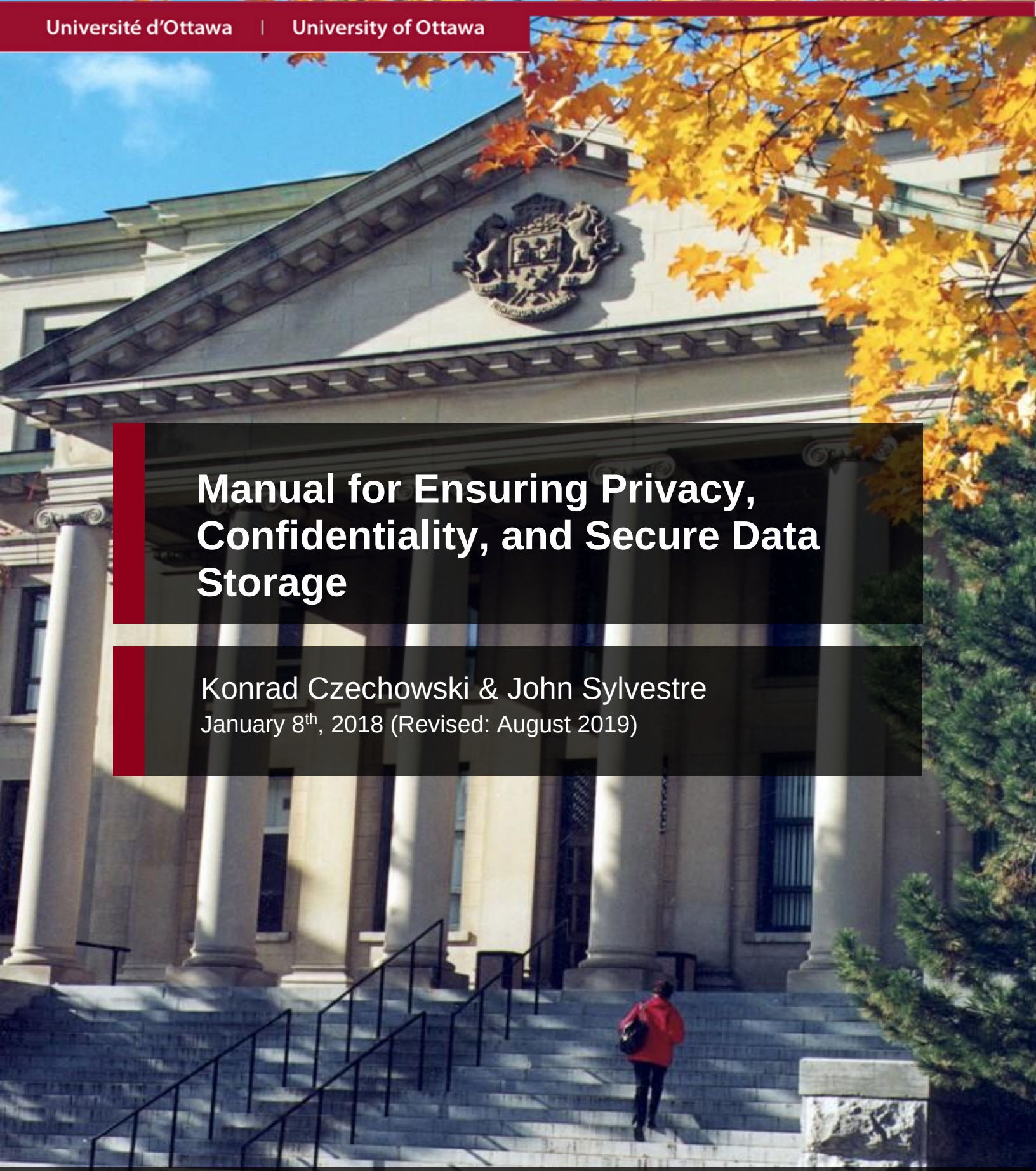

Manual for Ensuring Privacy, Confidentiality, and Secure Data Storage

Konrad Czechowski & John Sylvestre

January 8th, 2018 (Revised: August 2019)

Centre for Research on Educational and Community Services (CRECS)
613-562-5800, Ext. 1856 | crecs@uottawa.ca
crecs.uOttawa.ca

uOttawa

Table of Contents

Acknowledgements

We would like to thank Catherine Paquette and her team at the Office of Research Ethics and Integrity, our librarians Susan Mowers, Jessica McEvan and their team, and the uOttawa security architect Sandeep Gupta for their invaluable judicious feedback and comments on earlier versions of this document.

Introduction

This document outlines the Centre for Research on Educational and Community Services' (CRECS) recommended procedures for collecting, safeguarding, and working with confidential data. As a research centre located at the University of Ottawa, CRECS must hold itself to the highest standards for protection of personal information and confidential information. CRECS researchers, students and postdoctoral fellows regularly have access to confidential information such as personally-identifiable information, health information, and other kinds of non-public information.

As a research centre, it is essential that CRECS be able to assure both its institutional partners and research participants of the safety and confidentiality of the data they provide. The smallest lapse in vigilance may have consequences for the participants, CRECS, the University of Ottawa, or our community and institutional partners. It is the responsibility of everyone at CRECS to treat data security as a vital part of their work.

This manual begins with a rationale for maintaining privacy, confidentiality, and data security. Then, there is a brief description of how information for this manual was collected. Next, a review of applicable provincial and federal privacy legislation as well as professional codes of conduct is presented. Based on this review, a list of recommendations regarding privacy, confidentiality, and data security are provided.

General Principles

- These recommendations for collecting, safeguarding, and working with confidential data apply to all CRECS personnel, including professors, employees and research associates, visiting researchers, postdoctoral fellows, students, and trainees. Professors and others who have roles in supervising other CRECS personnel and students have the responsibility to ensure that the persons they supervise are aware of, understand, and follow these recommendations.
- Ethics approval from the University's Research Ethics Boards (REB) is required for studies involving human participants; researchers must obtain ethics approval before beginning any research activities involving humans.
    - Any research project involving direct data acquisition from human participants (e.g., through interviews, focus groups, surveys, etc.) must be reviewed and approved by the REB.
    - Projects whose primary purpose is quality assurance (such as program evaluation) and where there is no intention to publish the findings in a peer-reviewed journal may not require REB approval. However, the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (TCPS 2) guidelines must still be followed. Researchers should consult with the Office of Research Ethics and Integrity (OREI) to determine if REB review is required for their project.
    - Projects involving secondary data use may require REB approval depending on the nature and source of the data. Researchers should consult with the Ethics Office to determine if REB review is required for their project.

