Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

Stellenbosch University Research Data Management Regulation

January 2024

forward together
sonke siya phambili
saam vorentoe

Stellenbosch University Research Data Management Regulation

Type of document:	Regulation
Purpose:	Management of research data at Stellenbosch University (SU) ("the University") to ensure compliance with legislative frameworks, as well as protecting the University and staff and research participants involved in research through the mitigation and management of inherent risks.
Approved by:	Senate
Date of approval:	24/11/2023
Date of implementation:	01/01/2024
Date of next revision/frequency of revision:	Every five years
Previous revisions:	First version was approved on 27/11/2020 and implemented on 01/01/2021
Regulation owner[1]:	Deputy Vice-Chancellor: Research, Innovation and Postgraduate Studies
Regulation curator[2]:	Division for Research Development, and Library Information Service
Keywords:	Research, Data, Ethics, Personal information, Research participants, Intellectual Property, Confidential, Sensitive, Data classification, GDPR, POPIA, FAIR
Validity:	The English version of this regulation is the operative version, and the Afrikaans version is the translation.

[1] Rules Owner: Head(s) of Responsibility Centre(s) in which the rules functions.
[2] Rules Curator: Administrative head of the division responsible for the implementation and maintenance of the rules

1. Introduction

Stellenbosch University acknowledges that its research data are valuable assets and contribute to the knowledge economy and therefore need to be managed, protected and curated in an appropriate manner. As an institution, we need to interact with our research participants and the data in an ethical and responsible manner.

Research data management is a core part of responsible research conduct; it covers the entire research data lifecycle and ensures the confidentiality, integrity and availability of research data. Within the legislative realm of data protection, we need to understand our role and responsibilities in terms of various legislation and regulatory guidelines[1] on the protection of data, most notably the South African *Protection of Personal Information Act 4 of 2013* (POPIA).

Stellenbosch University, as a public entity, is committed to disseminating its research and knowledge as widely as possible through open access methods, and at the same time understands that there is a balance to be found between openness, good information security practices and compliance with the legislative framework (including, but not limited to privacy and intellectual property law), legal and ethical obligations where data should be kept confidential. Within this framework, we understand that our funders may require that we comply with their own research data management policies.

2. Application of the Regulation

2.1 The Regulation applies to all members of the University (See Section 3 below for a definition of 'member').

2.2 The Regulation should be read together with the University's other relevant policies, regulations, standard operating procedures (SOPs) and guidelines, in particular those specifically mentioned in this document.

2.3 The Regulation is intended for institutional use and does not confer any rights or privileges to a third party.

3. Definitions

Administrative data: Data that are derived from the operation of administrative systems at the university (e.g., data collected for the purposes of registration, transaction and record keeping).

Archiving: Retaining research data in a data centre, archive or repository, where it will be protected in the long term against loss, deterioration, unauthorised or inappropriate access, and future incompatibility.

Anonymous data: Data collected by researchers without any identifying information and without a link to a specific participant or donor.

Coded or de-identified data: Data of which the identifiers have been replaced with a unique number or code and a key exists to decipher the code, allowing the code to be traced back to a specific participant or donor.

---

[1] Examples: Department of Health (2015) *Ethics in health research: principles, structures and processes* (2nd Ed.); Department of Health and SAHPRA (2020) *South African Good Clinical Practice: Clinical Trial Guidelines* (3rd Ed.); *South African National Health Act 61 of 2003*; *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016* (GDPR) (EU); *Health Insurance Portability and Accountability Act of 1996* (HIPAA) (USA).

Confidential data: Data that is not accessible to everyone and is not intended for public dissemination. Refer to the *SU Information Classification Regulation* for more details.

Consent: Consent includes any voluntary, specific and informed expression of will in terms of which permission is given for the processing of personal information (POPIA, 2013). Research consent or 'informed consent' is an indication of agreement to participate in research, based on adequate knowledge and understanding of relevant information, and freely given as required by the Department of Health's, *Ethics in Health Research: Principles, Processes and Structures*, 2nd Edition (2015).

Data citation: Provision of accurate, consistent and standardised referencing for datasets that recognises the producers of the data and allows the impact to be tracked.

Data curation: Active and ongoing process of maintaining and adding value to data throughout its lifecycle. Data curation enables discovery, ensures quality and provides for reuse over time.

Data dissemination: The publication or public transmission, communication or distribution of data through some form of public setting.

Data management plan (DMP): A document describing the manner in which research data will be treated during as well as after the completion of research projects.

Data preservation: A series of managed activities to conserve and maintain both the safety and integrity of data for future use and easy access.

Data repository: A searchable and queryable interfacing entity that is able to store, manage, maintain and curate data/digital objects.

Data reuse: Use of research data for a research activity conducted by people other than those involved in the original collection, management and analysis of the data.

Data safe haven: A data repository in which research data can be stored and accessed in a manner that reliably maintains their fidelity and quality but also ensures that the data are 'safe' in the sense that all relevant social expectations and ethical and legal controls on their use and dissemination are appropriately met.

Data sharing: The practice of providing researchers and other parties with access to research data.

Data steward: Responsible for data management and the expert handling of data processing and administering in compliance with policy and regulatory obligations. The data steward knows how the data is collected, maintained, and interpreted. Stellenbosch University automatically assigns the Principal Investigator (PI) as a steward of the data unless otherwise agreed (legal or ethical requirements met). Stellenbosch University has also defined the role of information curator in the University's *Information Curatorship Regulation* . Any roles or responsibilities assigned to information curators are also assigned to data stewards as defined in this regulation.

