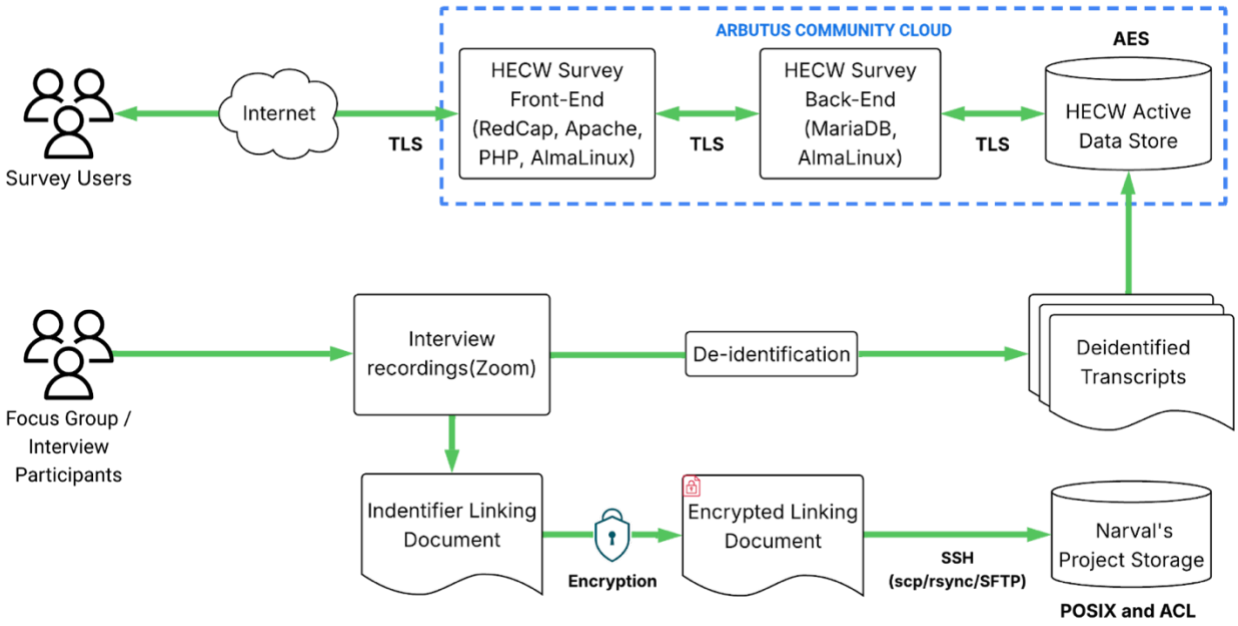

Fig. Project Architecture (Surveys + Focus Groups/Interviews - Identifiers retained)

Project number (provided at time of short form approval by CERC):

Date of receipt by CERC:

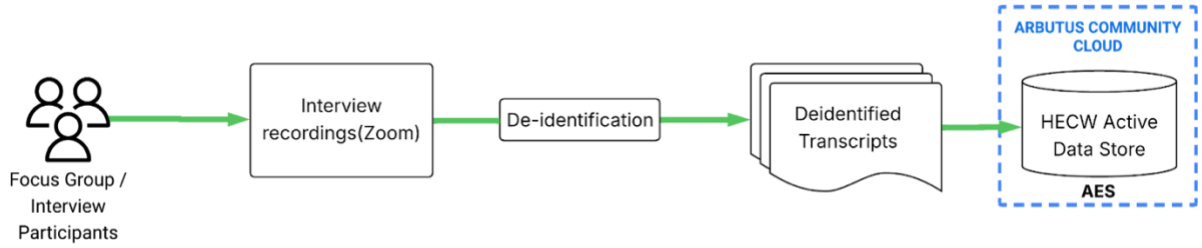

Fig. Project Architecture (Focus Groups/Interviews Only - Identifiers destroyed)

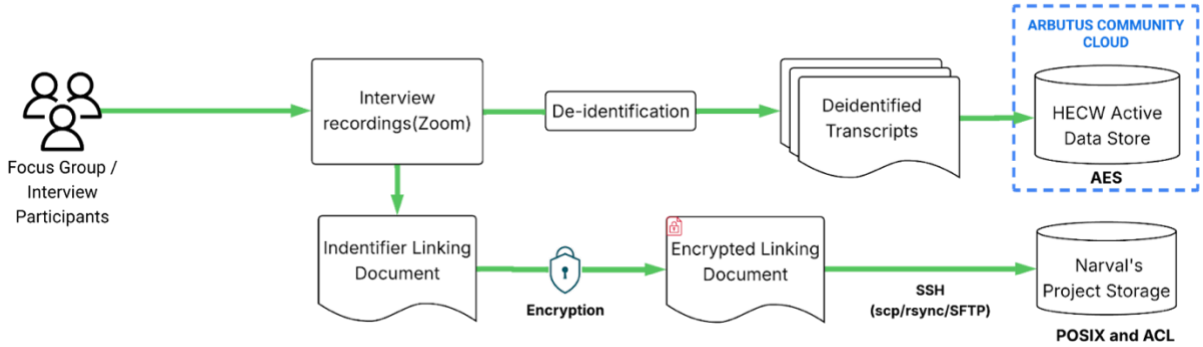

Fig. Project Architecture (Focus Groups/Interviews Only - Identifiers retained)

Each architectural component is described below in terms of how appropriate measures have been taken to protect sensitive research information:

3.1 Physical security

The infrastructure is housed at an Innovation, Science and Economic Development (ISED)/Digital Research Alliance of Canada funded data centre located at the University of Victoria. This data centre has physical security and environmental controls as required for an enterprise-grade data centre, reviewed, approved and governed by the Digital Research Alliance of Canada in accordance with the Physical Security of Data Centres Standard (SEC-08). The data centre is windowless with robust floor to ceiling walls. Access to the data centre is limited by access card. 24x7 alarm monitoring is provided by UVic Campus

Security. The data centre has redundant power supplies, cooling systems, and a backup electrical generator for high availability.

