



Policy

SAEON Data Policy

Context	The SAEON Data Policy addresses the terms and conditions under which SAEON data is made available to external users, and the terms and conditions that will apply to data hosted or archived on behalf of data providers.
Topics	Data Archiving, Data Dissemination, Data User Requirements
Intended Audience	Internal and external data providers
Stakeholders	SAEON staff, collaborators, data providers, students, and external data users

Revision History

Date	Authors	Notes
27 Jul 2020	L Chiloane	Draft for Comment
04 Aug 2020	L Chiloane, N Allsopp, J Slingsby, A Swemmer	Revision

License

Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)	Free to share, adapt, and apply the work, even for commercial purposes, provided that you give appropriate credit and distribute the adaptations under the same license.
http://creativecommons.org/licenses/by-sa/4.0/	

Table of Contents

Revision History	1
License	1
Table of Contents	2
Definitions	3
SAEON Data Policy	5
A. Conditions of Archiving	5
B. Responsibilities of the Data Provider	6
Metadata	6
Intellectual property rights	6
Persistent identifiers and data publication	7
Review of embargo/Annual review	7
Regular updates	7
C. Responsibilities of the SAEON ODP Gatekeeper	7
Basic services	8
Paid services	8
D. SAEON Data User Requirements	8
Annexures	9
Annexure I: Protection of Personal Information Act Procedures	9
Annexure II: Metadata Requirements and Guidance	11

Definitions

Archiving	Long-term preservation of data and information in a secure environment. Archived data are not publicly available.
Collaborators	Institutions, partners, students and staff with whom the Data Provider collaborates, and with whom access must be granted, in order to fulfill the Data Provider's intended use of the dataset. Collaborators must be specified in the project Data Management Plan.
Data	Any digital research output that can be processed as data - including text, numeric, image, video, and binary formats.
Data Provider	The individual or institution that wishes to archive, publish and share one or more datasets in the SAEON Open Data Platform. The Data Provider must demonstrate the right to share the data and to determine an appropriate license. Data Providers may be SAEON staff or non-staff.
Data User	Individual to whom access has been granted to a dataset, including their Collaborators.
Dataset	Any digital object or collection of digital objects that individually or collectively meet a set of granularity requirements. These requirements typically regard authorship, topics, spatial and temporal coverage, and data formats as a basis for separation or grouping of digital objects.
Embargo	A period during which open access to data is not allowed. Until the data are made open, permission to use these data will be given directly by the Data Provider.
External Data	Historical data or data that have been produced with no affiliations to SAEON, be it through registered projects, funding or resources hosted on behalf of data providers.
Dissemination/ Publication	The distribution of data and metadata in a form that is freely available to the public.
Gatekeeper	Individual or institution responsible for archiving and dissemination facilities.
Hosting	Storage of data on a stable and accessible web platform.
Intellectual Property	Any creation of the mind; where such creation may be embodied in, for example, know-how, software, trade marks, designs, names and insignia, copyright, patentable inventions as well as confidential matters and trade secrets; and may be capable of being protected by law from use by any other

person, whether in terms of South African law or foreign intellectual property law, and includes any rights in such creation.

License	The rights claimed by the Data Provider on behalf of the author/ creator of the dataset, and the optional limitations placed on application or use of the dataset. SAEON advocates the use of Creative Commons licenses, but also allows custom licenses. See SAEON's license guidance for more information.
Metadata	Descriptive information about one or more aspects of datasets; it is used to summarise basic information about data which can make tracking and working with specific data easier.
Open Data	Data that are made available to the public without restrictions or charges.
SAEON Affiliated Data	Data produced by SAEON, by projects registered with a SAEON Node or the SAEON National Office or with SAEON funding and resources.
SAEON Open Data Platform (ODP)	An information system developed by SAEON which provides an environment that allows aggregation of metadata from a variety of Data Providers. The metadata is used to describe datasets, and assists with discovery, re-use, and citation. Data Providers can optionally host datasets on SAEON Open Data Platform infrastructure.
Services	The SAEON ODP provides a portfolio of services, including metadata upload, editing, and harvesting, data upload, and data download / visualisation, and exploration services. Support services include download statistics and feedback mechanisms for end users.
Sensitive Data	Data that will not be released due to copyright agreements, local or international law and environmentally sensitive content. Some examples of sensitive data restrictions may include: locations of rare and endangered species, archaeological sites, data that are covered under prior licensing or copyright, or data where the protection of individual privacy is involved. In the interests of environmental responsibility SAEON reserves the right to restrict data as sensitive, even if the Data Provider does not stipulate it.
Terms and Conditions	Datasets may be subject to individual terms and conditions of use, provided as a non-standardised license. Selected open licenses (Creative Commons) also require the end user to abide by the Terms and Conditions.