Data subject: Any individual person who can be identified, directly or indirectly, via an identifier such as a name, an ID number, location data, or via factors specific to the person's physical, physiological, genetic, mental, economic, cultural or social identity. Data subjects may include, but are not limited to:

- prospective students;
- applicants;
- students;
- alumni;

- research participants;
- employees;
- employment candidates;
- visitors; and
- members of the public (POPIA, 2013).

Data transfer agreement (DTA): A contract between the providing and recipient institutions that governs the legal obligations and restrictions, as well as compliance with applicable laws and regulations, related to the transfer of such data between the parties and the use thereof.

Data types: Research data can be generated for different purposes and through different processes. It can include the following **types of data**:

- **Observational:** data captured in real-time through observation of a behaviour or activity, usually irreplaceable.
- **Experimental:** data collected through active intervention by the researcher to produce and measure change or to create a difference when a variable is altered.
- **Simulation:** data generated by imitating the operation of a real-world process or system over time using computer test models. For example, to predict weather conditions, economic models, chemical reactions, or seismic activity.
- **Derived or compiled:** involves using existing data points, often from different data sources, to create new data through some sort of transformation, such as an arithmetic formula or aggregation.
- **Reference or canonical:** a (static or organic) conglomeration or collection of smaller (peer-reviewed) datasets most probably published and curated. For example, gene sequence databanks, chemical structures, or spatial data portals.

De-identification: A process of detecting identifiers (e.g., personal names and identity numbers) that directly or indirectly point to a person (or entity), and replacing them with a number, symbol or letter. A key exists to decipher the code allowing linkage of the code to a specific individual.

FAIR Principles for scientific data management: Principles that propose that all scholarly output should be Findable, Accessible, Interoperable, and Reusable. The principles refer to three types of entities: data (or any digital object), metadata (information about that digital object), and infrastructure.

Fair use: A doctrine that permits the use of copyrighted material like books, journals, music and artwork without requiring permission from the copyright holder. It provides a balance between the just demands of rights holders and the need for people to use copyright material for education, research, in libraries and repositories.

GDPR: *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016*, also referred to as the European Union's (EU) *General Data Protection Regulations*, effective from 25 May 2018.

HIPAA: The United States of America's *Health Insurance Portability and Accountability Act of 1996* provides data privacy and security provisions for safeguarding protected health information (PHI).

Identifiable data: Data collected by a researcher with identifying details e.g., name, folder number or address.

Indigenous knowledge (IK): Local knowledge that is unique to a culture, society or community, has been developed by that indigenous community and is associated with the cultural and social identity of that indigenous community. IK includes—

- knowledge of a functional nature;
- knowledge of natural resources; and
- indigenous cultural expressions (*Protection, Promotion, Development and Management of Indigenous Knowledge Act 6 of 2019*).

Indigenous community: Any recognisable community of people—

- developing from, or historically settled in a geographic area or areas located within the borders of the Republic;
- characterised by social, cultural and economic conditions, which distinguish them from other sections of the national community; and
- who identify themselves as a distinct collective (*Protection, Promotion, Development and Management of Indigenous Knowledge Act 6 of 2019*).

Intellectual property (IP): IP encompasses registerable and non-registerable inventions, expertise, trademarks, trade secrets, copyrights, designs and plant breeders' rights which have come about through the mental efforts, insight, imagination, knowledge and creativity of humans (*SU Policy on Intellectual Property: Protection and Commercialisation* (2023)).

Member of Stellenbosch University (SU): All Stellenbosch University staff members, research students, post-doctoral fellows, external workers and research collaborators who carry out research under the University's auspices.

Metadata: Structured information about an information resource that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage that resource.

There are many different kinds of metadata. In the context of research data management, the following types of metadata might be captured:

- **Descriptive metadata** to help discovery, such as the author or title of a dataset - this metadata is often a mixture of manually and automatically generated information. Descriptive information might be as simple as a few fields or might be very complex if a rich domain ontology is used.
- **File metadata** embedded in files, such as information about the size and resolution of images, or the equipment used to generate the file - this kind of metadata can normally be automatically extracted.
- **Preservation metadata** to record information about the data itself, and any preservation actions that might have been performed - this metadata is normally created automatically by preservation software.
- **Structural metadata** about how a dataset has been assembled - this metadata would normally be generated automatically.
- **Machine-generated metadata** to record details about equipment or software and its setup as used for a particular research activity - this metadata would be generated automatically but may be in proprietary formats.

Open access (OA): A practice to provide unrestricted access to research works such as articles, data, software, etc., available freely, immediately, and permanently online to the public without financial and technical barriers.

Open data: Data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and share alike.

Personal information: Information relating to an identifiable, living, natural person, and where it is applicable, an identifiable, existing juristic person, including, but not limited to—

- information relating to the race, gender, sex, pregnancy, marital status, national, ethnic or social origin, colour, sexual orientation, age, physical or mental health, well-being, disability, religion, conscience, belief, culture, language and birth of the person;
- information relating to the education or the medical, financial, criminal or employment history of the person;
- any identifying number, symbol, e-mail address, physical address, telephone number, location information, online identifier or other particular assignment to the person;
- the biometric information of the person;
- the personal opinions, views or preferences of the person;
- correspondence sent by the person that is implicitly or explicitly of a private or confidential nature or further correspondence that would reveal the contents of the original correspondence;
- the views or opinions of another individual about the person; and
- the name of the person if it appears with other personal information relating to the person or if the disclosure of the name itself would reveal information about the person. (POPIA, 2013)

POPIA: *Protection of Personal Information Act 4 of 2013* (effective date 1 July 2020). ASSAF has developed a *POPIA Code of Conduct for Research* (2023).

