

Stellenbosch University Research Data Repository: Regulation¹

Stellenbosch University Research Data Repository: Regulation	1
1. Preamble	3
2. Definitions	3
3. Purpose.....	6
4. Service	7
4.1 Management and administration	7
4.2 Eligible Content	7
4.3 Users.....	7
4.4 Service levels	9
4.5 Disclaimer	9
5. Submissions.....	10
5.1 Eligible Depositors.....	10
5.2 Terms and conditions of deposit.....	10
5.3 Embargoes.....	11
5.4 Minimum requirements	11
5.5 Exemptions from DOI allocation	12
6. Selection and Appraisal	12
7. Processing and Treatment.....	14
8. Ingestion and Storage.....	15
8.1 Ingestion	15
8.2 Security of Data	15
8.3 Logical Security.....	15
8.4 Privacy and Confidentiality.....	16
8.5 Storage Quota	16
8.6 Compliance with this Regulation.....	17
9. Metadata Assignment	17
9.1 Metadata Records	17
9.2 Mandatory metadata	18
9.3 Language	18
10. Access	19
11. Reuse	19
11.1 Openly Accessible Datasets.....	19

¹ The bulk of this document has been adapted from the content found in the University of Reading’s Research Data Archive Policies. Use of the content has been granted subject to rightful acknowledgement of its source. These policies can be accessed at the following source: <http://researchdata.reading.ac.uk/policies.html>. In addition to this the document incorporates structural aspects pertaining to the Data Curation Lifecycle postulated by Lisa R Johnston in Curating Research Data Volume 2: A Handbook of Current Practice (2017).

11.2	Metadata	20
12.	Retention	20
13.	Withdrawal	21
13.1	Grounds for withdrawal	21
13.2	Grounds for complaint	21
13.3	Complaint procedure.....	21
13.4	Disclaimer	22
14.	Application of this regulation	23
14.1	Commencement	23
14.2	Duration.....	23
14.3	Amendments	23
14.4	Custodian.....	23
14.5	Interpretation.....	23



1. Preamble

Stellenbosch University adopts the following regulation for its institutional research data repository. The regulation itself sets out to govern the operation and use of the **Stellenbosch University Research Data Repository**, known as **SUNScholarData** (the repository) in order to facilitate the archiving of the University's research data associated with published research findings. It is acknowledged that this is an aspirational regulation, and that full implementation may take some time.

2. Definitions

- 2.1 Archiving:** A type of curation activity which ensures that data is properly selected, stored, can be accessed and that its logical and physical integrity is maintained over time, including security and authenticity.²
- 2.2 Authorising Agent:** A member of Stellenbosch University's staff who is authorised to act on behalf of right-holders, and who undertakes to grant or refuse requests for access. Repository administrators are authorising agents by default. Additional or alternative agents may be nominated.
- 2.3 Collection:** A group of one or more datasets. Large or multi-phase projects, or research that has given rise to multiple publications, may be decomposed into several datasets, which can be grouped together as collections. A collection may be stored in the repository or held elsewhere.
- 2.4 Content:** Any information found on the repository including but not limited to research data, datasets, metadata, models, templates, and interactive features.
- 2.5 Curation:** A process of maintaining and adding value to data throughout its lifecycle.³

² Research Data Alliance. 2016. *Term Definition Tool*. Available: <https://smw-rda.esc.rzg.mpg.de/index.php?title=Archiving> [2019, February 18].

³ Research Data Alliance. 2016. *Term Definition Tool*. Available: <https://smw-rda.esc.rzg.mpg.de/index.php?title=Curation> [2019, February 18].