3.2 Infrastructure

The project infrastructure runs on virtual servers on the Arbutus Community Cloud. This cloud service runs on compute and storage hardware dedicated to research. This infrastructure is managed by the Alliance-funded Research Computing Services team who are responsible to maintain the Arbutus Cloud infrastructure to professional standards.

3.3 Network

The dedicated research network is provided by BC Net and Canarie. There is a dedicated peering relationship with academic institutions all over the world, which allows data to be transmitted directly between research institutions via a private network, bypassing the public internet and thereby ensuring secure data transfer. And, there is a commodity Internet service provided by Rogers Telecommunications. Firewall rules called 'security groups' have been applied so that only necessary network ports are available to the open Internet and privileged ports are limited to known networks.

3.4 Operating Systems

The HECW virtual servers are running a patched and hardened version of AlmaLinux 9. They have been configured so that no unnecessary services are active. Administrative access is via Secure Shell (SSH) from known networks. SSH access requires asymmetric encryption keys, basic password access to SSH is not permitted. These software on these virtual servers will be patched and maintained on a regular basis.

3.5 Front-end software

REDCap is the front-end software used for surveys. REDCap runs on PHP and Apache. This software stack is up-to-date with the latest versions and will continue to be patched and maintained. A software agreement is in place with the REDCap Team at Vanderbilt University. End-user and privileged user access to REDCap is managed using role-based access controls.

3.6 Back-end software/Database

mariaDB is the back-end software used to house the REDCap configuration and the survey data. mariaDB was selected as a secure substitution for MySQL. The database has been configured to limit access to privileged connections to the local host and not permit

administrative connections over the network. Databases housed on the virtual database server have been protected with role-based access control.

3.7 Data in transit

Data in transit is encrypted with Transport Layer Security (TLS). This includes in transit between the Internet and the front end plus also in transit between the project's REDCap front-end and the back-end in mariaDB.

3.8 Data at rest

Data in the HECW active data store is encrypted with a symmetric Advanced Encryption Standard (AES) encryption key. The Storage Area Network (SAN) utilizes Redundant Array of Independent Disks (RAID) technology for high availability. Every unit of storage is written in two or more locations. Checksums are used to check and repair storage integrity issues (e.g. 'bitrot', disk failure). The disk and nonvolatile memory express (NVMe) storage media in the Arbutus service are shredded when they are decommissioned and not returned to the vendor when failure-related warranty issues occur.

3.9 Backups

Encrypted backups will be extracted from the HECW active data store and stored at a storage facility located at the ISED/Digital Research Alliance of Canada - funded data centre located at McGill University and operated by Calcul Quebec.

3.10. Secure Storage for Identifier Linking Sheets (Use ONLY if identifiers retained)

Narval is a general-purpose HPC cluster located at École de technologie supérieure in Montréal and operated by Calcul Québec under the Digital Research Alliance of Canada. Alliance resources follow academic best practices for integrity, confidentiality, and availability.

Access to the Narval cluster is via Secure Shell (SSH) with Alliance account credentials. The Alliance uses standard POSIX and ACL permissions on cluster filesystems. Access to encrypted linking sheets on Narval will be restricted to essential research personnel only through file permissions and access control lists. Identifier linking sheets will be encrypted using GNU Privacy Guard (gpg) by the research team before storage on Narval.

The infrastructure is compliant with the Digital Research Alliance of Canada's Physical Security of Data Centres Standard (SEC-08). The facility is a fully enclosed, windowless perimeter with floor-to-ceiling reinforced walls. Access is strictly controlled via electronic credentials and subject to 24/7 alarm and video monitoring. Storage devices removed from

Alliance systems due to hardware failure are either destroyed or already encrypted or erased before being returned to the vendor for replacement.

This separation of de-identified transcripts (stored in the HECW Active Data Store on Arbutus) from encrypted identifier linking sheets (stored on Narval) provides an additional layer of privacy protection.

4. Data Management Risk Assessment

The *TCPS 2 (2022)* requires that privacy risks and threats to the security of information for all stages of the research life cycle be identified. The research stages from the Tri-Council's *Research Data Management Policy* are:

- Data creation (DC)
- Processing (Pro)
- Analysis (A)
- Preservation (Pre)
- Storage and access (SA)
- Sharing and reuse (SR)

The privacy risks and threats identified for this project are:

[This risk assessment aligns with the standard CERC Data Management Plan (DMP). Ensure that the risk section in your project specific DMP document matches the table below.]

Risks	**Research Stage**	**Assessment of Risk**
Harm to individuals, legal liability, sanctions, and harm to reputation of researchers/institutions if personal information is improperly disclosed	DC, Pro, A, Pre, SA, SR	**Low**: with the safeguards and controls implemented for the project the risk of disclosure is low
Inability to meet project objectives due to loss of project data	Pro, A, Pre, SA	**Low**: with infrastructure being highly available and the off-site backup this risk is low
(USE ONLY IF IDENTIFIERS RETAINED)	Pre, SA, SR	**Low:** Focus group and interview transcripts are de-identified and stored in ADS. If a linking sheet is retained, it is encrypted and stored

Unauthorized re-identification of participants through linking de-identified data with identifier keys		separately on Narval, ensuring that re-identification would require unauthorized access to two distinct, geographically separated systems.