    - Further information can be obtained at: https://research.uottawa.ca/ethics/

- The Library provides public-use secondary data access and consultation services to the University of Ottawa community, and related services such as low-risk access to confidential Statistics Canada data through a Real Time Remote Access subscription service, as a complement to a secure Research Data Centre located in the Library. Please contact the library for more information.

Method

The following steps were undertaken for developing the manual:

1. Conducted an online search for terms such as “policy data security”, “privacy policy” and “data security”.
2. Reviewed relevant ethical codes both nationally and by discipline.
3. Reviewed the websites of Privacy Commissioners of Ontario and Canada.
4. Consulted with librarians specializing in data management.
5. Consulted with the uOttawa data security architect.
6. Consulted with the uOttawa legal counsel and the Office of Research Ethics and Integrity.
7. Presented draft findings to University researchers, students, and research ethics protocol officers.

Findings

Applicable Legislation and Professional Codes of Ethics

Legislation and professional codes of ethics were reviewed for information specific to privacy, confidentiality, and data security relevant to CRECS-related research. The following are findings from a review of six sources that are relevant to the research activities conducted at CRECS. Although CRECS researchers may also belong to foreign professional bodies (e.g., American Evaluation Association), only the most relevant Canadian sources were included in this section of the document. Other ethical principles may apply to work at CRECS and each CRECS researcher should be familiar with the full ethical codes that are applicable to them. This manual is intended to summarize the information most relevant to privacy, confidentiality, and data security in situations commonly encountered by CRECS researchers. It is not intended to replace, supersede, or in any other way override the more complete legislation or ethical codes.

Accordingly, we summarize the following: (i) the *Personal Information Protection and Electronic Documents Act*, (ii) the *Personal Health Information Protection Act*, (iii) the *Tri-Council Policy Statement: Ethical conduct for research involving humans*, written on behalf of the three primary funding agencies in Canadian research, (iv) the ethical guidelines from the Canadian Evaluation Society, (v) the ethical guidelines from the Canadian Psychological Association, and (vi) the ethical guidelines from the College of Psychologists of Ontario.

Personal Health Information Protection Act (PHIPA) and *Personal Information Protection and Electronic Documents Act* (PIPEDA)

PIPEDA is a 5-part federal statute covering the management of personal information. PHIPA is an Ontario provincial statute that has been declared “substantially similar” with respect to health information custodians to the federal statute (PIPEDA). As such, health information custodians need only comply with PHIPA for collection, use, and disclosure of personal information that occurs within the Province of Ontario (Privacy Commissioner of Canada, 2013). PIPEDA is still applicable to health information custodians within Ontario in some situations, such as disclosure of health information to agents outside of Ontario, to personal information that is not health related, and to activities covered in Parts 2-5 of PIPEDA (e.g., use of electronic documents; College of Psychologists of Ontario, 2004; Information and Privacy Commissioner of Ontario, 2015).

PIPEDA has been summarized into the following 10 “fair information principles” (College of Psychologists of Ontario, 2013; Office of the Privacy Commissioner of Canada, 2015):

1. **Accountability:** The organization [University of Ottawa] is responsible for the personal information under its control and shall designate an individual to be accountable for PIPEDA compliance.

2. **Identifying Purposes:** The purposes for collecting personal information shall be identified at or before the time of collection.

3. **Consent:** Informed consent of the individual is required for the collection, use or disclosure of their personal information, unless otherwise inappropriate.

4. **Limiting Collection:** Amount and type of personal information collected shall be limited to that which is necessary for the purposes identified.

5. **Limiting Use, Disclosure, and Retention:** Personal information shall not be retained longer than needed to fulfill the purpose it was collected for and shall only be disclosed or used for that purpose, except with the consent of the individual or as required by law.

6. **Accuracy:** Personal information shall be kept as accurate, complete, and up-to-date as is necessary for the purposes for which it was collected.

7. **Safeguards:** Personal information shall be protected by security safeguards appropriate to the sensitivity of the information.

8. **Openness:** Specific information about policies and practices relating to the management of personal information shall be made readily available to individuals. This can include contextual information about methods and purposes of the studies in which individuals participated.

9. **Individual Access:** Upon request, an individual shall be informed of the existence, use and disclosure of his or her personal information and shall be given access to that information. An individual shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate.

10. **Recourse (Challenging Compliance):** An individual shall be able to address through easily accessible procedures, a challenge concerning compliance with the above principles to the individual accountable for PIPEDA compliance.

Personal information. "Personal information" relates to clients, potential clients, or contract staff (e.g., non- employees, volunteers, students) and includes factual or subjective information about an identifiable individual related to their (College of Psychologists of Ontario, 2004):

- personal characteristics (e.g., sex, age, race, ethnicity, country of origin, education or training, family status, religion, occupation, sexual history or sexual orientation);

- health (e.g., health history, conditions, or services received)

Personal Health Information. Personal health information is broadly defined and includes (College of Psychologists of Ontario, 2004; Information and Privacy Commissioner of Ontario, 2015):

- information on an identifiable individual (including data that can be combined with other data to identify the individual);

- oral or recorded information (e.g., asking a question can constitute collecting personal health information even if it is not recorded); and

- information related to the individual's:
    i. personal or family history of their physical or mental condition;
    ii. health care (including maintenance, preventative or palliative measures);
    iii. health care providers;
    iv. payment for the health care or health card number;
    v. substituted decision maker; or
    vi. other information (e.g., phone number) mixed in with other personal health information.

The following are examples of personal health information from the College of Psychologists of Ontario (2004):

- **Personal Characteristics:** Name, home contact information, identification number (e.g., credit card, social insurance, website cookies), insurance benefit coverage, identifying features (e.g., fingerprints, blood type), gender, age, race, language, ethnic or country of origin, education, marital status, sexual history, sexual orientation, income, and social status.
- **Health information:** Health history, measurements, samples, or examination results, conditions, assessment results, diagnoses, services received, information collected during the course of providing services, prognosis or other opinions formed during assessment and treatment, compliance with assessment and treatment, reasons for discharge and discharge condition and recommendations, and bodily donations activities or plans for donations.