- 2.6 DataCite:** The organisation that manages the registration of digital object identifiers (DOIs), using the DOI system, and guarantees their persistent resolution.
- 2.7 Dataset:** A set or multiple sets of files or a database constituting a discrete and clearly defined set of data and/or the means of generating data related to a research activity or project, or supporting one or more research publications. It may include both data and the means to generate, interpret or validate data, such as computer models and software code. A dataset will include data files (including, where relevant, program files) and documentation. It is possible for a dataset to be stored in the repository, held elsewhere in other online data repositories, or to be published by other means – for example as supplementary materials alongside a journal article.
- 2.8 Data Dissemination:** The publication or public transmission, communication or distribution of data.
- 2.9 Data File:** For the purposes of the repository, any file containing research data or related documentation.
- 2.10 Data Repository:** A searchable and queryable interfacing entity that is able to store, manage, maintain and curate data/digital objects.
- 2.11 Data Sharing:** The peer-to-peer transmission or disclosure of data privately between researchers and other parties. Data sharing is not restricted to, but includes collaborative research.
- 2.12 Depositor:** Any Stellenbosch University user who uploads files into the repository and/or creates a metadata record in the repository for a dataset or collection. A depositor may register a dataset without uploading files, for example where the dataset is held outside of the repository.
- 2.13 Digital Object Identifier:** A character string used to uniquely and persistently identify and resolve to a digital landing page for an object. A DOI will be assigned to every metadata record for a digital dataset held in the repository.
- 2.14 Documentation:** Any digital supplementary context files such as README.txt files, laboratory notebooks, data dictionaries, metadata schemas, standards, administrative materials, codebooks, user manuals, workflows, protocols, technical specifications or methodologies which explain the production, provenance, processing or interpretation of research data.

- 2.15 Ingestion:** The process of receiving research data and preparing it for long-term archival storage and management within the repository.
- 2.16 License Agreement:** A legal statement of the rights that exist in an item of intellectual property and the permissions granted by the right-holders to any users of the item.
- 2.17 Mandatory Minimum Metadata:** The minimum set of metadata for a dataset that must be supplied to DataCite in order for a DOI to be assigned to the dataset metadata record.
- 2.18 Metadata:** Data that plays the role of documentation for data/resource discovery, description/documentation and contextualisation.⁴
- 2.19 Metadata Record:** A structured description of a dataset which enables the identification, discovery, interpretation, use and administration of a dataset or collection. Metadata records can refer to datasets held in the repository or elsewhere.
- 2.20 Open Access Setting:** Access setting for files that are openly accessible to all users without restriction.
- 2.21 Public User:** Anyone who may read and use all publicly-accessible content in the repository.
- 2.22 Private Access Setting:** Access setting for files that are only accessible via individual user accounts or collaborative spaces.
- 2.23 Published Research Findings:** Primary peer-reviewed research publications such as scholarly journal articles, books, chapters and conference papers. For the purposes of this Regulation theses and dissertations are deemed to be published research findings.
- 2.24 Repository Administrator:** A member of the University Library and Information Service staff or other identified University staff member who is authorised to register, deposit, modify, validate, curate, publish and delete data (datasets and metadata records) and perform any other activities necessary for the management and maintenance of the repository.
- 2.25 Research Data:** Recorded information, obtained during the research process, regardless of form or the media on which it may be recorded. The term includes computer software (computer programmes, scripts, databases and documentation thereof), and records of

⁴ Research Data Alliance. 2016. *Term Definition Tool*. Available: <https://smw-rda.esc.rzg.mpg.de/index.php?title=Metadata> [2019, March 7].

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 scientific data include both intangible data (statistics, findings, conclusions) and tangible data. Tangible data include, but are not limited to notes, text, formulae, models, photographs, slides, negatives, films, scans, images, graphs, graphics, audio recordings, 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.

- 2.26 Research Staff:** Stellenbosch University staff members who carry out research. This includes academics, researchers and post-doctoral fellows working in the University's research institutes, bureaus, units, centres and academic departments.
- 2.27 Restricted Access Setting:** Access setting for files that are accessible only to certain members of staff and students with Stellenbosch University accounts.
- 2.28 Stellenbosch University Research Data Repository:** A multidisciplinary online service for the registration, archival storage, sharing, and dissemination of research datasets produced, generated or collected in relation to research conducted under the auspices of Stellenbosch University.
- 2.29 Stellenbosch University User:** A member of Stellenbosch University's staff or student body who is entitled to register and deposit data by accessing the repository with their University account. Stellenbosch University users may be: current Stellenbosch University staff; Stellenbosch University visitors or collaborators associated with a Stellenbosch University research project as well as Stellenbosch University-registered postgraduate students.
- 2.30 Student:** A Stellenbosch University student enrolled for a postgraduate degree or a postgraduate student from an affiliated university who conducts research under the auspices of Stellenbosch University.
- 2.31 User:** Either a public user or a Stellenbosch University user.

3. Purpose

- 3.1 The repository is a multidisciplinary online service for the registration, archival storage, sharing and dissemination of research datasets produced or collected in relation to research

conducted under the auspices of Stellenbosch University. The repository will only publish content if it is associated with published research findings.