Primary research data: Research data that is collected by a researcher from original or first-hand sources, using methods like surveys, interviews, or experiments. It is collected with the research project in mind, directly from primary sources.

Principal investigator (PI): The PI, also referred to as the **researcher,** is directly responsible for the integrity and management of the design, conduct and reporting of a research project and for managing, monitoring, and ensuring the integrity of any collaborative relationships. The PI is accountable to the University and to external funders. The PI also provides scholarly leadership and bears primary responsibility for technical, programmatic, fiscal, and administrative requirements of the project, including direct responsibility for the collection, recording, storage, retention, and disposal of research data. Stellenbosch University automatically assigns the PI as the steward of the research data.

Research data: Recorded information, obtained during a research process, regardless of form or the media on which it may be recorded. The term includes computer software (computer programmes, databases and documentation thereof), and records of a scientific or technical nature. The term does not include information incidental to research administration such as financial, administrative, cost or pricing, or management information. In practice, research data include both intangible data (statistics, findings, conclusions) and tangible data. Tangible data include, but are not limited to notes, printouts, electronic storage, photographs, slides, negatives, films, scans, images, autoradiograms, electro-physical recordings, gels, blots, spectra, cell lines, reagents, modified organisms, specimens, consent forms, case report forms, collected organisms and other materials that are relevant to the research project.

Research data lifecycle: A model that describes and identifies the steps to be taken at the different stages of the research cycle to ensure successful data management. The data lifecycle starts before the research project begins and has six components:

- Data management planning
- Data collection
- Data processing and analysis
- Data dissemination
- Data curation
- Data reuse

Research data management (RDM): 1) Planning for the manner in which research data will be managed during and after the research process, and 2) controlling the collection, processing, analysis, sharing, dissemination, curation and reuse of research data.

Research ethics committees (REC): The Ethics Review Committees have a vital function in ensuring that all research activities at Stellenbosch University are conducted within national and international accepted standards and legislation with respect to ethics in research.

Research participant: A person who participates in research by being the target of observation by researchers.

Secondary use of data: Use in research of data originally collected for other purposes.

Sensitive data: Data that identifies individuals, species, objects or locations, and carries a risk of causing discrimination, harm or unwanted attention. There are three main types of sensitive information: personal information, business information and classified information.

4. Purpose of the Regulation

The purpose of the Regulation is to:

4.1 Enable Stellenbosch University to comply with all the legal, ethical and contractual obligations stemming from the relevant South African and international legislation regarding research data management.

4.2 Provide a foundation for defining the principles of governing research data management and protection of research subjects.

4.3 Provide a framework to define the responsibilities of all Stellenbosch University members regarding research data management so that research data generated at Stellenbosch University will always have a data steward responsible for its preservation.

4.4 Guide researchers and students in the best practices of managing research data, and making the data available, where appropriate, to the widest possible audience for the highest possible impact.

5. Objectives of the Regulation

The Regulation is intended to ensure that research data created, received or analysed as part of the research process are:

5.1 Compliant with legal obligations, ethical responsibilities, contractual requirements and the rules of funding bodies.

5.2 Findable, accessible, interoperable and reusable (FAIR).

5.3 In line with appropriate data sharing and open access principles.

5.4 Secure and safe with appropriate measures taken in handling internal, confidential and restricted data.

5.5 Preserved for its lifecycle with the appropriate high-quality metadata. In the absence of specific requirements (funder or legal), the default period for research data retention is ten years from date of project conclusion or publication.

6. Principles of the Regulation

The Regulation is based on the following principles:

6.1 Good management of research data assures that data are capable of supporting future analysis.

6.2 Good management of research data assures that all operations performed on data are traceable.

6.3 Good management of research data facilitates use of data by people other than those involved in their original collection, management, and analysis.

6.4 Good management of research data is an integral part of sound research practice.

6.5 Good management of research data allows reliable storage of and access to research data, protects the intellectual and financial investment made in their creation, enables data to be shared and cited (where appropriate), and prompts new and innovative research activities and outcomes.

6.6 The Regulation is, as far as possible, aligned with the FAIR data principles.[2]

7. Scope

7.1 The Regulation applies to all research data created, received or analysed as part of the research process as carried out by any member of the Stellenbosch University including contract research and research-related consultancy work carried out by University researchers.

7.2 The Regulation applies to all external research collaborators in cases where a member of the Stellenbosch University is the principal investigator of the study unless otherwise contractually agreed.

7.3 The Regulation does not apply to:

7.3.1 Physical materials, collections of items and software. Considerations regarding the curation and exploitation of these collections vary according to the specific discipline.
7.3.2 Administrative data.

[2] M.D. Wilkinson, M. Dumontier and B. Mons (2016) The FAIR guiding principles for scientific data management and stewardship. Scientific Data, Volume 3. DOI: 10.1038/sdata.2016.18.

8. Data acquisition and management

Best practices in research data management promote research integrity and collaborative opportunities. Clear and accurate records of the research methods and data sources, including any approvals granted, during and after the research process must be kept.