SAEON Data Policy

This data policy outlines the roles and responsibilities of SAEON and collaborating Data Providers, with the aim of making datasets submitted by the Data Providers discoverable, accessible, interoperable, and re-usable. This supports open science, and improves the utility of publicly funded research outputs for society.

SAEON promotes the responsible archiving and broadest possible dissemination and reuse of all research data while respecting the legal rights of the Data User and Data Provider.

This document contains four sections:

- A. Conditions of Archiving
- B. Responsibilities of the Data Provider
- C. Responsibilities of the SAEON ODP Gatekeeper
- D. SAEON Data User Requirements

A. Conditions of Archiving

1. The SAEON Open Data Platform (ODP) offers the ability to store metadata and datasets in a secure environment that protects the intellectual property rights of the Data Providers. All Data Providers are required to sign a Data Agreement or complete a Data Management Plan.
2. The Data Provider can archive their dataset(s) with the SAEON Open Data Platform, without payment of a monetary fee. SAEON agrees to archive these datasets provided they meet the conditions in (3), (4), (5), (6), and (18) below.

OR: The Data Provider can host the data in non-SAEON infrastructure of choice, provided access is aligned to the license conditions, and conditions (3), (4), and (6) are adhered to.

OR: The Data Provider can host larger datasets or collections in SAEON infrastructure on the basis of a service level agreement, for which a cost of service fee may apply.

3. The Data Provider must confirm by way of a written dataset submission that they are the creator/author and owner of the metadata and data that they upload to the ODP and that they have the right to specify the license that will be applied to it.
4. The Data Provider is aware that SAEON will offer applicable metadata records to global networks, including but not limited to [DataCite](#), [GEOSS](#), and [ILTER](#). Metadata is automatically made available to data systems linked to the SAEON ODP.
5. The Data Provider grants SAEON the permission to copy, store and transform the uploaded data in order to
 - publish data in standardised formats for access via services within license constraints,
 - perform format migrations as part of the preservation process.

B. Responsibilities of the Data Provider

Metadata

6. All datasets that are archived in the SAEON ODP must be accompanied by metadata. All metadata provided to SAEON will be made publicly accessible, to allow potential users to locate datasets of interest to them. Metadata must be maintained using the service stipulated by (23) or (24) below.

Intellectual property rights

7. All intellectual property predating the submission of data shall be the sole property of the party that discovered and/or created and/or acquired such intellectual property (Background Intellectual Property). The owner and/or creator and/or holder of the background intellectual property must determine an appropriate license in consultation with SAEON.
8. SAEON allows access by default to datasets by application of a Creative Commons License (see <https://creativecommons.org/choose/>).
9. The standard license choice allows sharing, as long as subsequent users can share alike, and allows commercial application. The terms of this license can be found here: <http://creativecommons.org/licenses/by-sa/4.0/>. This license protects the rights of IP owners across multiple jurisdictions.
10. If the recommended license as per point 9 is deemed not appropriate by the Data Provider, they should indicate either:
 - a. If another Creative Commons License is acceptable based on
 - i. Whether use is limited to non-commercial applications;
 - ii. Whether a publication embargo is applicable;
 - iii. Whether the data refers to species that are endangered;
 - iv. Whether the data is private to an individual;
 - v. Whether there are ethical considerations attached to the data;
 - vi. Whether end users should be identifiable.
 - b. If no Creative Commons License is applicable, in which case they must
 - i. Provide an end user license (terms and conditions) instead; or
 - ii. Disallow access to the data.
11. SAEON Licensing options can be found here: <https://goo.gl/HaJiDb>
12. The responsibility falls on the Data Provider/ IP Owner to choose the appropriate license for their data, taking into account legal and ethical considerations. SAEON will maintain the license

of the dataset as received, and any added value products will be made available under the same license.