Personal Health Information Disclosure, Storage, and Security. Health information custodians (usually, a hospital) have a responsibility to protect personal health information from theft, loss, unauthorized use, disclosure, copying, modification, or disposal. There is a positive obligation for custodians to notify individuals of any breach of their privacy.

PHIPA does provide for disclosure of personal health information for research purposes under certain conditions. Under section 44, subsection 1 in PHIPA, (entitled "Disclosure for Research"), the following guidelines are presented:

> *A health information custodian may disclose personal health information about an individual to a researcher if the researcher,*
>
> *(a) submits to the custodian,*
>
> *(i) an application in writing,*
>
> *(ii) a research plan that meets the requirements of subsection (2), and*
>
> *(iii) a copy of the decision of a research ethics board that approves the research plan; and*
>
> *(b) enters into the agreement required by subsection (5)*

Subsection 2 (entitled "Research Plan") refers to a research plan in writing that must set out

a) *the affiliation of each person involved in the research*

b) *the nature and objectives of the research and the public or scientific benefit of the research that the researcher anticipates and*

c) *all other prescribed matters related to the research.*

Subsection 5 (entitled "Agreement Respecting Disclosure') states that:

> *Before a health information custodian discloses personal health information to a researcher under subsection (1), the researcher shall enter into an agreement with the custodian in which the researcher agrees to comply with the conditions and restrictions, if any, that the custodian imposes relating to the use, security, disclosure, return or disposal of the information.*

If a CRECS researcher is obtaining information from another institution, it is important to note that CRECS is subject to the same privacy and confidentiality requirements as the institution from which the data originate.

Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans.

The Tri-Council Policy Statement was initiated by the presidents of the individual research funding agencies (Canadian Institutes of Health Research [CIHR], Natural Sciences and Engineering Research Council [NSERC], and Social Sciences and Humanities Research Council [SSHRC]). The objective of the policy statement was to ensure that institutions funded by the federal research councils and its researchers conform to ethics policies set by the Tri-Councils for conducting research with humans. The guidelines are occasionally updated, with the most up-to-date version of the Tri-Council Policy Statement published in December of 2014 (TCPS 2, 2014). Chapter 5 speaks specifically to privacy and confidentiality.

According to the TCPS 2, it is important that any exceptions to privacy and confidentiality, whether legal or ethical, be outlined in the process of free and informed consent and approved by the REB prior to the beginning of the data collection, unless the information is publicly available. When conducting program evaluation, approval by an REB may not be necessary, but the same standards concerning privacy and confidentiality should be maintained. Researchers who rely exclusively on secondary use of non-identifiable information are not required to seek participant consent but are required to seek REB review. Identifiability can be context specific (e.g., use of coded information is considered non-identifiable only if researcher does not have access to the key) and consent from participants for secondary use of identifiable information is not always necessary if there is REB approval (for more about secondary use of information see Articles 5.5 and 5.6 of the TCPS 2, 2014).

In addition to the general Tri-Council Policy Statement, one of the three organizations (CIHR) also produced a document on best practices for protecting privacy in health research (Canadian Institutes of Health Research, 2005; http://www.cihr-irsc.gc.ca/e/29072.html). According to the document, data should be assessed using a threat-risk vulnerability assessment, which includes seven steps (i.e., determine what assets need to be protected, determine what to protect against, assess probability of threat occurring, assess magnitude of the impact if threat occurs, assess existing safeguards, recommend appropriate additional safeguards, and update regularly).

Canadian Evaluation Society

The Canadian Evaluation Society provides general ethical guidelines for its members who are conducting evaluation research. Guidelines are divided into three sections including competence, integrity, and accountability. These general guidelines indicate that evaluations should be designed and conducted in a manner that protects the rights (including the right to privacy) and welfare of all participants. In general, evaluators should act with integrity and confer with clients on contractual decisions such as confidentiality, privacy and ownership of findings and reports (Canadian Evaluation Society, 2014).

Canadian Psychological Association

The Canadian Psychological Association has a set of ethical guidelines, which are presented in the document "Canadian Code of Ethics for Psychologists (Fourth Edition)". This ethical code applies to all professional activities for psychologists affiliated with CPA, including clinical work, consultation, and research (Canadian Psychological Association, 2017). According to the CPA, psychologists should seek and collect only information that is germane to the purpose for which consent has been obtained. Personal information is collected and documented only if it is required for provision of future services, required for a research study being conducted, or required or justified by law. Moreover, clients should be informed of the measures taken to ensure the confidentiality of their personal information. Finally, in settings where personal information is shared or collected in group format (e.g., focus group, group psychotherapy), all clients should be informed of their responsibilities for maintaining confidentiality for others in the group (Canadian Psychological Association, 2017).

College of Psychologists of Ontario

The College of Psychologists of Ontario (CPO) has ethical guidelines presented in their document "Standards of Professional Conduct". This ethical code applies to all professional activities for psychologists registered with CPO, including clinical work, consultation and research (College of

Psychologists of Ontario, 2005). According to the CPO, psychologists should make a reasonable effort to ensure records are complete and accessible. For corporate clients (e.g., companies, organizations), a record should be maintained including the name of the corporate client, the name and title of the persons authorized to release confidential information about the corporate client, the date and nature of each service provided to the corporate client, a copy of all agreements and correspondence with the corporate client, and a copy of each report prepared for the client.

Recommendations

CRECS senior researchers are involved in data collection in the field which sometimes entails the collection of confidential information. Therefore, it is important that researchers are mindful of risks that may come up during the collection and handling of their data and that they take appropriate steps to protect the privacy of their participants. A lack of appropriate safeguards and measures for data handling may result in the loss of data, accidental disclosure to an unauthorized party, or theft of data. Appropriate data handling practices are therefore essential for each project and it is the responsibility of all research team members to follow procedures that minimize the risk of a data breach (TCPS 2, 2014).

Understanding Identifying and Anonymous Data

Researchers may seek to collect, use, share, and access different types of information about participants. Such information may include personal characteristics or other information about which an individual has a reasonable expectation of privacy. Before acquiring their data, researchers should determine what type of data they are working with when considering steps to safeguard that data (see Table 1 for more detail). The TCPS 2 (2014) has provided the following five categories to assess the extent to which information can be used to identify a participant.

- **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; a master list of codes/identities is retained. Participants can be re-identified by linking codes to identifying data.
- **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.