8.1 Institutional requirements

8.1.1 Data acquisition and management should be performed in accordance with the *Policy for Responsible Research Conduct at Stellenbosch University* as well as other measures and guidelines where appropriate.
8.1.2 Where personal information will be collected the *Stellenbosch University Personal Information Impact Assessment* must be considered as per the *Stellenbosch University Data Privacy Regulation*.
8.1.3 Researchers must collect data using appropriate methodology and recording practices and apply appropriate quality assurance mechanisms and systems.
8.1.4 Raw data must be recorded in hard copy or electronically as appropriate for each research field and with due consideration given to the advantages and disadvantages of different methods.
8.1.5 Data must be properly stored and protected in order to allow for the validation of research findings, to establish the priority of the data, allow for reanalysis, if necessary, to comply with requirements of funders etcetera. Please refer to Section 8.5 for more information.
8.1.6 Processes should be established to protect data from accidental loss, damage or theft.
8.1.7 The duration of appropriate data archiving must be determined by each research environment, giving due consideration to requirements of all stakeholders, including funders, collaborators and legal requirements. In the absence of specific requirements, the default period for research data retention is ten years from date of publication.

8.2 Institutional ownership of research data

8.2.1 Both the researcher and the University have responsibilities and rights regarding access, usage and maintenance of primary research data.
8.2.2 Research data belong to the University unless there are specific terms regarding intellectual property rights in a research-related contract. Before research is initiated, it is important to delineate the rights, obligations, expectations, and roles played by all interested parties.
8.2.3 Stellenbosch University can therefore be held accountable for the integrity of the data even after the researchers have left the university.
8.2.4 The primary responsibility for the management of primary research data remains with the laboratory or department or research environment where it was created, in compliance with this regulation. However, in accordance with principles of academic freedom and intellectual integrity a researcher may be allowed to retain copies of the research records and portions of materials created by him/her in the course of the research.
8.2.5 Data created or collected in the course of research may be transferred to another institution. However, in all cases the transfer shall be subject to the terms of a data transfer agreement negotiated by the Research Contracts Office, Division for Research Development, and signed by the authorised representatives of the relevant parties.
8.2.6 Researchers must familiarise themselves with the *Stellenbosch University's Policy on Intellectual Property: Protection and Commercialisation* *of* which is also applicable to the context of data ownership, sharing and dissemination.

8.3 Data Sharing

8.3.1 Validated research data can be shared where appropriate once researchers have had the opportunity to establish the priority for their work through publication. Data sharing must comply with

ethics, legal and contractual requirements (e.g., data privacy, intellectual property rights, access to information, etc.).
8.3.2 The conditions for sharing and use through transfer of data to other institutions must be stipulated in a data transfer agreement (officially signed by authorised representatives of the relevant parties), to ensure that all the appropriate legislative and regulatory considerations as well as the requirement of informed consent from research participants for data sharing are correctly addressed.
8.3.3 Collaborative research databases or repositories should be managed according to the principles contractually agreed upon for managing research collaborations. Where appropriate, collaborative data repositories should be formally managed by the appointment of a steering committee and the development of written operating procedures that set out the conditions for the use and transfer of data.

8.4 Data Dissemination

8.4.1 Certain funders have specific data dissemination policies and researchers should acquaint themselves fully with the requirements and comply where applicable.
8.4.2 The University encourages researchers to publish their research data, where appropriate, through means which are supportive of the FAIR Data Principles, subject to contractual requirements. More specifically, the University encourages its researchers to publish their research data in the following ways:

a) as supplementary data published with journal articles;
b) data articles published in data journals;
c) data published via third-party digital data repositories; and
d) data published via Stellenbosch University Research Data Repository *(SUNScholarData)* under the *Stellenbosch University Research Data Repository: Regulation* which regulates its operation.

8.5 Data categories and storage requirements

The University subscribes to the open access principle, and data should therefore be as accessible as possible, providing that the University complies with relevant legislation and contractual obligations. For data storage, the University prioritizes the storage options in the following order as described in Sections 8.5.1 – 8.5.3:

8.5.1 Researchers must ensure that both active and published research data are securely stored. The data must be protected against loss and unauthorised access, by making use of institutional storage of the University.
8.5.2 Data may only be stored on stand-alone drives or portable media if no institutional data storage or data transfer solution is feasible, or where a specific funder requires such (for instance during a pharma sponsored clinical trial), and after confirmation received from the Information Technology Division that the storage complies with the specific requirements set out by the funder or legislation, as the case may be.
8.5.3 If non-institutional cloud services are considered for storage of data, researchers should consult the Information Technology Division to ensure that it complies with the specific requirements set out by the funder or legislation, as the case may be.
8.5.4 Researchers must ensure that any form of personal identifier is removed from data to which wider access is given so that data cannot be linked to an identifiable living individual.

8.6 Levels of access

Different levels of access including public, internal, confidential, and restricted access, can be placed on stored research data. Researchers must ensure that they assign the most relevant level of access to their data and consider the type of consent that was given by research participants, where relevant.

Access to datasets may be controlled according to the following access settings:

a) Open Access Setting: Publicly accessible subject to licence conditions.
b) Private Access Setting: Access to individual user accounts or collaborative spaces. This setting may only be accessible to members of Stellenbosch University or external research collaborators from other institutions.
c) Restricted Access Setting: Available only to users who request access to research data and are authorised to access such data by an authorising agent.
d) Access controls should apply to:
   - i Research data with commercial potential.
   - ii Medical research containing confidential patient data.
   - iii Non-medical related research data that contains personal identifiers.
   - iv Research data containing culturally sensitive information.
   - v Third party data which have contractual agreements.

e) Under embargo: Research data held under embargo for a fixed period and made available on either an open or restricted basis upon the expiry of the embargo period.

9. Ethical consideration in use of data

Core ethical principles apply to all forms of research and therefore the general requirements for research participant engagement, social value, scientific validity and integrity, informed consent, risk/benefit ratio, protection of privacy and confidentiality are the same for all research. Both current and future uses of data should be considered.

