Title: Use of Cloud Storage for Research Data Involving Human Participants

1.0 PURPOSE

The Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2 2022) establishes Concern for Welfare as a core principle. A key component of this principle is ensuring the privacy and secure control of participant information throughout its lifecycle. Cloud storage solutions offer researchers a convenient and often cost-effective method for securely housing large datasets, but their use introduces specific risks related to data security, privacy, and sovereignty.

This document is a recommendation from the Research Ethics Board (REB) - it is not a mandated university policy. This document guides researchers at Lakehead University who choose to use cloud storage services for housing their research data involving human participants. It details procedures for selecting appropriate platforms and implementing safeguards to protect participant data. **This guidance does not apply to researchers storing all data on local, physical, or university-managed hard drives or servers that are not internet-accessible.**

This guideline should be applied to all human participant research data, regardless of classification level. Additional considerations may apply when working with First Nations communities, particularly in relation to OCAP® (Ownership, Control, Access, and Possession) principles. **If a researcher suspects they are handling Level 3 (high-risk) data, they must contact the Research Security and Data Management Specialist** (security.research@lakeheadu.ca | rdm.research@lakeheadu.ca; (807) 343-8010 ext. 8190)**, before storing their data in the cloud.**

2.0 DEFINITIONS

- **Research Ethics Board (REB):** The body of researchers, community members, and others with specific expertise established by Lakehead University to review the ethical acceptability of all research involving humans conducted within Lakehead University's jurisdiction or under its auspices.
- **Cloud Storage Solution:** An online service that allows for the storage, management, and access of research data over the internet, rather than directly on a local computer (e.g., Google Drive, Sync.com, Microsoft OneDrive).
- **Personal Information (PI):** Any information about an identifiable individual.

  Note: Data is often considered PI even if direct identifiers (like names) are

removed, especially if it can be re-identified using an identifier list or by linking it with other publicly available data.

- **Data Encryption:** The process of converting data into a code to prevent unauthorized access. Data can be encrypted "at rest" (while stored) and "in transit" (while being uploaded or downloaded).
- **Anonymous information:** "information never had identifiers associated with it (e.g., anonymous surveys) and risk of identification of individuals is low or very low" (Chapter 5, Section A).
- **Anonymized information**: “the information is irrevocably stripped of direct identifiers, a code is not kept to allow future re-linkage, and risk of reidentification of individuals from remaining indirect identifiers is low or very low” (Chapter 5, Section A).
- **De-identified Information:** Confidential information from which direct identifiers have been removed, but a code or key may exist that could allow re-identification in the future (Chapter 5, Section A).

3.0 RESPONSIBILITIES

Researchers have an ethical duty to protect participant information (Article 5.1, TCPS 2 2022). This includes assessing potential privacy risks in data storage, implementing robust security safeguards such as password protection, encryption, and access controls, and being transparent with participants about where their data will be stored and who may have access to it.

Data stored with any online service provider may be subject to various laws and jurisdictions depending on the server location and the provider's ownership. All online storage introduces a potential risk of access by external entities, regardless of where the data resides (NIST 2011). This general risk should be disclosed to potential participants.

4.0 PROCEDURES FOR CLOUD STORAGE SELECTION AND USE

4.1 Selecting a Cloud Storage Provider

Lakehead University’s preferred and institutionally supported cloud storage platform is Google Drive. This platform offers enhanced security protocols under our institutional agreement with Google.

Lakehead University does not support the use of Microsoft OneDrive.

Researchers are encouraged to assess the **class/type of data** they are working with using the **Data Classification Tool**. If researchers have questions about their data's sensitivity, researchers should contact Andrew Austin, Research Security and Data Management Specialist (security.research@lakeheadu.ca | rdm.research@lakeheadu.ca; (807) 343-8010 ext. 8190).

Most importantly, researchers must ensure that participants are fully informed of any privacy risks associated with the chosen platform.

4.2 Data Management Best Practices

For research data management best practices, visit our cybersecurity webpage. Regardless of the platform chosen, researchers should adhere to the following practices:

- **De-identify Data:** Before uploading any research data to a cloud service, you must remove all **direct identifiers** (e.g., names, specific addresses) and replace them with a unique code or pseudonym. This process results in **de-identified data**.
    - The **master list** that links these codes back to the personal identifiers must be stored **separately** from the research data and secured locally (e.g., on an encrypted, local university computer).
    - To minimize risk, researchers can **anonymize the data**. This occurs when the master list linking the codes to the participants is **permanently destroyed** after the need for linkage has passed, making re-identification practically impossible and thus lowering the privacy risk (**TCPS 2 2022**).
- **Use Strong Passwords & Multi-Factor Authentication (MFA):** Protect the account used to access the cloud storage with a strong, unique password and enable MFA if possible.
- Prioritize Platforms with **End-to-End Encryption (E2EE):** Where data is sensitive, choose providers that offer E2EE, meaning the data is encrypted on the researcher's device *before* uploading and can only be decrypted by the research team. Please contact the Research Security and Data Management Specialist (security.research@lakeheadu.ca | rdm.research@lakeheadu.ca; (807) 343-8010 ext. 8190) if you believe your research data is sensitive.
    - For example, Lakehead University's default Zoom meetings are at an "enhanced" privacy setting. If you require your Zoom meeting to be encrypted, please contact the Research Security and Data Management Specialist.