- 3.2 The repository can be used by Stellenbosch University academic staff and students to register, share and disseminate data supporting research undertaken at the University in accordance with the University's Research Data Management Regulation and the policies of public research funders.
- 3.3 Subject to 5.3.2, the primary purpose of the repository is to enable discovery and re-use of research data. While access to some data held in the repository may be restricted or subject to embargo for a period, all other data in the repository will be accessible under certain conditions.

4. Service

4.1 Management and administration

- 4.1.1 The repository is managed by Library and Information Service's Division of Research Data Services, and administration is undertaken by repository administrators. There are no costs associated with the use of the service. Enquiries and support requests should be directed to rdm@sun.ac.za.

4.2 Eligible Content

- 4.2.1. Items that fall within the definition of a dataset are eligible for inclusion in the repository.
- 4.2.2. Metadata records and datasets may be registered and deposited in the repository if the data contained in the dataset arise from research undertaken at the University by staff, affiliates or students of the University, and the dataset supports published research.

4.3 Users

Users are divided into five different categories each with a different set of access privileges as indicated in the table below:

Level	User Categories	Privileges
1)	<ul style="list-style-type: none">Repository administrators	<ul style="list-style-type: none">May modify, distribute and delete metadata records and datasets registered and deposited in the repository to ensure the storage and management of content in accordance with the repository's regulation and in compliance with any legal considerations that apply.

Level	User Categories	Privileges
2)	<ul style="list-style-type: none"> • Researchers – Stellenbosch University • Postdoctoral fellows • Doctoral students • Masters students 	<ul style="list-style-type: none"> • Can register and describe datasets and collections in the repository through the creation of metadata records. • May deposit datasets for long-term storage and publication. • Can manage access to data files through the setting of access controls, including specified embargo periods and access conditions. • May access and use metadata records and datasets recorded and held in the repository. Access to some data files may be restricted or may require users to obtain authorisation from an authorising agent.
3)	<ul style="list-style-type: none"> • Researchers – other organisations 	<ul style="list-style-type: none"> • May access collaborative research projects if invited by Stellenbosch University researchers, post-doctoral fellows or doctoral students. Such access is limited to either viewing or commenting on datasets or their associated metadata records. • May access and use metadata records and datasets recorded and held in the repository. Access to some data files may be restricted or may require prior authorisation from an authorising agent.
4)	<ul style="list-style-type: none"> • Professional, administrative and support services • Honours students 	<ul style="list-style-type: none"> • May access and use metadata records and datasets recorded and held in the repository. Access to some data files may be restricted or may require prior authorisation from an authorising agent • Under exceptional circumstances level 4 users may be granted the same privileges as level 2 users. As a prerequisite for this an application demonstrating reasonable justification would need to be submitted to the following e-mail address: rdm@sun.ac.za. Such reasonable justification will exist if such users possess research data which supports published research findings.

Level	User Categories	Privileges
5)	<ul style="list-style-type: none"> Undergraduate students Public users 	<ul style="list-style-type: none"> Users may access and use metadata records and datasets recorded and held in the repository. Access to some data files may be restricted or may require prior authorisation from an authorising agent.

4.4 Service levels

- 4.4.1 The repository will be accessible twenty-four hours a day, seven days a week, with 99% target uptime during any calendar year.
- 4.4.2 Any unscheduled service outages will be managed to minimise impact on the service and to bring the repository back to full operation as soon as possible. In the event of an unscheduled loss of service please contact rdm@sun.ac.za.
- 4.4.3 User support is provided by the Library and Information Service's Research Data Services Division.
- 4.4.4 Support requests and enquiries should be sent to, rdm@sun.ac.za. E-mail requests and enquiries will be acknowledged within two working days. Requests will be prioritised for resolution according to urgency and levels of demand.
- 4.4.5 Consultations or extended support can be provided on request.

4.5 Disclaimer

- 4.5.1 By making use of the repository all users agree to the repository's General Terms and Conditions of Use.
- 4.5.2 Stellenbosch University makes no undertakings as to the accuracy of any dataset or metadata record stored in or supplied by the repository, or their conformity with any relevant laws, and accepts no liability for mistakes, omissions, or legal infringements within any dataset or metadata record.
- 4.5.3 Use of any information supplied by the repository is at the user's own risk. Users should read carefully any information supplied with a dataset which qualifies the accuracy or viability of the data supplied.
- 4.5.4 Stellenbosch University accepts no liability for any errors or losses caused by any inaccuracies in information supplied by the repository, or for the consequences of any person acting on or otherwise relying on such information.