9.1 Ethical use

9.1.1 Researchers should manage research data in an ethical manner. Also refer to the *Policy for Responsible Research Conduct at Stellenbosch University*.
9.1.2 Any conflicts of interest associated with research data should be declared to the University.
9.1.3 Research that involves interaction with humans, use of animal material, as well as research that may be hazardous to the environment or may lead to potential biosafety risks, are subject to ethical review as well as legislative requirements. The Division for Research Development can provide guidance on the process.
9.1.4 The detailed ethical aspects relating to the treatment of research data at Stellenbosch University are regulated by the standard operating procedures and guidelines associated with the University's various research ethics committees. These standard operating procedures and guidelines provide assurance that human and animal use is properly governed and conforms to the highest ethical standards.

9.2 Sensitive data should be managed accordingly

9.2.1 Where research data is obtained from human participants, researchers must obtain ethical approval from a National Health Research Ethics Council (NHREC)-registered Research Ethics

Committee as well as informed consent from such participants beforehand. The Divisions for Research Development and Information Governance can provide guidance.
9.2.2 When personal information is collected from human participants the confidentiality of such participants should be safeguarded through the de-identification or anonymisation of any personally identifiable information when appropriate.
9.2.3 As sensitive data carries a risk of causing discrimination, harm or unwanted attention, researchers are encouraged to conduct disclosure risk assessments before, during and after research data are collected in order to ensure that personal information pertaining to research participants remains confidential. A *Personal Information Impact Assessment* is available from the Division for Information Governance.
9.2.4 Researchers should also take reasonable steps to ensure that sensitive data is not breached accidently or as a result of negligence in the handling of the data. Identifiable data must always be stored separated from other data and should not be stored on portable devices. The Information Technology Division can provide guidance and assistance.

9.3 Ethical reporting of research data

9.3.1 Researchers should report research data in accordance with internationally accepted research practices.
9.3.2 Researchers should explain the processes used to protect research participants including the steps taken to ensure anonymisation and de-identification of data.
9.3.3 Researchers should refrain from plagiarising research data which has already been produced by others. In the event that researchers make use of research data which had been produced by others the researchers should credit their sources by acknowledging them in citations and ensure that secondary use of data is permitted (also refer to Section 9.4, Secondary Use).
9.3.4 Researchers should take care to avoid falsifying research data. Any falsification of data will be regarded as research misconduct.

9.4 Secondary Use

9.4.1 A data transfer agreement should accompany secondary, closed data that comes from a source other than the researcher. The Research Contracts Office, Division for Research Development manages all data transfer agreements and can provide guidance on the process.
9.4.2 Review by the relevant research ethics committee is required for all research that relies exclusively on secondary use of confidential information, even when the information will be deidentified and the process of data linkage, recording or dissemination of results does not generate identifiable information. Identifiable information that is available in the public domain is exempt from ethics clearance, but a data management plan will still be required.
9.4.3 Research ethics committees can approve a waiver of informed consent if research is retrospective and uses anonymous and aggregated data, but they may not approve an informed consent waiver for data that has only been de-identified, unless under very specific conditions stipulated in the National Department of Health's *Ethics in Health Research* document (2015).

9.5 Informed Consent

9.5.1 According to the principles of lawfulness, fairness and transparency, personal information cannot be collected, shared or distributed without informed consent. Researchers must acquaint themselves with the provisions of the relevant national and international legislation. The Stellenbosch University

Health Research Ethics Office has developed informed consent guidelines and templates that researchers should use[3].
9.5.2 Research participants should be unambiguously informed of what will be done with their data and give consent. Subsequently, data processing should be done accordingly.
9.5.3 The informed consent of participants in a project needs to include sharing, preservation and long-term use of their personal data as well as any limitations to these actions.
9.5.4 Signed consent forms and other documents and records linked to the ethical conduct of research must be stored by the principal investigator in a safe and secure manner according to the Department of Health's guidelines for research ethics.

9.6 Cultural sensitivities and indigenous knowledge

9.6.1 Research with and about indigenous communities should be conducted in a way that ensures a process of mutual engagement between the researcher and the indigenous people. Free, prior and informed consent (FPIC) should be obtained. Adhering to the principles of FPIC, the communities should be in a position to make an informed decision on the potential risks and benefits of proposed research.
9.6.2 Indigenous people and local communities could develop their own ethical guidelines for external researchers. Researchers must be aware of and adhere to these Codes.
9.6.3 Research data management planning requires the inclusion of indigenous intellectual and cultural property rights and sensitivities. Special attention must be given to data collection, storage, disclosure and reuse of indigenous-related data.
9.6.4 Refer to the *Protection, Promotion, Development and Management of Indigenous Knowledge Act (Act no. 6 of 2019)* and the *United Nations Declaration on the Rights of Indigenous Peoples* of which South Africa is a signatory[4].

10. Data retention

10.1 Research data related to publications may be made discoverable or accessible to other researchers via the institutional *SUNScholarData* repository. Datasets deposited in this repository will be retained for a minimum of ten years from the date of publication or for the retention period specified in any applicable third-party policy or contract term, whichever is longer. The retention will be performed according to the *Stellenbosch University Research Data Repository: Regulation*.

10.2 Research data not in a repository, must be retained securely according to applicable retention and disposal requirements considering future research needs, complying with legal, ethical, research funder and collaborator requirements, and with particular concern for the confidentiality and security of the data. In the absence of specific provisions, the default period for research data retention is ten years from the date of deposit.

10.3 The appropriate data preservation format should be selected for each research dataset. Digitisation of print data should be considered for long-term preservation purposes provided that the storage format is robust and long-term storage data integrity mechanisms can be ensured (prevention of bit-corruption, e.g., write-once, read many storage).

