

Research Data Guidance for REB Applicants

Introduction

Open Science

Storing your data in a repository makes it easy for other researchers, publishers, and collaborators to find, use, and cite your data. It supports the open science movement goals of transparency, accessibility, and collaboration, and can increase attention to your work.

[Borealis](#) is free and available to VCC employees. The platform makes it possible for researchers to deposit data, create appropriate metadata, obtain DOIs for permanent links, and maintain version control of their datasets. All files are held in a secure environment on Canadian servers. While you are encouraged to make your data available publicly, you can choose to restrict access to the data if you wish.

Funding

Moreover, the Tri-Agency in Canada (NSERC, CIHR, SSHRC) have introduced new policies which will soon require grant recipients to deposit research data in repositories for a period of 10-years following publication or the funding period end date. Researchers must provide availability statements on accessing data sets, in all publications resulting from funded research. Researchers at VCC can deposit data sets with Borealis: The Canadian Dataverse Repository.

The Tri-Councils additionally require that all findings be made available (outside a pay wall) within 12 months of publication of results. Researchers at VCC can make publications openly available through the [VCC Digital Repository](#).

Security of Data & Recommendations

- Use a secure, locked location for hard copies of your data (this includes paper copies of consent forms, field notes, etc.)
- Ensure you have password protection and encryption for all computer files and back-up files.
- Use of cloud storage requires information about user agreements and password protection. At VCC we generally recommend SharePoint/Teams. This allows for

role-based access, and folder or document-level access restrictions, version history, logs, and data are stored within Canada.

- Use of any non-FIPPA compliant cloud storage platform must be explained to participants in the consent form, and approved by the REB.
- Transcription of video and audio files is best done

Confidentiality of Data

- Confidential information must not be collected or exchanged via e-mail unless the information has been encrypted (e.g., documents are password-protected, or via Outlook).
- Web-based questionnaires must use encryption software.
- VCC has Survey Monkey and MS Forms licenses which can be configured to collect responses anonymously as wished.
- The use of synthetic data requires strict ethical oversight and is subject to an additional application form. Researchers must transparently justify its use and explain how the data was generated to avoid perpetuating biases, ensuring compliance.

Specific Data Collection Techniques & Confidentiality

Focus groups

- Include a statement in your consent form such as:
"We encourage all participants to refrain from disclosing the contents of the discussion outside of the focus group; however, we cannot control what other participants do with the information discussed."

Interviews

- Interview participants may prefer to have their comments attributed in publications rather than to remain unnamed. However, if disclosure of one participant's identity may jeopardize the welfare of other participants who wish their identities to remain confidential, it may not be appropriate to attribute comments to those wishing to be identified in publications or dissemination of research results (see TCPS 2 Articles 5.1 and 10.4).
- If you are planning to disclose the identity of participants, this should be specified in the consent process.

Video/Audio

- Consider whether video or audio recording is needed. Is audio sufficient? Will you destroy the video/audio recording after the transcript is prepared?
- Consent must state the plans for the recordings, and if they are not destroyed, it should be clear that the identity of participants will not be anonymous.

Transcripts

- When creating transcripts from audio or video files any tool in the Microsoft 365 environment is recommended. ClipChamp is also an option but be aware that it often also stores files locally.
- Review Terms of Service/Use to ensure the security of the content you input before using transcription tools and be sure to review output for accuracy.

Describing Data Collected

When it comes to the collection of data, please ensure that you are very specific in describing the data you are collecting. You may wish to utilize the definitions below.

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

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 risk of re-identification 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.

Source: TCPS2 (2022) Chapter 5

Please also refer to this Research Data Management in the Canadian Context, [Chapter 13](#), for a discussion of indirect and direct identifiers, as well as information about sensitive data.

Future Use and Confidentiality

Informed consent must specify how data will be handled during the study and for future use. If datasets will be published, researchers must clearly explain how participant identity will be protected.

Two general approaches include:

De-identification involves removing or masking direct identifiers (e.g., names, contact info) and reducing indirect identifiers that could be used in combination to re-identify individuals (e.g., postal codes, birthdates).

Anonymization goes further by ensuring that re-identification is no longer reasonably possible, even when combining datasets.

Source: [Data De-identification | Research Data Management](#)

Access to Data

- The principal investigator is responsible for the collection, maintenance and retention of research data.
- Limit access to data to only those necessary in terms of raw data, consent forms, and lists of participants
- All personnel who are associated with a research project and who will have contact with the research participants and/or data are required to complete the TCPS online tutorial 'CORE' (Course in Research Ethics') before the ethics application will be approved. Copies should be retained by the Principal Investigator and should be available on request. The tutorial is free. It takes a minimum of three hours to complete and can be found at <http://tcps2core.ca/welcome>.

Storage of Data

Researchers have a duty to keep complete and accurate records of data, methodologies and findings, including graphs and images, in a manner that will allow verification or replication of the work by others.

This includes:

- consent forms and research communications,
- recording all primary data in clear, adequate, original and chronological form,
- retention in a repository from which it cannot be removed.

These records should be in sustainable file formats (PDF/A, TIFF, CSV, etc.)

Records of data should normally be retained in the department in which they are produced for at least seven years after the work is published or otherwise presented.

Researchers must specify how information and data will be disposed of at the end of retention, or once the period of open publication has come to an end.

Research Supported by Tri-Council Agency Funding (SSHRC, CIHR, NSERC)

Researchers and trainees (students and postdoctoral fellows) are required to retain all digital research data, data documentation and code arising from agency-supported research for a period of 10 years post end of grant, award or publication, whichever comes last.

All deposits must include a README file. At minimum, quantitative data must include explanation of variables, codes, and abbreviations, as well as statistics, anonymization procedures, and links to output. At minimum, qualitative data must include interview protocols, transcription conventions, and other contextual information.

Participants and Data

Consent Forms & Future Use

Please include the following details in the consent form. Use non-technical/lay language suitable for the intended recipients:

1. Describe what open access means, i.e. who will have access, and where/how data will be stored.

2. Describe the data that will be made available, e.g. that identifiers will be removed.
 - Do you have a data sharing statement indicating where your data can be found, and under what conditions it can be reused?
 - For example, with Borealis you may want to choose one of the two most common licenses: [CC License 1.0 Universal](#) or [CC License 4.0 Universal Use with Attribution](#)
3. Acknowledge (if applicable) that opening access to data has the potential for increasing participant risk.
4. Explain that once the data is made available, the participant will not be able to withdraw their data.

Withdrawing

Generally we advise participants that they may withdraw any time from a study without consequence. But when collecting various kinds of data, it is important to spell out when this is possible, and when this is no longer possible in your consent form.

Example (Focus Group)

During the Session: Participants may leave the focus group at any time.

After the Session: Participants have up to seven days following the focus group to contact the facilitator to withdraw, revise, or omit their specific comments.

Data Removal: If you withdraw from a focus group we must remove your contributions from the study. Your survey data however is anonymous so it cannot be withdrawn after you submit your answers.

Choosing not to participate or withdrawing from the study will not affect your professional relationships or standing with Vancouver Community College.