    - If researchers are using a physical storage device, they must ensure that the data is properly encrypted and this responsibility lies with the researcher.
- **Manage Access:** Provide data access only to authorized research team members. Use the platform's tools to set permissions (e.g., "view only" vs. "edit") and regularly review access lists.
- Researchers are required to retain research data for a **minimum of seven (7) years as per Lakehead University's policy** following the completion of the study. This applies to all types of data, including those stored using online cloud platforms.

5.0 INFORMED CONSENT REQUIREMENTS

The information letter and consent form must clearly describe where the research data will be stored and outline any potential risks associated with that storage. Importantly, researchers should distinguish between **anonymized** and **de-identified** data when describing the state of the data being stored (Canadian Institutes of Health Research et al., 2022). *Anonymized information* refers to data from which all reasonable risk of re-identification has been removed, and the process is irreversible. *De-identified information* refers to data that has had identifiers removed but can still be re-linked using a code key. Researchers must ensure that the consent language in their materials accurately reflects the appropriate term for the data.

The following descriptions are suggested language to include in both the information letter and the consent form bullet points. Please customize the wording to accurately reflect the specific details of your study.

A. For Storage on the Lakehead University Google Drive:

Scenario 1: Level 3 Data (Most Secure)

> "Research data will be manually encrypted and uploaded to the secure cloud server. This data will be **stored** on my **two-factor authenticated**, password-protected Lakehead University Google Drive, which operates under an institutional agreement with enhanced security."

Scenario 2: Level 2 Data (No Manual Encryption)

> "Research data will be uploaded to the secure cloud server, which encrypts data in transit and at rest. This data will be stored on my two-factor authenticated, password-protected Lakehead University Google Drive, which operates under an institutional agreement with enhanced security."

Details on these steps can be found on the Research Services Cybersecurity Webpage

B. If data will be deposited in a repository or retained long-term beyond the LakeheadU required 7 years:

"Your [*select one: de-identified/anonymized*] data may be stored in a secure data repository (e.g., Lakehead University Knowledge Commons) for future research use. Access to this data may be open or limited depending on repository policies. All efforts will be made to protect your privacy, and no identifying information will be included. Please note that any future use of your data may be for research purposes that are entirely different from the original purpose for which it was collected."

For further resources on research ethics in data sharing and research data management, please visit the Metaresearch & Open Science Program's resources.

C. When using institutionally approved online tools/services, the information/consent letter for participants must contain the following statement:

> Because this study uses online tools (*e.g. Google Drive*), there is a small risk that your data could be accessed by third parties, including government authorities, under applicable privacy and national-security laws. While all reasonable efforts will be made to protect your confidentiality and anonymity, absolute confidentiality cannot be guaranteed when data is transmitted or stored online. With your consent to participate in this study, you acknowledge this potential risk.

7.0 REFERENCES

- Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada. (2022). Privacy and confidentiality [TCPS Interpretations]. Panel on Research Ethics. https://ethics.gc.ca/eng/policy-politique_interpretations.html
- Digital Research Alliance of Canada. (n.d.). DRAC/Alliance. https://alliancecan.ca/en
- Government of Canada. (2021). 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
- Government of Canada. (2022). Bill C-27: Digital Charter Implementation Act. https://parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading
- Lakehead University. (2023). Institutional Research Data Management (RDM) Strategy/Policy. https://www.lakeheadu.ca/research-and-innovation/research-services/research-data-management-rdm-institutional-strategy-2023
- Metaresearch & Open Science Program. (n.d.). Ethics and harmonized consent language. Journalology Training. https://journalologytraining.ca/ethics-and-harmonized-consent-language/
- National Institute of Standards and Technology. (2011). Guidelines on security and privacy in public cloud computing (Special Publication 800-144). https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-144.pdf
- Panel on Research Ethics. (2022). Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2 2022). https://ethics.gc.ca/eng/tcps2-eptc2_2022.html

8.0 ACKNOWLEDGEMENT

This SOP was developed with the assistance of AI tools (Google Gemini and Grammarly) to support clarity, consistency, and alignment with current best practices.

Author: Research Ethics Facilitator and Research Security and Data Management Specialist

REB Chair: Dr. Monica Flegel

Date of approval: February 27, 2026