The TCPS 2 explains how ethical concerns regarding privacy decrease as it becomes more difficult (or impossible) to associate information with a particular individual. These concerns also vary with the sensitivity of the information and the extent to which access, use or disclosure may harm an individual or group. The easiest way to protect participants is through the collection and use of anonymous or anonymized data, although this is not always possible or desirable. If identifying information will be anonymized, it should be done as soon as possible after it is acquired and often to the greatest extent possible (note that data may not always need to be automatically de-identified; the decision to de-identify or anonymize data will depend on factors such what participants may have consented to, degree of sensitivity of the data, etc.).

Data can also be anonymized by re-grouping (and recoding) raw data into groups within a certain range. For example, the indirectly identifying information of age could include an outlier of 89 years of age, where the overall average age of the participants is 34. Grouping this data into a “65 and over” category can mask the identity of the outlier. Such anonymization is common practice by

agencies such as Statistics Canada, in producing public-use microdata files. The University of Ottawa Library outlines some recommendations for anonymizing data and even links to a data anonymization helper tool (an add-on to MS Word) for qualitative data researchers may find useful.

Where it is not feasible to use anonymous or anonymized data for research (and there are many reasons why data may need to be gathered and retained in an identifiable form), the ethical duty of confidentiality and the use of appropriate measures to safeguard information become paramount. Researchers are expected to consult their REB if they are uncertain about whether information proposed for use in research is identifiable.

In the United States, the *Health Insurance Portability and Accountability Act* (HIPAA) Privacy Rule considers that data are de-identified if **all** of the following is removed from the data:

1. Names.
2. All geographic subdivisions smaller than a state, including street address, city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code if, according to the current publicly available data from the Bureau of the Census:
    a. The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and
    b. The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000.
3. All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older.
4. Telephone numbers.
5. Facsimile numbers.
6. Electronic mail addresses.
7. Social security numbers.
8. Medical record numbers.
9. Health plan beneficiary numbers.
10. Account numbers.
11. Certificate/license numbers.
12. Vehicle identifiers and serial numbers, including license plate numbers.
13. Device identifiers and serial numbers.
14. Web universal resource locators (URLs).
15. Internet protocol (IP) address numbers.
16. Biometric identifiers, including fingerprints and voiceprints.
17. Full-face photographic images and any comparable images.
18. Any other unique identifying number, characteristic, or code, unless otherwise permitted by the Privacy Rule for re-identification.

According to HIPAA, data is not considered fully anonymized unless all the above information is removed from a document. However, broadly speaking in the Canadian context, if, for example, there is a “large” sample at both the province and city level, geographic information smaller than that of the province, (e.g., the city level) may be permissible.

Assessing the Risk of Data to Be Acquired, Handled, or Stored

A first step in planning for the safe collection, handling and storage of data is to determine what kind of data are being collected, and what are the risks associated with those data. Based on the Harvard Information Security Policy’s five data security levels (Harvard University, n.d.), we have developed four levels of risk associated with research data. When collecting data, researchers should limit the data they collect to only gather the minimum amount of personal and identifiable information necessary to achieve research objectives. Collection of personal information not relevant to research objectives may unnecessarily lead to a higher level of risk associated with possible breach of confidentiality, which may make participants vulnerable to a greater risk of harm. The following applies to circumstances where conditions of participation in research determined that participants’ information would not be made public and that confidentiality would be kept.

Low-risk confidential research information

Confidential research information is information that in its present form, would not cause harm to an individual or group if disclosed but which researchers have nevertheless decided to keep confidential. Information at this level may include anonymous information, anonymized information, or coded information that if de-coded and disclosed would not cause serious harm to participant.

Examples include:

- Anonymous survey data
- Completely anonymized data (see anonymization guidelines below)
- Unpublished intellectual property (e.g., manuscript drafts)

Sensitive confidential information

Sensitive confidential information is information that if disclosed in its present form can reasonably be expected to cause some damage to an individual’s reputation, or cause embarrassment. Information at this level may include low risk data that has not-yet been anonymized and also coded information that if de-coded and disclosed could cause serious harm to participant.

Examples include:

- Bank account numbers
- Student records (e.g., transcript)
- Information supplied in confidence (information understood by the respondent and so noted to be suppled in confidence)

Information that would likely cause harm if disclosed.

Information at this level includes information that if disclosed could create risk of social, psychological, reputational, financial, legal, or other harm to an individual or group. This level may include data that has not-yet been anonymized or cannot be anonymized because confidential information is needed for analysis but if disclosed, may cause harm to the participant.

Examples include:

- Health card information
- Diagnosis of mental or physical illness
- Credit card numbers

Information that would cause severe harm if disclosed.

Information at this this level includes information that if disclosed could create risk of criminal liability, loss of employment, or severe harm to an individual or group. This level is reserved for data of the most sensitive nature and should be anonymized as soon as possible. Highly confidential information that cannot be fully anonymized because it is needed for analysis must be handled with the extreme care.

Examples include:

- Social insurance number
- Information about illegal activity

How to Acquire, Handle, and Store Data Associated with Different Levels of Risk

Table 1 provides an overview of recommended steps to acquire, handle, and store data appropriate for the different levels of risk described above. These guidelines are based on recommendations outlined in the Harvard Information Security Policy and consultations with experts in data security at the University of Ottawa. This table was designed as a tool to aid researchers in deciding on what steps to take to secure their data. Ultimately, it is up to the researcher to judge the level of risk associated with her data. It is important to note that data can move from one risk category to another throughout the research project, most often from higher levels down (e.g., a database containing participants' names and health card numbers may at first be a level 3, but after the data is fully anonymized, it may be moved down to a level 2, or even level 1.)

In subsequent sections we describe specific steps that can be taken that are appropriate for various stages of the research process, from data acquisition to storage.