4.5.5 While every care will be taken to maintain the integrity of all datasets, Stellenbosch University is not liable for loss of or damage to any datasets stored in the repository.

5. Submissions

The submission process will take the form of mediated deposits.

Submissions will be published only when they have been validated against minimum requirements and approved by a repository administrator.

Researchers are expected to inspect their data prior to submission in order to detect and fix any errors or issues associated with their data. Such inspection should cover the following aspects:

- a) a review of the content of the data files;
- b) checking for quality and usability issues; and
- c) detection and extraction of any hidden documentation inherent to the data files that may facilitate reuse.

5.1 Eligible Depositors

5.1.1 Only eligible depositors can create metadata records for datasets and collections and deposit associated datasets.

5.1.2 The term eligible depositor refers to one of the following users from Stellenbosch University:

- a) Academics
- b) Researchers
- c) Post-doctoral fellows
- d) Doctoral students

5.1.3 Although Stellenbosch University support staff and honours students are not regarded as eligible depositors they may be deemed to be eligible depositors under certain exceptional circumstances. These circumstances are identified in the service provisions.

5.2 Terms and conditions of deposit

5.2.1 Any Stellenbosch University user who deposits data files into the repository must accept the repository's General Terms and Conditions of Use.

5.2.2 Deposit of data in the repository does not transfer ownership of the data or any rights in the data. All rights in the data are reserved by the rights-holders, including any moral rights under the laws of intellectual property, such as the right to be acknowledged as the creator or source of the dataset.

5.3 Embargoes

- 5.3.1 Files in a dataset can be placed under embargo for a specified period, so that the metadata record and file names are visible, but the file contents cannot be viewed or downloaded.
- 5.3.2 Files should not be embargoed for more than 3 years from the date of deposit, and the embargo should last no longer than the date of publication of any research findings that rely on or reference the dataset, unless there is good reason for continuing to withhold access to data. Examples of these would be the following:
- a) compliance with the terms of a commercial, sponsorship, research, confidentiality or non-disclosure, or any other agreement;
 - b) protection of intellectual property rights pending commercial exploitation or patent registration; or
 - c) in order to enable misconduct enquiries to be carried out.

5.4 Minimum requirements

- 5.4.1 A metadata record must be created for each dataset deposited. The metadata record must be a valid record in accordance with the requirements of the metadata assignment provisions of this regulation. A mandatory minimum amount of metadata must be supplied in order to create a valid metadata record.
- 5.4.2 To be accepted in the repository and published, a dataset must conform to the following requirements:
- a) Permission must be obtained from all right-holders.
 - b) Data files must be virus-free and uncorrupted. It must also be possible to open all the files, given the required software.
 - c) Each data file must have an associated description that indicates its contents.
 - d) Each file must have an appropriate access setting.
 - e) Each dataset must include at least one documentation file. The recommended minimum documentation is a README.txt file providing a brief description of the dataset. A README file template is provided for depositors to use.
 - f) If any data files contain personal data, this must be declared. A repository administrator will ask any depositor submitting personal data to separately provide evidence that the publication of such personal data will not contravene any privacy laws such as the Protection of Personal Information Act 4 of 2013 or the European Union's General Data Protection Regulations.

- g) Each published file must have an associated license statement which specifies the terms according to which the data file(s) can be used.

5.5 Exemptions from DOI allocation

A DOI will not be assigned to a metadata record if the dataset is held outside the repository and is not under the authority of Stellenbosch University.

6. Selection and Appraisal

- 6.1 Research data is not automatically published on the repository. Instead submissions for the publication of research data are made by Stellenbosch University users. These submissions are then subject to a curation process that is managed by repository administrators. During the review process submissions are appraised and then subjected to meta-data enrichment.
- 6.2 After submitting research data for publication on the repository each depositor must complete and submit the repository's data appraisal form.
- 6.3 Data submitted to the repository for publication must be within scope of the eligible contents and of sufficient substance and value to merit inclusion in the repository. Data will be assessed against the repository's data appraisal and selection checklist, and may be rejected if judged to be confidential, ineligible, trivial, of poor quality or without practical re-use value.
- 6.4 Where a policy, legal or contractual requirement to preserve and disseminate Data applies, Data collected or created by members of the University will be eligible for deposit in the repository. Where there is no policy requirement for the data to be preserved and disseminated, or any other reason which may prevent dissemination, such as:
 - a) disclosure restrictions imposed by research contracts,
 - b) prejudicial information that could cause harm if released publicly,
 - c) possibility of copyright infringement,
 - d) patentability of the research data,
 - e) potential for commercialisation,
 - f) ethical considerations, or
 - g) data privacy requirements (in cases where data contains personally identifiable information)

the decision on whether to accept a dataset submitted for deposit will rest with the Manager: Research Data Services. While every effort will be made to accommodate eligible research data in the repository, there is no guarantee that all submitted datasets will be accepted.