---

[3] *SU Health Research Ethics Office documents for researchers*

[4] Also refer to the *Protecting and Promoting Indigenous Peoples Rights in Academic Research Processes: A Guide for Communities in South Africa*.

10.4 If the research data will be retained for the long term, researchers must make allowances for storage costs and metadata maintenance in their research data management planning and in the project budget. Access permission can be granted to collaborators (as specified in research collaboration contracts).

11. Data destruction

11.1 Destruction of research data, regardless of format, created and kept by members of Stellenbosch University cannot be undertaken without first considering the relevant contractual obligations or the commitments made during the informed consent process with research participants.

11.2 Secure destruction of research data involves using irreversible methods to ensure that the data is no longer usable. It is particularly critical that confidential or sensitive data remains unreadable. Depending on the storage medium, it may be necessary to utilise software that permanently erases data. Researchers should consult with the Information Technology Division for recommendations.

11.3 Secure destruction services should be employed for the destruction of physical records, including paper and digital media such as CDs, DVDs, flash drives and other media formats.

12. External funder and external collaboration requirements

12.1 The researcher acting as principal investigator is responsible for all obligations that are outlined in the research contract. The Research Contracts Office in the Division for Research Development provides advice to researchers in meeting their responsibilities for all contractual obligations.

12.2 All contractual obligations should be adhered to, such as obligations regarding confidentiality, intellectual property requirements, security of research data and conflict of interest.

12.3 Researchers must be aware of, and comply with, external funder requirements on research data management. However, external funding and collaboration arrangements must comply with all applicable Stellenbosch University policies and procedures, such as ethics approval and this regulation. Many funders have open research policies and require a research data management plan and open access publication of supporting research data. Funders may provide their own data management plan templates and specified terms of storage of research data in repositories.

12.4 The required data management plan, which may be part of the final evaluation of the project, must be updated regularly and exist as a versioned document.

12.5 Researchers must know whether the funders offer separate financing for the storage of research data during the research or for sustainable storage afterwards.

12.6 For collaborative projects involving one or more external partners, access to and publication of research data should normally be included in collaboration agreements, and data sharing agreements should be entered into where required.

13. Legal aspects

13.1 Research data should be managed in compliance with all relevant legislation, regulatory requirements and contractual obligations (See Section 12, External funder and external collaboration requirements).

13.2 The ownership of research data must be clarified prior to the commencement of a project and should be documented via a research data management plan and contractually agreed.

13.3 Future storage and reuse are directly affected by the intellectual property rights applicable to the research data. Intellectual property law in South Africa refers to all legislation concerning copyright, patents, designs, and trademarks, and more specifically intellectual property created through public funding.

13.4 Where research involves the use of research data owned by a third party, researchers must abide by any applicable data transfer agreement, licences or terms of use governing the research data.

13.5 Where multiple requirements apply to a particular research project, the applicable legal requirements supersede all other requirements.

13.6 Data Transfer Agreements

13.6.1 A data transfer agreement represents both legal and ethical aspects associated with research data being shared between parties. The purpose of a data transfer agreement is to govern the transfer of research data between the providing and recipient institutions, recognising that such data will be used according to the description of a research project or study protocol. Where the research data being shared comprises personal information, the data transfer agreement ensures that parties to the agreement process such data in a lawful and ethical manner.

13.6.2 The data transfer agreement specifically governs the parameters for access, storage, duration of use, destruction of the data within the context of the specific project and the purpose of transfer/sharing.

13.6.3 The data transfer agreement needs to consider local data privacy laws and, if applicable, the data privacy laws of the territory from which the data was collected and thus the researcher/research team needs to be clear about the encryption format of the data (whether deidentified, anonymised, or pseudonymised).

13.7 Fair Use

13.7.1 In instances where Stellenbosch University researchers acquire research data from/through external parties towards a secondary use, such researchers should be aware of the fair dealing or fair use doctrines applicable to the use of such data or the restrictions of such use as detailed in a data transfer agreement (also refer to Section 9.4, Secondary Use).

13.7.2 In the case of external parties acquiring Stellenbosch University's research data (most likely although not exclusively via the University's institutional research data repository) use of such research data will be subject to the fair dealing doctrine as embodied in Section 12 of the *Copyright Act 98 of 1978*. (Such use is also subject to the *Stellenbosch University Research Data Repository: Regulation* or a signed data transfer agreement).

14. Regulation governance

14.1 Roles and responsibilities

All staff and research students involved in research under the University's auspices have a responsibility to manage research data they create or acquire, and to maintain it effectively and in line with the University's regulations.

Research data management is the responsibility of the Stellenbosch University researcher who acts as the principal investigator of any research project, including collaborative projects. Where research is undertaken by a research student, it is the responsibility of the supervisor to ensure that the student has a clear understanding of appropriate research data management practice in line with this regulation.

A. Researchers acting as principal investigators are automatically assigned as stewards of the research data and are responsible for:

1. Ensuring that researchers, staff and students understand the requirements for research data management and engage with training and development support in the handling, curation and archiving of research data as necessary.
2. Ensuring that researchers, staff and students comply with the requirements for research data management as specified by the University.
3. Collecting and managing research data in accordance with the principles and requirements in this regulation.
4. Producing and adhering to a data management plan to address the requirements.
5. Planning for archiving and curation of data after the completion of research.
6. Ensuring that active research data are stored securely and protected from loss and unauthorised access and that access is not limited to a single person.
7. Ensuring that active research data are accessible by another authorised person during the course of research to guarantee access to the data by the University in case of need.
8. Producing metadata and documentation to describe data, sufficient to understand what research data exist, why, when and how it was generated and access restrictions and mechanisms.
9. Ensuring that on completion of research, all relevant research data are archived, maintained, deposited or disposed of appropriately in accordance with the provisions of this regulation.
10. Ensuring that, should they leave the institution before completion of a research project, copies of any data produced under the auspices of the University are deposited in an appropriate medium of storage at Stellenbosch University, before their researchers' departure.
11. Meeting all the requirements in relation to research data placed on their research by funding bodies, regulatory agencies, third party data providers and collaborating institutions or under terms of a research contract with the University.