Persistent identifiers and data publication

13. Digital Object Identifiers should be issued for published datasets as the preferred means of tracking the data. This will normally be done by default for all publicly available data under the Creative Commons License, and allows citation of data in journals and scholarly articles.
14. If, for any reason, the Data Provider does not want a Digital Object Identifier issued for the data, they must provide an alternative method for tracking updates to their data.

Review of embargo/Annual review

15. The Data Provider may stipulate an embargo period of a maximum period of three years, starting from the date the data is provided.
16. The data will be made publicly available at the end of the embargo period unless an extension is requested.
17. An extension to the embargo period will be granted on an annual basis, provided that it is requested before the termination date of the current embargo.

Regular updates

18. The Data Provider will notify the SAEON Open Data Platform Gatekeeper if they have any updates to add to a particular dataset archived in the SAEON Data Portal. Such updates include changes of contact details; changes in or additions to metadata information; changes in or additions to the dataset.

C. Responsibilities of the SAEON ODP Gatekeeper

19. SAEON will comply with the Protection of Personal Information (POPI) Act with regards to the personal information provided by Data Providers when registering on SAEON's online platforms. Personal information included in metadata records intended for publication are not subject to the POPI Act and once published are not solely under the control of SAEON. See Annexure I for details.
20. SAEON does not make any express or implied commitment in terms of service levels in support of the data deposit.

OR: SAEON will adhere to the service level agreement entered into with the data provider for a paid service portfolio.
21. SAEON will apply due diligence in respect of backups of the data and metadata supplied by the Data Provider. Backup frequency is routinely on a daily basis but at least once a week.

OR: SAEON will develop a backup and failover strategy with the Data Provider as a part of a paid service portfolio.

22. SAEON will endeavour to provide high availability for its servers and services associated with metadata and data deposit, but does not make any express warranty in respect of this.

Basic services

23. The SAEON Data Policy makes provision for the following specific services:
- a. SAEON will create a custodian area for the Data Provider, for the purpose of management and harvesting of metadata.

Paid services

24. A paid service portfolio can stipulate any or all of the following requirements, in addition to the above, as appended to the service level agreement:
- SAEON will create a custom portal or gateway for the service.
 - SAEON will provide operational database access and support for the purpose of publication of a variety of supported data formats.
 - SAEON will define specific development tasks based on user requirements in respect of gateway and portal functionality.

D. SAEON Data User Requirements

25. In order to download SAEON data, Data Users must accept an agreement to properly cite those data (e.g., text will be presented which must be accepted before the download initiates). In addition, the acceptance of the following Acknowledgment statement will be required:

These data are made available with the express understanding that any such use will properly acknowledge the originator(s) and publisher and cite the associated Digital Object Identifiers (DOIs). Anyone wishing to use these data should properly cite and attribute the data providers listed as authors in the metadata provided with each dataset. It is expected that all the conditions of the data license will be strictly honoured. Use of any material herein should be properly cited using the dataset's DOIs. SAEON cannot be held responsible for the quality of data provided by third parties, and while we take reasonable care in referencing these datasets, the content of both metadata and data is under control of the third-party provider.