- 6.5 Where more suitable external services exist for the dissemination of research data, such as funder-maintained data centres or trustworthy disciplinary data services, users may be directed to offer their datasets to these repositories. For more information, visit the Research Data Management Libguide.
- 6.6 Datasets submitted to the repository will also be assessed against the minimum acceptance criteria for datasets specified in the submission provisions, and may be rejected if they fail to meet these criteria.
- 6.7 Any prospective depositor who is unsure whether a dataset submission will be eligible for admission to the repository should contact rdm@sun.ac.za for advice.
- 6.8 All submissions will be validated against the **minimum requirements** by a repository administrator.
- 6.9 If the submission is deemed out of scope of the repository's eligible contents, it will be rejected.
- 6.10 If the submission falls within the scope of the eligible contents but fails to meet the **minimum requirements** it will be returned to the depositor with a request for the required information or data.
- 6.11 A repository administrator may modify metadata elements to correct minor errors, ensure consistency with the SUNScholarData Regulation, and add administrative metadata, but will not make substantive modifications to descriptive metadata without notifying the depositor first.
- 6.12 The accuracy and authenticity of the content of submissions will not be checked by the repository administrator and are the sole responsibility of the depositor.
- 6.13 Once data has been submitted by a depositor, curation workflows will be utilised in order to accept/reject/edit submissions. The repository's administrators or designated authorising agents will be responsible for reviewing and approving submissions prior to Ingestion.
- 6.14 After a submission has been approved by a repository administrator it will be accepted into the repository. At this stage:

The metadata record will be made publicly available in the repository;

The metadata record for a dataset will be registered in the DataCite Metadata Store under the terms of the particular licensing agreement(s) selected and a DOI will be assigned to and published on the metadata record in the repository.

7. Processing and Treatment⁵

- 7.1 Whenever repository administrators make changes or additions to submissions they must create working copies of the files in question in order to protect the originals.
- 7.2 A curation log will be used in order to track any changes made to the submitted data. The curation log will keep a record of the correspondence between the repository and data submitters.
- 7.3 Repository administrators may inspect the file representation and organisation with the intention of accomplishing the following two objectives: understanding the directory structures, file relationships and any naming conventions; and preserving any relationships in the files or generating documentation to help others understand how the files relate to one another.
- 7.4 Repository administrators may inspect submitted data in order to detect and fix any errors or issues associated with the data. Such inspection would cover the following aspects:
 - a) a review of the content of the data files;
 - b) checking for quality and usability issues; and
 - c) detection and extraction of any hidden documentation inherent to the data files that may facilitate reuse.
- 7.5 Repository administrators may work with data creators in order to enhance data submissions.
- 7.6 Repository administrators may assess the formats of submitted files in order to:
 - a) Verify technical metadata that would optimise the files for reuse;
 - b) Transform files into open, nonproprietary file formats;
 - c) Retain original files if data transformation is not perfect.
- 7.7 Repository administrators may rename files to optimise their description and organise them in a manner which might facilitate their reuse.

⁵ Johnston, L.R. 2017. *Curating Research Data Volume 2: A Handbook of Current Practice*. Chicago: Association of College and Research Libraries.

8. Ingestion and Storage

8.1 Ingestion⁶

8.1.1 After they have been processed and treated data files will be ingested.

8.1.2 Throughout the ingestion process the integrity and fixity of the data files will be maintained and verified.

8.1.3 Some changes to files may have to be made post-ingestion. In such instances a separate notification will be sent to the depositor noting this process.

8.2 Security of Data⁷

8.2.1 The repository will be managed in accordance with Information Security Standard ISO27001. The repository holds two copies of deposited data in separate data centres.

8.2.2 When data are first deposited in the repository they will be automatically transferred to secure storage. The depositor should retain their copy of the data during this transit phase. Once the two copies of the data have been deposited into the long-term storage data centres the relevant depositor will be notified of the successful ingestion of the data.