Table 1: Levels of Risk and Corresponding Steps to Safely Handling Data

Level of Risk	Steps to Securing Data
4 Information that would cause severe harm if disclosed If disclosed, could create risk of criminal liability, loss of employment, or severe harm to an individual or group. This level is reserved for data of the most sensitive nature and should be anonymized as soon as possible. Highly confidential information that cannot be fully anonymized because it is needed for analysis must be handled with the extreme care.	Field collection: Data should be collected on an encrypted and password-protected device. Use of paper material is discouraged, but if used should be handled with extreme care and not left unattended unless in a locked and secure environment. Storage: Data must be stored in a physically locked room (preferably secured by an alarm) on a password-protected and encrypted hard drive or password-protected computer not connected to a data network. Sharing: Sharing at this level should be limited, data should only be accessed in a secure location. Access: Controlled by PI, keeping a list of individuals who have been granted access to data.
3 Information that would likely cause harm if disclosed If disclosed, could create risk of social, psychological, reputational, financial, legal, or other harm to an individual or group. This level may include data that has not-yet been anonymized or cannot be fully anonymized because confidential information is needed for analysis and if disclosed to an unauthorized party, may cause harm to the participant.	Field collection: Data should be collected on an encrypted and password-protected device. Use of paper material is discouraged, but if used should be handled with extreme care and not left unattended unless in a locked and secure environment. Storage: Data must be encrypted and password-protected. Sharing: Data should not be shared by email. Files must be encrypted when using Docushare. Access: Should be controlled by PI, keeping a list of individuals who have been granted access to data.
2 Sensitive confidential information If disclosed in its present form, can reasonably be expected to cause some damage to an individual's reputation, or cause embarrassment. Information at this level may include low risk data that has not-yet been anonymized and also coded information that if de-coded and disclosed could cause serious harm to participant.	Field collection: Data should be collected on a password-protected device. Storage: Data must be password-protected, encryption is recommended. Sharing: Files sent via e-mail should be password-protected and encrypted. Password must be sent through a different medium. Use of Docushare, however, is preferred to e-mail.
1 Low-risk confidential research information Information that in its present form would not cause harm to an individual or group if disclosed but researchers have nevertheless decided to keep it restricted. Information at this level may include anonymous information, fully anonymized information, or coded information that if de-coded and disclosed would not cause serious harm to participant.	Storage: Data must be stored on a password-protected computer or drive. Sharing: It is recommended that files sent via e-mail be password-protected. Password must be sent through a different medium.

Note. Levels of risk adapted from the Harvard Research Data Security Policy

Data Collection

The collection of data from human participants is an important aspect of research and evaluation. CRECS senior researchers may work with two kinds of data.

Primary data include data that are directly collected from participants in research or evaluation projects. Because data collection often entails the collection of confidential information, researchers should endeavor to have the most secure means of collecting such data. As paper surveys can be lost, data collection using more secure devices that can be password-protected and encrypted is recommended (e.g., laptops or tablets; note that Apple products are generally more secure, iPads have strong default encryption, other tablets may be less secure). Moreover, a tablet or mobile device is only considered secure if it is password-protected with a strong password and if only properly validated apps are installed (i.e., only certified apps from an official app store).

Secondary data are data that were originally collected by a community organization, hospital, or other institution, which is then transferred to CRECS researchers for analysis.

The use of both types of data must be undertaken respecting the principles of privacy, confidentiality, and data security. As such, the following recommendations are made:

Recommendations for Collecting Primary Data

- Always ensure that all research projects involving human participants or the use of their data have received all required approvals from the uOttawa REB.
- If necessary, ensure you have obtained consent from each research participant as to the collection of personal information, together with the use and purpose of such collected information. Each participant has a right to know what they are agreeing to and to choose whether they want to participate.
- When collecting data, always make sure you only collect data that you need, and after that data is collected and safely stored, do not remove data from your secure workplace unless absolutely necessary. If working under supervision, removal of data should be done with the knowledge of your supervisor.
- The University of Ottawa provides some helpful guidelines for safeguarding records when they must be taken off campus. Some include:
    - Records on a portable electronic device (e.g., laptop) should never be left unattended. When not used, they should be stored in a secure location
    - A portable electronic device containing data should be password-protected and encrypted
    - Passwords should not be easy to guess, should be kept confidential.

Recommendations for Acquiring Secondary Data

- Always ensure that all research projects involving human participants or the use of their data have received all required approvals from the uOttawa REB.

- Always ensure that all projects involving the use of aggregate data obtained from research partners (such as hospitals, government agencies, and others) have been approved by the appropriate authorities of the research partner and that required consents have been obtained from the research partner and, if needed, from individual participants, for release of the data to CRECS researchers.

- Never accept data with identifying information unless such information is absolutely required for your project and appropriate measures are implemented to safeguard the confidentiality of the data, including password protection and encryption of files. Wherever identifying information is not needed, you will require research partners to remove all identifying information before providing it to CRECS researchers.

- The uOttawa REB has in some cases, waived ethics approval for access to secondary data at the record-level, where the data protection measures in place are deemed to be in full compliance to the governing data privacy and protection laws. Public-use microdata files provided by Statistics Canada and available on odesi.ca or archived in the Inter-University Consortium of Political and Social Research (ICPSR) require no prior uOttawa REB approval or waiver process.

Data Use, Handling and Transportation

CRECS has researchers and staff members all working in various capacities on a variety of research projects. Data from these numerous projects must be collected offsite at times and may need to be returned to CRECS and secured after normal business hours. Additionally, because of limited resources and space combined with a large number of researchers, the computers and offices are often shared among numerous CRECS researchers. As with data acquisition above, these issues present challenges that must be met with solutions that ensure the principles of privacy, confidentiality, and data security are upheld. As such, the following recommendations are made:

<u>Recommendations for Data Use and Handling</u>

- Minimize the use of data with identifying information. To minimize the risk of accidental confidentiality breaches:
    - Use an anonymized or anonymous dataset wherever possible;
    - Where identifying data are needed, always keep identifying information in a separate ID list (e.g., a spreadsheet or database with unique identification codes);
    - Ensure computers and files are password-protected and encrypted;
- Only use data for the approved purposes.
- Never enter data outside the prescribed spaces for data entry. Moving hard copies of data

around CRECS increases the risk of accidental confidentiality breaches;

- Safely store paper material in locked cabinets in locked offices whose access is limited to staff members;
- Handling of data (e.g., survey data) in hard copy format is discouraged. Secured (encrypted, password-protected) electronic data collection tools such as laptops or tablets are preferred;
- Never share personal authentication credentials (user-IDs, passwords, etc.) or use your credentials to give another person access to information systems or computers containing confidential information.