B. Heads of Department are responsible for:

Governance and oversight of research data management in the department, including compliance with this regulation, departmental standards and procedures, and professional frameworks and standards. Departments will be supported in the development and implementation of local arrangements in fulfilment of this responsibility.

C. Stellenbosch University research supporting divisions/entities:

The Division for Research Development and Library and Information Service are co-custodians of this regulation.

C.1 Division for Research Development is responsible for:

1. The Regulation's formulation, approval, review and communication as well as the interpretation and guidance in respect of the implementation of this regulation.
2. Institutional awareness through researcher workshops and support (e.g., researcher grant applications, ethics requirements and research contract-related services).

C.2 Library and Information Service is responsible for:

1. Providing research data management support services in collaboration with other divisions within the Library and Information Service.
2. Establishing standards for data management for research irrespective of funding including requirements for producing data management plans.
3. Provide specific operational services relating to research data management:

   3.1 Advocacy:
   a) Increasing awareness of research data management and demonstrate the value of research data management services to the academic community.
   b) Promoting data dissemination and re-use.
   c) Articulating the benefits of data dissemination and re-use.

   3.2 Research data management information services:
   a) Providing assistance and advice to researchers and research students on the following research data management-related activities by means of consultation and training, and providing relevant information via library guides and research data management webpages:
      i. Creating data management plans and providing access to data management plan tools where applicable.
      ii. Finding data (e.g., searching data repositories).
      iii. Analysing data (including visualisation).
      iv. Organising data (including advice on file naming conventions and metadata).
      v. Storing data (including directing researchers to various existing data repositories).
      vi. Disseminating data (where applicable, advise researchers on platforms to share datasets).
      vii. Citing data.

   3.3 Hosting and managing the Stellenbosch University Research Data Repository, known as *SUNScholarData*, in which the University's research data associated with published research findings are archived:
   a) Maintaining the repository platform.
   b) Managing regulations for the use of the repository.
   c) Managing ingest of datasets.
   d) Managing metadata and description of datasets.
   e) Overall content management of repository.

C.3. Division for Information Governance is responsible for:

1. Managing data privacy regulations.
2. Providing guidance and training for the use of Personal Information Impact Assessment.
3. Managing the personal information incident and breach handling procedure.

C.4 The Information Technology Division of the University has a responsibility to ensure that systems are in place to support and reinforce good research data management. They are also responsible for:

1. Providing access to infrastructure, active data management platforms and services for the secure capture, storage, backup, computing and sharing of research data that allows researchers to meet the requirements under this regulation and that of the funders of their research.
2. Assist and advise on the costing for active research data storage.
3. Providing advice on the implementation of technical data security mechanisms.
4. Providing advice on the management and storage of big data.
5. Assist and advise on the selection of ICT infrastructure and applications for RDM.

14.2 Implementation of the Regulation

14.2.1 The Division for Research Development is responsible for coordinating the implementation of this regulation.
14.2.2 Training will be provided by the Library and Information Service and the Division of Information Governance.
14.2.3 The Library and Information Service and Information Technology Division will provide tools for data classification, data management plans, etc.

14.3 Support and advisory services

An RDM Advisory Group has been established and will provide collective advice on a case-by-case basis on research data management requirements. Researchers can contact the RDM Advisory Group with regard to specific questions.

14.4 Actions taken in case of non-compliance with the Regulation

Non-compliance with this regulation will be considered by the RDM Advisory Group and appropriate steps to rectify the situation will be recommended.

15. Data management infrastructure

Where specific data management solutions such as preservation platforms, safe havens, repositories, etc. are required, researchers should consult with the RDM Advisory Group before acquiring bespoke solutions.

16. Data management plans

Research data management covers all phases of the research data lifecycle. Protocols that relate to ownership, documentation, security, sharing and disposal of research data must be implemented throughout each stage of the research process. Research data can continue to be used, reviewed and modified via follow-up projects beyond the scope of the original research project. The data management plan can help researchers document every stage of the research data lifecycle. Furthermore, a data management plan is an important tool for ensuring that researchers are aware and have a plan to adhere to policy and regulation requirements before they start collecting their data.

16.1 Data management plans must be compiled for all research projects that are expected to process data, irrespective of whether submission of such plans is required by research funders. Researchers

should consult funder guideline documents directly in order to assess the specific requirements of the different research funders.

16.2 Data management plans must explicitly address data classification, data acquisition, data curation, processing, analysis, storage, transfer and use (including sharing) according to data type, confidentiality, retention, publication of data, and, where applicable, other valuable outputs, created or collected during a project. The plans should also include measures that will be put in place in order to facilitate the reuse of research data and set out responsibilities for the management and control of the research data.

16.3 For collaborative projects involving one or more external partners, access to and publication of research data should be included in collaboration agreements, and data transfer agreements should be entered into where required.

16.4 The university's generic data management plan templates should be used unless funders prescribe their own template. Templates can be retrieved via the Library and Information Service website.

