RHODES UNIVERSITY
Where leaders learn

RHODES UNIVERSITY HUMAN RESEARCH ETHICS COMMITTEE

SOP 4.7 COLLECTION AND STORAGE OF RESEARCH DATA

Approved by:	Name	Signature	Date
Human Research Ethics Committee	Dr Janet Hayward (Chair)		26 / 07 / 2024
Endorsed by:			
VC Legal Unit	Ismail Mnjoli		02/12/2024
DVC: RISP	N. Mzilikazi		09/12/2024

Compiled by Dr Janet Hayward, Chair, Rhodes University Human Research Ethics Committee and Samuel Simango, Senior Manager of Academic Support Services, Rhodes University Library.

DOCUMENT HISTORY
Version 1.0 (July 2024)

COLLECTION AND STORAGE OF RESEARCH DATA

1. Purpose

The purpose of these guidelines is:

1.) To formalise a procedure for the collection and storage of research data (in written/video and/or audio form) that is collected from or relates to human participants.

2.) To ensure compliance with legislative frameworks, specifically the Protection of Personal Information Act (POPIA).

For guidelines regarding the collection and storage of biological material, refer to SOP 4.8 COLLECTION AND STORAGE OF HUMAN BIOLOGICAL SAMPLES.

2. Introduction

2.1. Research data are valuable assets that contribute to the knowledge economy.
2.2. They need appropriate management, protection, and curation.
2.3. Correct and appropriate management of research data must cover the entire lifecycle of the data to ensure its integrity, confidentiality, and availability. [1]

3. Definitions

3.1. Personal information: includes (but is not limited to) information about a person's demographics, background, biometrics, contact details, identifiers (e.g., ID number, photos), opinions and preferences, financial information, correspondence, and criminal record. [2]
3.2. Special personal information: includes information about children, and information about a person's race or ethnic origin, health, DNA, religious or philosophical beliefs, political opinions, sex life, and criminal behaviour. [2]
3.3. Data subject: the research participant whose information is being processed during the research project. [2]
3.4. Responsible party: the responsibilities of the responsible party outlined in POPIA will fall on the researchers designing and leading the research study. Legally, ultimate responsibility lies with the research institution by which the Principal Investigator is employed or with which they are affiliated. [2]
3.5. Data processing: the carrying out of operations on data, especially by a computer, to retrieve, transform, or classify information.
3.6. Data storage: the retention of information using technology (computers and other devices) specifically developed to keep that data and have it as accessible, as necessary.
3.7. Data-sharing: the process of making data resources available to multiple users, applications, or organizations. It involves various technologies, legal frameworks, cultural elements, and practices to facilitate secure data access. Data sharing aims to enhance collaboration, decision making, and analysis, as well as promote transparency, research reproducibility, and increase the impact of research data. Factors including due diligence, data processing principles, and documentation of actions and decisions must be considered before data sharing takes place.

---

[1] Stellenbosch University Research Data Management Regulations (2021:2). Available online at chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.sun.ac.za/english/Documents/Terms_and_conditions/Current/RDM_Regulations_English.pdf

[2] Protection of Personal Information Act (POPIA), 2013:14-16, 38 (South African Government Gazette Vol 581 No 37067).

3.8. Open data: data that is available without any restrictions.

4. Collection of research data from human participants

4.1. This must only go ahead where:

1.) Ethical clearance has been obtained from a Research Ethics Committee (Refer to SOP 2.3 HUMAN RESEARCH ETHICS COMMITTEE REVIEW PROCESSES).
2.) Any conflicts of interest have been disclosed.
3.) Informed consent has been obtained from research participants (where applicable)
4.) Personal identifiers are anonymised or deidentified where anonymisation was not implemented.
5.) Gatekeeper permission has been obtained (where applicable).