Data sharing

- Researchers should only share data with approved external groups and researchers and should ensure they use a means to share data that is sufficiently secure.
    - Email is inherently unsecure and encrypting emails is very difficult and often impractical. It is recommended that researchers use DocuShare, which offers a more secure means of sharing files, although confidential information put up on Docushare must be encrypted. Cloud-based storage (e.g., DropBox, Google Drive) should not be used for storage or sharing of confidential data. Although email should not be used for confidential data sharing, email can be used when communicating with researchers or participants, or sharing very low risk anonymized data.
    - USB sticks should not be used unless data is deemed to be very low risk because they can easily be lost or misplaced, and even after data is deleted there are ways that it could be retrieved, making it very difficult to permanently delete data from USB. When a USB is used, data files must be encrypted and password-protected;
    - Only share confidential data with approved external groups once proper permission has been obtained from the original source, and if relevant, approved by the REB;
    - Only share data with agencies after their privacy, confidentiality and data security policies have been determined to be adequate; and
    - Ensure agencies acknowledge in writing their and CRECS' responsibility to preserve the confidentiality of all information shared.

Data transportation

- Avoid, where possible, taking data offsite, whether in hard copy or in electronic format.
- When transporting data, always maintain the security of data collected offsite:
    - Never leave data unattended or in non-secure locations (e.g., a locked car) as this increases the risk of accidental confidentiality breaches;

    - Ensure computers and files are password-protected; and
    - Ensure electronic data that include any identifying information are encrypted.
- Never remove an ID list or other identifying information from CRECS.
- Never transport ID lists, or any other data or information that affords the possibility of identifying individuals from the anonymized dataset, together with the anonymized data.
- Always bring data collected offsite to CRECS as soon as possible. Always ensure that such data is de-identified as soon as possible.
- If data collected offsite cannot be brought directly to CRECS, maintain security practices as listed above, and if working under supervision, always get prior written approval from your supervisor for the data handling and security procedures.

Data Storage

The following steps ought to be taken when very confidential data is being stored. These steps will likely be taken in the early stages of the research process before the data can be anonymized.

Step 1: Find a clean computer / create a safe environment

First and foremost, it is essential that computers are up-to-date. System providers release operating system updates on an ongoing basis to protect users against vulnerabilities hackers have identified and exploited.

Next, it is important to scan computers for malware; computers are easily infected by malware. Even reputable websites such as that of a reputable news organization may unknowingly host advertisements that are infected with malware. Therefore, it is preferred that a computer that handles confidential data (i.e., pre-anonymized data) be only used for data handling. To be considered safe, a computer should first be scanned for malware and if any threats are detected, they should promptly be cleaned or removed. Sophos Home has been recommended by the experts in data security at the University of Ottawa.

Step 2: Protect and encrypt

Once a computer has been scanned and any threats removed, a researcher may load the confidential data onto the computer (i.e., interview recordings, non-anonymized data, etc.). Once this is done it is recommended that the computer's entire drive is encrypted and strongly recommended that the individual files be encrypted (see the Resources section for encryption software).

Step 3: Permanent file removal

Files with confidential information should never be deleted using standard "trash bins" installed on computers to "remove" files. This only removes the file directory, yet the file and its information is

still stored physically on the computer's hard drive, and remain accessible. "File shredding" applications can be used to permanently delete files. These programs overwrite the file with random letters and/or numbers, often multiple times, before removing them. It is important to note that this applies to files that contain confidential information and are no longer needed, consistent with REB approval (e.g., an audio recording that has been transcribed.) A file that has been anonymized and will later be used for research does not need to be destroyed, it can simply be password-protected, and encrypted.

Creating a Strong Password to Protect Accounts and Documents

Not all passwords are equally secure. If a password is not strong enough, encryption will not matter. An individual can bypass a weak password with ease using a "brute-force" software using one of many brute-force applications easily accessible online. These applications can guess a weak password in minutes, if not seconds. It is also important to never store a password on the same computer confidential data is stored on, unless it is stored in a secure password manager (examples below).

Software can be used to quickly run through countless possible passwords, and the software is constantly being improved to check passwords for:

- Dictionary words
- Words spelled backwards
- Personal information such as names, birth dates, local street names

The University of Ottawa provides guidelines for stronger passwords; the following is a summary of some of those guidelines.

- Length: at least 10 characters long (this will take longer for the hacker to decipher)
- Complexity: at least three uppercase/lower case letters, numbers, and special characters.
- Variation: change a password at least every three months
- Variety: use a different password for every site you use (e.g., different sites = different passwords).

If you are not sure if your password is strong, there are various resources online (i.e. Password Meter and Password Checker) to check password strength. There are also random password generators (i.e. Password Generator and Password Generator Online) that some researchers may find useful.

Keeping track of strong passwords can be difficult. In fact, if passwords are sufficiently strong, it should be nearly impossible to remember them all. A password manager can be a useful tool in helping one keep track of all their complicated passwords. The following is a list of password

managers (all are either free of charge or include a free version) recommended by the University of Ottawa's security architect:

- Dashlane
- KeePass (open source)
- Sticky Password
- LastPass
- Password Safe (open source)
- RoboForm Password Manager
- SplashData SplashID

Of the above options, we recommend LastPass, which offers a complex password generation tool, automatic form and login/password completion, is available for both Windows and Mac operating systems, and also can be downloaded on smartphones (including Apple's iOS and Android). LastPass is free to use, but the premium version is very affordable, at $2/month. The premium version includes the ability to share passwords with others with their "one to many sharing" function and 1GB of free encrypted cloud-based storage. LastPass is easy to use and automatically registers and remembers passwords as one logs in to their accounts on their computer or smartphone, and even includes an automatic password changing tool. Many tutorials can be found online.

Additional Recommendations for Data Storage

- Always store data and ID lists on "safe" computers. Password protection, data encryption, firewalls, updated virus protection, and a restricted number of users will minimize the risk of accidental confidentiality breaches.
- If possible, do not use hard copy media to store confidential data. If this is not possible, ensure that such media are stored in a secure and locked environment.
- Never store anonymized data in the same place as ID lists or other identifying information.
- Remove any confidential information when it is no longer needed and ensure that it is properly cleansed from any electronic storage media (including portable media) or, if in hard copy, securely destroyed.
- Always archive data as required. Keep old data in a safe place for the proper length of time. Needs for preserving data beyond the life of the project and research publications should be considered, will the data be needed for replication, further documentation, reference purposes? These and other considerations raised in the Tri-Agency Statement of Principles on Digital Data Management, may necessitate consideration beyond the five years after the completion of the research project.
- When possible, store files in more than one format. QDA Miner and SPSS Files can also be saved in Excel format. This is also useful if data is being shared with another researcher who

may not have QDA Miner. While data from software packages such as SPSS and QDA Miner can both be exported to and saved in Excel format, some such as NVivo cannot.