16.5 Data management plans must be completed prior to the commencement of data collection/generation/creation.

16.6 In instances where research funding is sought, and data management plans must be completed and submitted as part of research grant proposals, funder-specific templates and guidelines for data management plans must be used where applicable.

16.7 In instances where ethical clearance is required, data management plans must be submitted as part of the ethics approval process.

16.8 In instances where specific legislation (e.g., GDPR) requires the submission of data management plans, the final document must be signed by an authorised official. Researchers must consult with the Research Contracts Office, Division for Research Development for assistance.

16.9 Data management plans must be reviewed at least annually during the lives of research projects and be updated where necessary as per funder requirements in order to ensure that they continue to be relevant for the purposes of their associated research.

16.10 Costs of research data management

16.10.1 Appropriate resources (time and financial resources) associated with delivering a data management plan should be considered at the earliest opportunity and should be allocated in grant proposal budgets, where possible.

16.10.2 The University may provide a limited amount of central storage for free, but charges may be incurred for requirements exceeding those limits.

16.10.3 Legitimate costs associated with managing and publishing data include:

a) Storage and computing of active data in volumes that exceed institutional provision.
b) Digitisation.
c) Data publication, where there is a charge associated with deposit in a data repository, or where volumes exceed institutional provision.
d) Transcription of interviews.
e) Secure destruction services.

f) Recruitment of specialist staff to support the management of complex datasets and the preparation of these data for archiving and publication should be budgeted and allocated for in grant proposals.

16.10.4 Researchers should consult with the RDM Advisory Group for recommendations.

17. Incident reporting

In the event of possible data leaks, personal information breaches or other incidents that need to be reported. The Office for Research Integrity and Ethics, Information Technology Division as well as the Division for Information Governance should be informed.

18. Related documentation

Table 1 is a list of currently existing regulatory instruments that are linked in some way to the research data management governance framework. This framework relates to the internal governance framework that exists at Stellenbosch University.

Table 1. Summary of institutional regulatory documents related to research data management.

Name	Document custodian(s)	Type of document	Purpose
Policy on Mandatory Self-Archiving of Research Output	Library and Information Service	Policy	To regulate the submission of electronic copies of theses and dissertations and the depositing of full text peer-reviewed journal articles into the university's institutional repository.
Information Technology (IT) Policy Definitions	Information Technology Division	List of policy definitions	To consolidate certain terminology definitions used in different ICT policies.
Information Security Regulations	Information Technology Division	Regulation	To provide an outline for the scope of information security at the university.
Identity and Access Management Policy	Information Technology Division	Policy	To establish principles and provisions by which the identities, specifically the electronic identities, of natural persons who have a relationship with the Stellenbosch University as well as their access privileges are managed across the university.
Stellenbosch University Password Regulations	Information Technology Division	Regulation	To establish standards and guidelines for the creation of strong passwords, the protection of those passwords, and the frequency of change.

Policy for Responsible Research Conduct at Stellenbosch University	Division for Research Development	Policy	To promote and ensure research integrity and the ethical conduct of research.
Stellenbosch University Health Research Ethics Committee (HREC): Terms of Reference and Standard Operating Procedures	Division for Research Development and Support, Faculty Medicine and Health Sciences	Procedure	To protect the dignity, rights, safety, and well-being of all human participants in health-related research.
Research Ethics Committee: Social, Behavioural and Education Research (REC: SBE): Terms of Reference and Standard Operating Procedure	Division for Research Development	Procedure	To promote and ensure a culture of ethically responsible research at Stellenbosch University in the Humanities.
Research Ethics Committee: Animal Care and Use Standard Operating Procedures and Guidelines	Division for Research Development	Procedure	To ensure that all animals used in research and teaching are cared for and used in ways judged to be scientifically, technically, and humanely appropriate.
Research Ethics Committee: Biological and Environmental Safety (REC:BES) Standard Operating Procedures	Division for Research Development	Procedure	Management of Biosafety and Environmental Ethics in Research at Stellenbosch University.
Policy on Contract Research Management at Stellenbosch University	Division for Research Development	Policy	To stipulate the way in which research related contracts at Stellenbosch University shall be dealt with
SU Policy on Intellectual Property: Protection and Commercialisation	InnovUS	Policy	To regulate and provide for the identification, protection and commercialisation of Intellectual Property by Stellenbosch University, and in particular as may arise in the course of teaching and learning and/or research and development activities conducted at Stellenbosch University, and in compliance with the applicable regulatory and legislative framework.
Risk Management Policy	Risk Management and Campus Security	Policy	To guide the management of risks at Stellenbosch University
Regulations for recruiting Stellenbosch University persons as research participants and for conducting research on Stellenbosch University-held personal and institutional information	Division for Information Governance	Regulation	Regulations covering institutional and gatekeeper permission requirements at Stellenbosch University

SU Information Classification Regulation	Division for Information Governance	Regulation	To establish a classification framework that enables information curators to identify and classify the information for which they are responsible.
SUNScholarData Repository: Regulation	Library and Information Service	Regulation	To govern the operation and use of the Stellenbosch University Research Data Repository
SU Data Privacy Regulation	Division for Information Governance	Regulation	To establish and enable an institutional framework for the processing of personal information
Stellenbosch University Curatorship Regulation	Division for Information Governance	Regulation	This regulation, within the context of Stellenbosch University: clarifies the information- governance and management responsibilities of Responsibility Centre heads; defines the role of information curators and deputy information curators; establishes the mandate for an information curators oversight committee; establishes responsibilities for defining information curator required competencies and capabilities, and establishes responsibilities for ensuring the provision of adequate training for information curators.