8.3 Logical Security

8.3.1 Host systems, operating systems, and externally sourced applications software must have the most recently available and appropriate software security patches within the relevant software/OS distribution, commensurate with acceptable risk.

8.3.2 Applications must be designed, implemented, and managed to minimise the risk from malicious or accidental misuse of the application and any associated data.

8.3.3 Adequate authentication and authorisation functions must be provided, commensurate with appropriate use and the acceptable level of risk. In general, a least privileges model will be followed, in terms of which privileged users of the system will be granted the least amount of privilege necessary to fulfil their responsibilities.

⁶ Ibid.

⁷ University of California. 2016. *California Digital Library Information Technology Key Security Guidelines & Baseline Supporting Practices*. Available: https://www.cdlib.org/about/it_sec_guidelines.html [2018, September 13].

8.3.4 Adequate electronic perimeter controls must be in place to control and manage the origins, targets, and allowable connections to the repository’s server(s) and desktop systems.

8.3.5 System monitoring must be designed to detect configuration changes and intrusions commensurate with acceptable risk. System logs must be kept and examined as required to detect unusual behavior and conduct forensics for incident analysis and reporting (following a principle of least privilege necessary for such detection and analysis).

8.4 Privacy and Confidentiality

Applications must be designed and systems must be used so as to protect the privacy and confidentiality of the various types of electronic data they process, in accordance with Stellenbosch University’s [Privacy Regulation](#) as well as applicable laws such as the [Protection of Personal Protection Information Act](#) and the [General Data Protection Regulations](#).

8.5 Storage Quota

8.5.1 The storage quota is split between individual storage and research project storage. Whereas the individual storage is assigned to eligible depositors the research project storage is allocated to research projects. The project storage has been assigned to the University’s research institutes, bureaus, units, centres and academic departments on a pro rata basis.

The table below illustrates the manner in which the storage quota has been partitioned between the user and research project quotas:

Storage quota	User category	Allocated storage (Gigabytes) per User
Individual storage	Academics or Researchers	20
	Post-doctoral fellows	20
	Doctoral and Masters students	5
Research project storage	Academic department, research institute, bureau, unit or centre	5

Notwithstanding the existence of the allocated storage quotas, Stellenbosch University users can still apply for the allocation of more storage. The success of these applications is not guaranteed. Each application will be assessed by the repository’s administrators on a case-by-case basis.

8.6 Compliance with this Regulation (in terms of ingestion and storage)

The following activities are specifically prohibited under this regulation:

- a) Interfering with, tampering with, or disrupting resources.
- b) Intentionally transmitting any computer viruses, worms, or other malicious software.
- c) Attempting to access, accessing, or exploiting resources without the prerequisite authorisation.
- d) Knowingly enabling inappropriate levels of access or exploitation of resources by others.
- e) Uploading sensitive or confidential electronic information/data to computers that are not adequately configured to protect it from unauthorised access.
- f) Disclosing any electronic information/data without possessing the right to make such disclosure.
- g) In addition to any possible legal sanctions, violators of this regulation may be subject to disciplinary action up to and including dismissal or expulsion, pursuant to Stellenbosch University's policies, collective bargaining agreements, codes of conduct, or other instruments governing the individual's relationship with the University. Recourse to such actions shall be as provided for under the provisions of those instruments.

9. Metadata Assignment

9.1 Metadata Records

- 9.1.1 A metadata record will exist for each dataset and collection.
- 9.1.2 Each metadata record will be displayed on its own landing page at a unique URL. Each landing page for a digital dataset held in the repository or a non-digital dataset stored and managed at Stellenbosch University will be assigned a DOI, which will resolve to the metadata record URL. DOIs are not assigned to metadata records for collections or for datasets that are not under the authority of the University.
- 9.1.3 A mandatory minimum set of metadata elements must be recorded for each metadata record. Other, optional metadata may be provided at the discretion of the depositor or a repository administrator.

9.1.4 All metadata records registered with the repository are openly available online and can be accessed by anyone without restriction and free of charge.

9.2 Mandatory metadata

9.2.1 The mandatory minimum metadata that must be recorded for each dataset or collection metadata record are:

- Title
- Author(s)
- Categories
- Keywords
- Item type
- Description
- License
- Academic group
- URL/DOI (for Datasets held in external repositories)

9.2.2 The depositor must provide the mandatory metadata. For datasets held in the repository, the data publisher and DOI will be assigned in the process of publishing the metadata record.