Some best practices outlined by the University of Ottawa include:

- When naming files, use of spaces and special characters (?, &, !) is discouraged. Hyphens, underscores, and capitalization are best when separating elements in a file name. A file name should include key information about the project such as the phase the project is in, research team working on the project, language used in file, etc. (More on file and folder structure here)
    - E.g., SecondCodeList-HousingProject-Dec2016 or Research_Proposal_CRECS2_Eng
- File formats should be carefully chosen with share-ability and long-term access in mind (list of preferred file formats here)
- To minimize the chance of accidentally working on an outdated version of a file, researchers are encouraged to practice "file versioning". Some suggested methods:
    - Include information about version in file name as well as in the document
    - Use of sequential numbering is encouraged (e.g. 0.1, 0.2, 0.3 for drafts, 1.0 for a final version, 1.1, 1.2 may signify revisions to a final version until version 2.0 is complete)
- Data should be regularly backed up, a regular back up schedule should be maintained. It is recommended at least three geographically distributed copies are regularly backed up: original, external/local (e.g., external hard drive in locked room), external/remote (e.g., remote back up of data on secure and approved platform such as DocuShare)

Training and Supervision

CRECS is a multidisciplinary centre that employs a large number of researchers, including employees, post-doctoral students, graduate students, and undergraduate or honours students. As such, researchers beginning work at CRECS can have varying backgrounds in ethical and research training. This variety of foundational knowledge in ethics and research indicates that training plays an important role in ensuring that the principles of privacy, confidentiality, and data security are upheld at CRECS. Additionally, CRECS functions both as a research centre as well as a training centre for numerous students at the University of Ottawa. As such, supervision of researchers is important to ensuring both the quality of work as well as adherence to the ethical principles of privacy, confidentiality, and data security. To ensure that proper training and supervision are conducted at CRECS, the following recommendations are made.

Employees and students who will have access to and use data:

- Must read this manual on privacy, confidentiality, and data storage at CRECS.
- Must complete the Introductory Tutorial for the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans 2 (TCPS 2) together with any additional training program the University of Ottawa or CRECS may require from time to time to ensure a basic foundational knowledge on research ethics.
- Must complete the confidentiality and non-disclosure agreement provided in Appendix 1: The agreement outlines the ethical guidelines for privacy, confidentiality, and data storage at CRECS.

Additionally,

- Supervision of every project must include oversight of confidentiality, privacy and data security procedures.
- All ethical issues must be raised with a supervisor as soon as possible. All ethical decisions must be made in consultation with a CRECS supervisor and be documented in the project file altogether with approvals from the CRECS supervisor in writing.
- Any unauthorized use, access or loss of confidential information must be reported as soon as possible to one's immediate supervisor (if working under supervision) and the Office of Research Ethics and Integrity; the Office should also be advised if the information in question is directly identifiable.

Documentation of Services

With the numerous researchers at the CRECS working in various capacities on several research projects, documentation of services is vital to operations at CRECS. Ensuring duties are being performed properly, that ethical principles (including privacy, confidentiality, and data security) are being adhered to and that work is being completed are all important aspects of CRECS work that must be appropriately documented. To ensure that services are being appropriately documented at CRECS, the following recommendations are made:

1. All services for corporate clients will be documented in line with CPO standards summarized earlier in this manual (College of Psychologists of Ontario, 2005).
2. Each researcher and supervisor must keep a record confirming that issues related to privacy, confidentiality, and data security were discussed in relation to the project they are working on and what decisions were made.
3. If working under supervision, all deviations from the guidelines relating to privacy, confidentiality and data security must be approved in writing by the CRECS supervisor (and possible the REB).

Developing Data Management Plans

A data management plan (DMP) refers to an outline, often in written form, that summarizes how data is handled before, during, and after a research project. The usefulness of a data management plan (DMP) is recognized by the Tri-Agency Statement of Principles on Digital Data Management which envisions making data management plans mandatory for research grant applications in the future. Developing a DMP can help address many of the concerns raised in this manual and implement many of the recommendations outlined above. It is strongly recommended that CRECS researchers develop a DMP prior to beginning a new project. The University of Ottawa has dedicated a section of their website to helping researchers develop and adjust DMPs appropriate for their research project, including helpful tips, examples, links to helpful articles, and even an online DMP tool. As described on the University of Ottawa website, a typical DMP includes information about:

- How the data are to be collected or acquired
- Conventions and procedures to be used to structure, store, backup, and preserve the data
- The assignment of responsibility for the production of documentation or metadata
- Security measures, if any, to protect confidential data
- Legal, ethical or intellectual property issues, if any, and how they will be addressed
- Resources needed to implement the DMP

The University's DMP website also contains a section on long-term storage of data outlining various data repositories used to provide infrastructure for such storage and archiving of data.

Assistance in the preparation of DMP's is available through the uOttawa library rdm@uOttawa.ca

Resources

Below we list some resources that may be helpful in ensuring the security of research data.

> *Note.* Links to resources may eventually not be supported or may evolve into different software (or versions). Most of the software we recommend serve basic functions and there exist countless acceptable replacement options and a multitude of tutorial options. Please consult the uOttawa IT department or a librarian specializing in data management for more information about specific resources.

Malware Scan + Removal Application

Sophos Home is the application we recommend. Note that the "home" version is free of charge. The University of Ottawa can install Sophos Business on machines connected to the school's domain free of charge. For more information, contact the Information Technology Service Desk.

Encryption

Full drive encryption is very simple. Mac users can use FileVault while Windows users can use BitLocker. Most computers will have these installed by default and can be turned on with a click of a

button. We caution however, that drive encryption can give one a false sense of security. When logged in, the drive is decrypted and can be accessed by a hacker. This is why file-level encryption is essential for very sensitive data.

MS Word allows users to encrypt a file with a password. This encryption is strong, but it is important to remember to use a strong password.

VeraCrypt (formerly known as TrueCrypt) is recommended for a higher level of file-level encryption. This is an open-source encryption software known to many as the "gold standard" for open-source encryption, and is very easy to use. VeraCrypt will create a strongly encrypted and secure "container" – this is a file where other files can be safely stored.

It is important to review the beginner's tutorial before using the software. After some practice researchers will see this software is very easy to use. There are many other step-by-step guides that can be found by searching Google.

Latest versions of other software (e.g., SPSS) also have encryption capabilities and there exist many tutorials and how-to guides easily found on Google.