5. Conditions for the lawful processing of data

5.1. The Protection of Personal Information Act (POPIA) defines eight conditions for the lawful processing of data:

1) Accountability: The Responsible party must ensure compliance with POPIA. They must be identified in the research protocol and ensure that all the conditions for lawful processing are complied with through all phases of the research project. [3]
2) Processing limitation: Personal information must be processed lawfully and in a manner that does not infringe on the privacy of the data subject. It may only be processed if the purpose for which it is processed is adequate, relevant, and not excessive. Personal information must be collected directly from the data subject who must be a competent person who consents to the processing and may withdraw consent at any time. [4] (Refer to SOP 4.3 RESEARCH INVOLVING VULNERABLE PERSONS for further details regarding the 'competency' of research participants).
3) Purpose specification: Personal information must be collected for a specific and lawful purpose which must be outlined in the research protocol. The data subject must be made aware, in the consent form, of the explicit research-related purpose for which their information is being collected. Records of personal information must not be retained longer than necessary and should be destroyed or deleted after authorisation to retain such records has expired. Destruction or deletion must be done in a manner that prevents reconstruction. [5]
4) Further processing limitation: Further processing of personal information must be in accordance with the purpose for which the data was originally collected. [6]
5) Information quality: The responsible party must have regard to the purpose for which the personal information is collected and must take steps to ensure that the personal information collected from data subjects is complete, accurate, not misleading and updated where necessary. [7]
6) Openness: The responsible party must maintain documentation of all processing operations. They must take steps to ensure that the data subject is aware that the information is being collected, and the purpose for which it is being collected. The

[3] POPIA, 2013:22
[4] POPIA, 2013:24-26
[5] POPIA, 2013:26-28
[6] POPIA, 2013:28-30
[7] POPIA, 2013:30

information must be used for historical, statistical or research purposes and may not be used in a form that will allow for the identification of the data subject. [8]

7) Security Safeguards: The responsible party must secure the integrity and confidentiality of personal information by taking appropriate, reasonable, and technical measures to prevent the loss, damage, or unauthorised destruction of personal information. They must also ensure that unlawful access to or processing of personal information is prevented.

8) Data subject participation: A data subject has the right to request the responsible party to confirm whether personal information about the data subject is held and to request a record of such information. A data subject may request the responsible party to correct or delete personal information if this is inaccurate, irrelevant, excessive, out of date, incomplete, misleading, or obtained unlawfully. On receipt of such a request the responsible party must correct, destroy, or delete the information as requested and notify the data subject of the action taken in response to the request. [9]

6. Data storage

6.1. Storage period.

6.1.1. Research data must be stored securely for a minimum of five years after completion of the research project to ensure that reassessment/analysis of the data can take place should this be deemed necessary.

6.1.2. Extended periods of data storage are permissible where disciplinary practice or compliance with terms imposed by research funders requires this.

6.1.3. Data can be retained for follow-up or longitudinal studies, provided subsequent studies are in accordance with the purpose for which the data was originally collected (see 5.1. 4) above).

6.1.4. In the case of 6.1.2 and 6.1.3, the extended storage / retention of data for future use must be explained and justified in the research protocol and made clear in the informed consent process and documentation.

6.1.5. Upon expiry of the 5 years post-completion of the research project, data that are not to be stored for an extended period or retained for future use should be destroyed so that they cannot be retrieved in the future.

6.2. Storage of sensitive research data.

6.2.1 Sensitive data is information that can be used to identify an individual and introduces a risk of discrimination, harm or unwanted attention. It includes for example, identifiable personal data, health and medical data, genetic data, data concerning a person's sex life or sexual orientation.

6.2.2 The storage of physical (paper based) sensitive data requires adequate employee training in the implantation and maintenance of security measures including the use of secured storage equipment behind locked doors and preferably with physical intrusion detection.

6.2.3 The storage of digital data should take place on a Rhodes University server secured by regularly changed strong passwords with additional security measures including backups, authentication, access control, and encryption in place.

[8] POPIA, 2013:30-32

[9] POPIA, 2013:36-38

7. Data-sharing

7.1 Data-sharing refers to the process whereby data resources are made available to multiple users, applications, and/or organizations for purposes of among other things, research collaboration, publication and/or postgraduate examination.

7.2 Data can be shared by means of 1) direct file transfer to one or more individuals or 2) software applications that permit the collaboration of several individuals on research projects.

7.3 Data that have been collected may be shared intra or inter-institutionally. The latter should be governed by the use of Data Transfer Agreements.

7.4 Participants need to have been informed that their (anonymised) data will be shared and to have consented to the sharing.

7.5 Data that was not anonymised at the point of collection must be de-identified before it is shared.

7.6 Data should be shared in accordance with the Five Safes Framework. In terms of this framework secure data access can be provided if five risk dimensions can be addressed:

1. Safe projects: the appropriate use of the data in question.
2. Safe people: the trustworthiness of the people using the data.
3. Safe setting: the degree to which unauthorised use has been limited.
4. Safe data: the existence of disclosure risk in the data.
5. Safe outputs: the degree to which statistical results are non-disclosive.

7.7 Safety aspects relating to projects, people and settings can be addressed through managerial controls, while safety pertaining to data and outputs should be addressed through statistical controls.

8. Effective date of this SOP

26 July 2024 and the next revision date is 26 July 2027, or as deemed necessary by a quorate meeting of RU-HREC.