9.2.3 In addition to the mandatory minimum metadata elements, other metadata elements may be entered in order for the metadata record to be accepted as a valid record. In this respect provision can be made for the inclusion of domain-specific metadata schemas.

9.3 Language

9.3.1 The primary language of the metadata record is English.

9.3.2 The mandatory metadata fields must be completed in English.

9.3.3 A title in an alternative language may be entered in the alternative title field, and an alternative-language description of the dataset can be entered alongside the English-language version in the description field.

9.3.4 Any other free text fields that will be displayed on the public metadata record must also have English language entries, although alternative-language entries may be included where relevant.

10. Access

- 10.1 Access to files within a dataset may be controlled by depositors. Files deposited with the repository may have the following access settings:
- a) Open Access Setting: Openly accessible without restriction.
 - b) Private Access Setting: Access to individual user accounts or collaborative spaces. This setting is only accessible to members of staff and students with a Stellenbosch University account 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) 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.
- 10.2 Open Data Files are made available under open licenses. The default recommended licenses for data files are the [Creative Commons Attribution and Open Data Commons Attribution](#) licenses.
- 10.3 Open software files are made available under one of a variety of standard [Open Source software licenses](#).
- 10.4 Data can be held under embargo for up to 36 months after the date of deposit. After this date data files must be made available, unless there is good reason to continue withholding access to the data.
- 10.5 The level of access is specified by the depositor when data files are uploaded to the repository.
- 10.6 All metadata records registered with the repository are openly available online and can be accessed by anyone without restriction and free of charge.

11. Reuse

11.1 Openly Accessible Datasets

- 11.1.1 Data and software files made available under the terms of a Creative Commons Attribution, Open Data Commons Attribution or standard Open Source license can generally be:

- reproduced, displayed or performed, given to third parties, and stored in a database in any format or medium;
- used for personal research or study, educational, non-commercial purposes without prior permission or charge;

provided that:

- a citation to the dataset is provided, including at least the following metadata elements: the Author(s), Title, Data Publisher, Publication Year, and DOI of the Metadata Record;
- the original copyright statement and license/rights permission statement are given.

11.2 Metadata

11.2.1 The metadata published in the metadata records may be re-used in any medium without prior permission for not-for-profit purposes.

11.2.2 Any person who re-uses metadata is asked to identify the Stellenbosch University Research Data Repository as the source of the metadata and/or to provide a reference to the original metadata record by means of its DOI or URL.

12. Retention

12.1 Once deposited, a dataset can still be updated by the addition, removal or modification of data files. Data files are allocated a checksum to facilitate the detection of alterations. A new version of a dataset may be deposited. In this case:

- The earlier version may be withdrawn from public view;
- There will be links between earlier and later versions of a dataset, with the most recent version clearly identified.

12.2 Datasets deposited in the repository will be retained for a minimum of ten years from the date of deposit or for the retention period specified in any applicable third party policy or contract term, whichever is longer. Exceptions to this policy will be made only if the repository is required to withdraw data in accordance with any applicable laws or policies or in response to a proven complaint.

- 12.3 Once the minimum retention period has expired, a dataset will be subject to review by repository administrators, and may be withdrawn after the depositor or relevant Head of Department has been consulted.
- 12.4 Metadata records will be retained indefinitely, including those for withdrawn datasets. Where a dataset or collection has been withdrawn from the repository, the associated metadata record will not be deleted. Instead the metadata record will be annotated to record the change in status of the dataset or collection and the reason for this change.

13. Withdrawal

13.1 Grounds for withdrawal

Materials deposited in the repository are made available based on the terms specified by the depositor of the dataset, with the implied permission of all right-holders, and will not be withdrawn from the repository except under the following circumstances:

- a) in keeping with one of the University's policies or regulations;
- b) upon the expiry of the minimum or specified retention period following standard collection management procedures;
- c) in response to an identified breach of contract, law or policy; or
- d) on receipt of a valid and proven complaint.

13.2 Grounds for complaint

Acceptable reasons for the exceptional withdrawal of data include:

- a) violation of intellectual property rights;
- b) breach of contract;
- c) legislative breach;
- d) matters of national security; and
- e) falsification or misrepresentation of data, plagiarism or failure to follow ethical guidelines.