Permanent File Removal

File Shredder is a very simple to use and effective application for Windows users.

Mac users can use apps such as Incinerator or Permanent Eraser. While Permanent Eraser is free, Incinerator costs only $1. Most may prefer Incinerator since it is easier to use (just drag and drop file you want deleted onto the icon in your dock) while Permanent Eraser will erase everything in one's Trash. However, more powerful apps such as AweEraser may be necessary for deletion of larger files.

Glossary

Anonymization – The process of permanently removing identifying information from data with the intent of privacy protection. This includes methods, in the context of the data sample, to suppress outliers, aggregating and recoding identifying information, such as participants' exact age, income, occupation groups, and addresses.

Confidential Information (or data) - protected due to proprietary, ethical, or privacy considerations. (uOttawa Information Classification and Handling Policy).

Confidentiality – An ethical and/or legal responsibility of individuals or organizations to safeguard information entrusted to them, from unauthorized access, use, disclosure, modification, loss or theft (TCPS 2).

De-identification – The process of removing identifying information from data with the intent of privacy protection. The data can still be re-identified; direct identifiers are removed from the information and replaced with a code; a master list of codes/identities is retained.

Encryption – The process of encoding information in a way that can only be accessed by authorized individuals. To access an encrypted file, an individual will need to access a "key" generated by a complicated algorithm – this key is accessed using a password.

Malware – Short for *malicious software*, which is any software designed to disrupt a computer's normal operations, gather sensitive information, or grant an unauthorized third party access to an infected computer.

Personal Information – Information that may be reasonably expected to identify an individual, alone or in combination with other available information. Also referred to as Identifiable information (TCPS 2).

Privacy - An individual's right to be free from intrusion or interference by others (TCPS 2).

Secondary Use (Secondary Data) – The use in research of information or human biological materials originally collected for a purpose other than the current research purpose (TCPS 2).

Many of the above definitions are taken directly from the TCPS 2 Glossary.

References

Canadian Evaluation Society (2014). *CES guidelines for ethical conduct.* Retrieved from https://evaluationcanada.ca/ethics

Canadian Institutes of Health Research (2005). *CIHR best practices for protecting privacy in health research*. Retrieved from http://www.cihr- irsc.gc.ca/e/29072.html

Canadian Psychological Association (2017). *Canadian code of ethics for psychologists (fourth edition)*. Retrieved from http://www.cpa.ca/docs/File/Ethics/CPA_Code_2017_4thEd.pdf

College of Psychologists of Ontario (2016). *The Personal Health Information Protection Act, 2004: A guide for regulated health professionals.* Retrieved from http://www.cpo.on.ca/WorkArea/DownloadAsset.aspx?id=1591

College of Psychologists of Ontario (2005). *Standards of professional conduct*. (Revised March 2009) Retrieved from http://www.cpo.on.ca/WorkArea/DownloadAsset.aspx?id=335

College of Psychologists of Ontario (2013). *Privacy code of the college of psychologists of Ontario.* Retrieved from http://www.cpo.on.ca/WorkArea/DownloadAsset.aspx?id=653

Harvard Information Security (n.d.). *Data Classification Table.* Retrieved from https://security.harvard.edu/dct

Information and Privacy Commissioner of Ontario (2015). *Frequently asked questions: Personal Health Information Protection Act*. Retrieved from https://www.ipc.on.ca/wp-content/uploads/Resources/phipa-faq.pdf

Office of the Privacy Commissioner of Canada (2013). *Provincial legislation deemed substantially similar to PIPEDA*. Retrieved from https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/legislation-related-to-pipeda/provincial-legislation-deemed-substantially-similar-to-pipeda/

Office of the Privacy Commissioner of Canada (2015). *A guide for individuals - Protecting your privacy*. Retrieved from https://www.priv.gc.ca/en/about-the-opc/publications/guide_ind/

Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, December 2014. Retrieved from http://www.pre.ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2018.html

Appendix

Centre for Research on Educational and Community Services

Faculty of Social Sciences and Faculty of Education

University of Ottawa

Agreement of Confidentiality and Non-Disclosure

The purpose of this document is to clarify and acknowledge understandings about the nature and handling of confidential information and material that an individual (academic, staff person, student, volunteer or other person) might encounter at the Centre for Research on Educational and Community Services (CRECS) at the University of Ottawa.

Where specific instances are not stated in this agreement, the general intent of the document is to safeguard the confidentiality of information received by CRECS researchers and to ensure ethical behaviour in research. More specific conditions, for example, as may be required by CRECS procedures or the requirements of Research Ethics Boards may also apply.

I understand and agree that compliance with this agreement is a prerequisite for my participation in CRECS research activities.

I am undertaking a task and/or participating in research activities that may include interviewing human participants, transcribing interviews, entering data or other research-related activities at CRECS at the University of Ottawa. I understand that my activities will support research study(ies) that involve(s) human participants.

In conjunction with the understandings that I have gained from an orientation and information session on the research activities of CRECS, I agree to hold in confidence all matters that come to my attention in the line of duty at CRECS including material from and about research participants, community agencies and clients.

I agree to abide by all CRECS procedures on collecting, safeguarding and working with confidential data.

I understand that any verbal discussion of human participant research is confined to the research team and as designated by the Principal Investigator. The “Principal Investigator” is the individual directing a research project or program and who is responsible and accountable for the proper conduct of the research project or program.

I shall hold all information derived from research activities in trust and confidence and agree that any of the information will be used only for the purposes of the research and/or theses, and shall neither be used for any other purpose nor disclosed to either the sponsoring (funding) agency or any other third party unless approved by the Principal Investigator.

No copies shall be made or retained by me of any original information in either soft or hard copy form for my personal use. Any information that is removed from the premises must always be anonymous with all potential identifying characteristics removed. Any information that is transferred in this manner must be cleared with the Principal Investigator in writing and

permission received in writing. This includes all forms of data including paper, electronic and audio or visually recorded data.

At the conclusion of my activities with CRECS, any information, data, computer passwords and any other related document shall be returned to the Research Centre. I will also erase any electronic data bases that are in my possession or under my control.

I agree to strictly observe and implement the requirements of the Research Ethics Board and or other bodies in all of the research activities in which I participate.

Dated at Ottawa, Ontario, this _____ day of _____ 20__. Agreed to and accepted by

_____

Signature

_____

Name

_____

Supervisor/Employer

_____

Signature

_____ _____ _____

Name Address Occupation