Annexures

Annexure I: Protection of Personal Information Act Procedures

#	Aspect	Description and Procedure	Reference
EAP-01-01-15	Information Officer	SAEON must designate a responsible Information Officer to implement the provisions of the act. <i>This will be the node manager or CDIO ex officio.</i>	POPI
EAP-01-01-01	Consent	Data can only be processed if consent is given. <i>Data Providers will be asked for their information when registering on the SAEON online platforms.</i>	POPI
EAP-01-01-02	Withdrawal of Consent	Consent can be withdrawn at any time. <i>If a Data Provider requests that we remove their registration information from our database the Information Officer will do so in a timely manner.</i>	POPI
EAP-01-01-03	Obtain Directly	Obtain directly from the subject, as attested by email validation. <i>The Information Officer will obtain direct consent from Data Providers to collect or use their personal information in any way that falls outside of the uses related to the online platform registration.</i>	POPI
EAP-01-01-04	Defined Purpose	Purpose of data collection must be defined explicitly and the provider must be aware of this. <i>The purpose of the data collection, i.e. for registration on the online platforms, will be made explicit in the registration process.</i>	POPI
EAP-01-01-05	Retention	Retention has to be qualified and the provider has to assent to use of inactive records for statistics and reporting. <i>Information will be kept in the databases related to those online platforms until the Data Provider deregisters from the platform.</i>	POPI
EAP-01-01-06	Proof of Removal	Proof has to be provided of removal of records for whatever reason – by request form data provider or through lapse of registration period. It must not be possible to reconstruct the record. <i>The Information Officer will send an email as proof of removal should Data Providers request that their information be removed from the database.</i>	POPI
EAP-01-01-07	No other uses	Specific conditions need to be met for the use of the information in a different context. <i>The registration information will be used solely to allow the Data Providers access to the platform and for any platform</i>	POPI

		<i>related communications. Consent will be obtained by the Information Officer should any other uses of the information be required.</i>	
EAP-01-01-08	Quality	Reasonable care must be exercised to ensure that the data is complete and accurate. <i>Data Providers are responsible for ensuring that their registration information is up to date. The Information Officer will ensure that the Data Providers are able to easily update their information on the platforms.</i>	POPI
EAP-01-01-09	Documentation	Processing history and documentation must be maintained. <i>All email correspondence relating to the collection of personal information for the registration process will be kept by the Information Officer.</i>	PAIA
EAP-01-01-10	Notification	The data subject/ provider must be provided with information about the system/ responsible party. This should form part of the contracting between the parties. <i>The Information Officer will ensure that all necessary information is included on the online platforms.</i>	POPI
EAP-01-01-11	Security	Prevent loss of or damage to personal information, and prevent unlawful access to such information. <i>The Data Providers' information collected for registration on the platforms will be stored in the SAEON uLwazi online databases.</i>	POPI
EAP-01-01-12	Delegation	Subcontractors, employees, and operators with access to the data are bound by the same provisions. <i>Employees with access to the data will be made aware of the provisions surrounding it by the Information Officer.</i>	POPI
EAP-01-01-13	Notification of Breach	Where there are reasonable grounds to believe that personal information has been compromised, the responsible party must communicate this to the Regulator and the data subject in the prescribed way and within the prescribed time period. <i>The Information Manager will contact the Regulator and Data Provider should the data be compromised.</i>	POPI
EAP-01-01-14	Notification Content	The content must enable the data subject to understand the consequences and take action to take protective measures, and describe corrective steps taken by the responsible party. <i>This will be taken into account by the Information Officer when writing the notification.</i>	POPI

Annexure II: Metadata Requirements and Guidance

Metadata documenting archived datasets of all data types will be made publicly available regardless of any restrictions on access to the data. All metadata will follow SAEON recommended standards and, at a minimum, will contain adequate information on proper citation, access, contact information, and discovery. Complete information including methods, structure, semantics, and quality control/assurance is expected for most datasets and is strongly encouraged. Guidance in this respect can be accessed below:

Data Family	Metadata Standards	Guidance Documents
Spatial Data	<u>ISO 19115</u> <u>SANS 1878</u>	<u>https://bit.ly/303Nrb6</u>
Multidimensional Data	ISO 19115 SANS1878	<u>https://bit.ly/2vKKOgB</u>
Ecosystem Observation and Biodiversity Data	<u>EML</u> <u>Darwin Core</u>	<u>https://bit.ly/302FJOv</u>
Digital Objects	<u>DataCite</u> DublinCore	<u>http://bit.ly/2OssKEc</u>