13.3 Complaint procedure

13.3.1 Anyone who wishes to lodge a complaint on any of the grounds specified above must contact the repository in writing at rdm@sun.ac.za with the following information:

- The name and contact details of the complainant;

- A reference to the dataset in question, including a DOI or URL for the relevant metadata record;
- Details of the complaint, including any evidence or proof.

13.3.2 A repository administrator will acknowledge receipt of the complaint in writing, and the dataset in question will be immediately removed from public view pending further investigation. The depositor of the dataset as well as any right-holders will be notified about the complaint, and the matter will be investigated by the Manager: Research Data Services and the University's legal officers. The Manager: Research Data Services will notify the complainant and the depositor of the outcome of the investigation and of any action taken as a result.

13.3.3 If the complaint is found to be invalid, the dataset will be re-instated.

13.3.4 If it is found that there are reasonable grounds for a complaint, or that there is proof of infringement, the dataset in question will be permanently withdrawn from the repository.

13.3.5 Withdrawn datasets will not be deleted, but merely removed from public view. The metadata record for a withdrawn item will be retained indefinitely, and modified to include:

- a link to a replacement version, where available; or
- a note explaining the reasons for the withdrawal in question.

13.4 Disclaimer

13.4.1 Stellenbosch University takes reasonable steps to ensure that any datasets and associated metadata records published in its repository are an accurate and authentic output of research activity undertaken at the University. Furthermore, the University takes reasonable steps to ensure that datasets published in its repository do not breach any law, and that they do not violate or infringe any copyright, trademark, patent or other intellectual property rights whatsoever of any person(s) or organisation.

13.4.2 Notwithstanding the above, to the extent permitted by law, the repository shall not be held liable for content or material which infringes the rights of any third party. The repository seeks to retain the good will and favour of rights-holders. To this end every effort will be made to accommodate requests and negotiate compromises. This may include the replacement of an offending dataset with a modified version that is acceptable to the complainant, amendment

of the terms of access to a dataset or the licensing agreement(s) associated with its data files or upholding an embargo period before making a dataset available.

14. Application of this regulation

14.1 Commencement

The commencement date of this regulation has yet to be confirmed.

14.2 Duration

This regulation will be applicable until it has been amended or repealed.

14.3 Amendments

Notices of any amendments to this guideline will be publicised. Such amendments will come into effect once they have been published, on the Library and Information Service's website, as part of the regulation.

14.4 Custodian

The custodian of this regulation is the Library and Information Service.

14.5 Interpretation

The rules embodied in this regulation should be read with or interpreted within the broader context of the University's governance or regulatory framework. More specifically, the regulation should be read with the regulatory documents set out in the list below:

1. Information Governance Policy
2. [Mandatory Self-Archiving of Research Output](#)
3. [Privacy Regulation](#)
4. Research Data Management Regulation
5. [IT Policy Definitions](#)
6. [Information Security Regulations](#)
7. [Identity and Access Management Policy](#)
8. [Password Regulations](#)
9. [Audit Trail Logging and Monitoring Regulation](#)
10. [Electronic Communication Policy](#)
11. [Electronic Identity Validation Regulation](#)
12. [Records Management Policy](#)
13. [Super User Policy](#)
14. [Research Policy](#)
15. [Policy for Responsible Research Conduct at Stellenbosch University](#)
16. [Policy on Plagiarism](#)

17. [Stellenbosch University Procedure for the investigation and management of allegations of plagiarism](#)
18. [Postdoctoral Policy](#)
19. [Contract Research Policy](#)
20. [Policy on Conflict of Interest](#)
21. [Stellenbosch University's \(SU\) procedure for the investigation of allegations of breach of research norms and standards](#)
22. [Policy regarding the Research Committees and Subcommittees](#)
23. [Health Research Ethics Committees Human Research: Standard Operating Procedures and Guidelines](#)
24. [Research Ethics Committee: Human Research \(Humanities\) Standard Operating Procedure](#)
25. [Research Ethics Committee: Animal Care and Use Standard Operating Procedures and Guidelines](#)
26. [Research Ethics Committee: Biosafety and Environmental Ethics Standard Operating Procedures and Guidelines](#)
27. [Stellenbosch University Policy in Respect of Exploitation of Intellectual Property](#)
28. [Risk Management Policy](#)

SUNScholarData Regulation

Accepted and Approved by Senior Management Team, Library and Information Service
on 15 October 2019

Owner: Stellenbosch University Library and Information Service