

International Science Council

ISSN 1998-0337

SCARreport

No 42
September 2022

SCAR Data Policy (2022)



Published by the
SCIENTIFIC COMMITTEE ON ANTARCTIC RESEARCH
at the
Scott Polar Research Institute, Cambridge, United Kingdom

Table of Contents

1. Definitions	2
2. Background.....	2
3. Purpose.....	3
4. Resourcing	3
5. Policy Statement.....	4
5.1 <i>Data must be ethically open</i>	4
5.2 <i>Data should be free</i>	4
5.3 <i>Data must be provided in a timely manner</i>	4
5.4 <i>FAIR Principles should be applied to the greatest extent practicable</i>	4
5.5 <i>All data must be accompanied by a complete set of metadata</i>	4
5.6 <i>Data should have persistent and globally unique identifiers</i>	5
5.7 <i>Data must be labelled as reusable</i>	5
5.8 <i>Data sources should be attributable and attributed</i>	5
5.9 <i>Data must be appropriately preserved for the long term</i>	5
5.10 <i>Data management and long-term curation must be planned and resourced</i> ..	5
5.11 <i>Quality assurance and fitness-for-purpose</i>	6
5.12 <i>Reporting</i>	6
6. Related Legislation and Documents	6
7. Feedback	6
8. Approval and Review Details.....	7
 Appendix – List of Acronyms	 8

SCAR Data Policy (2022)

1. Definitions

Data: A set of values, symbols, or signs (recorded on any type of medium) that describe one or more properties of an entity. For example, the numbers generated by a sensor, values derived from a model or analysis, text entered into a survey, audio or video files, or the raw text of a document. Generally speaking, data are used to quantitatively or qualitatively describe one or more persons or objects. Research data provide the evidence base for supporting or refuting ideas in a scientific manner.

SCAR Data: SCAR data are those data generated under the auspices of a SCAR-sponsored research project. Data used in SCAR projects that are not SCAR-generated, but are managed by National Antarctic Data Centres, should be treated similarly to SCAR data to the greatest extent practicable, subject to ownership, licensing, and operational requirements.

Data Management Plan: A document describing how an individual collection of data or linked data collections (e.g. under a single project or programme) will be managed, described, and stored, the standards the data conform to, and how data will be handled and protected during and after the completion of the project.

Information: Products derived from data that lead to a greater understanding of an entity. For example, the interpretation of a range of data from an array of conductivity sensors that informs us about the ocean's salinity range, or the narrative text of a report on algal blooms that informs the reader regarding their timing.

Metadata: Metadata provide information that describe the data source, and the time, place, and conditions under which the data were created. Metadata inform the user of who, when, what, where, why, and how data were generated. Metadata allow the data to be traced to a known origin and known quality. Metadata can be used for discovery and identification of data collections, to provide information on structural aspects of the data, and to provide administrative information on aspects such as ownership and licensing.

2. Background

The International Science Council (ISC, formerly ICSU and ISSC) is the parent body of the Scientific Committee on Antarctic Research (SCAR). In its Assessment on Scientific Data and Information in 2004, ISC observed that "science has long been best served by a system of minimal constraints on the availability of data and information", and that a strong public domain for scientific data and information promotes greater return from investment in research, stimulates innovation and enables more informed decision-making. Thus, one of the fundamental recommendations of the assessment is that "ICSU should continue to actively promote the principle of full and open access to scientific data".

Additionally, countries working in the Antarctic operate within the framework of the Antarctic Treaty System. The cornerstone of the system is the Antarctic Treaty, which was signed on December 1, 1959, and came into effect on June 23, 1961. Of particular relevance for polar data management and delivery is Article III, section

1(c), which stipulates that “scientific observations and results from Antarctica shall be exchanged and made freely available”. This Article has been followed up by ATCM Resolutions, such as:

- ATCM Recommendation XIII-5 (1985), which invites SCAR to offer advice “on steps that possibly could be taken to improve the comparability and accessibility of scientific data on Antarctica.”
- ATCM XXII Resolution 4 (1998), which recommends that Consultative Parties establish National Antarctic Data Centres and link these to the Antarctic Data Directory, and that they give priority consideration as to how the requirement for freedom of access to scientific information is achieved within their national data management systems.

Effective data management is a critical part of operating responsibly within the Antarctic environment, and it is essential that SCAR is able to provide clear guidance to researchers and research programmes on data management requirements.

In 2019, the Standing Committee on Antarctic Data Management (SCADM) initiated a process to align the data policies of international scientific bodies coordinating research in the polar regions. Following this, a working group under SCADM, SOOS, IASC, SAON, and the Arctic SDI published the report “Alignment of Polar Data Policies - Recommended Principles” in November 2021 (DOI: [10.5281/zenodo.5734900](https://doi.org/10.5281/zenodo.5734900)). The report examines external policy drivers and overarching global and regional data policies, notably those of the Antarctic Treaty, ISC, UNESCO, WMO, IOC, UN-GGIM, OECD, and GEO, and concludes by recommending and explaining ten fundamental principles for polar data policies. This data policy is based on those ten principles.

SCADM is responsible for coordinating the activities of National Antarctic Data Centres (NADCs), and SCAR activities relating to the curation and delivery of data.

Matters relating specifically to cartographic information and mapping – for example relating to maps, mapping standards or infrastructure for maintaining geolocation information – are referred to the Standing Committee on Antarctic Geographic Information (SCAGI).

3. Purpose

The purpose of this document is to identify roles and requirements relating to the management of data and information collected as part of SCAR initiatives, and to provide guidance regarding best practice for managing and delivering Antarctic and Southern Ocean data.

4. Resourcing

It is essential that data be effectively curated to ensure their long-term integrity, to ensure their accessibility and reusability, and to maximise their value for a broad range of uses. To support these enduring objectives, it is important that data management activities be supported over the long-term, beyond the lifespan of individual projects. This will include costs for hardware and software, as well as supporting staff with specialist skills in data preservation, data curation, as well as the necessary infrastructure to support the effective delivery of data services.

5. Policy Statement

5.1 *Data must be ethically open*

SCAR data will be made accessible in a full, free, and open manner for all users in keeping with Section III.1.c. of the Antarctic Treaty, excepting cases where data must be limited for ethical, cultural, or legal reasons. Data should be “as open as possible, as closed as necessary”.

5.2 *Data should be free*

Distribution and reuse of SCAR data should be free of charge, delivered at no more than the cost of reproduction and delivery. With modern digital communication technologies, the distribution costs for modest data volumes have largely been eliminated, and typically do not justify any cost recovery on the distributor side. The costs of open data processes should be regarded as an intrinsic part of the cost of doing the research, and supported appropriately. Where the management and handling of very large data volumes (“big data”) may incur significant costs and such costs cannot be funded as part of the original research activity or through the operating budget of the data centre, some cost recovery may be justified even under a free and open access data policy.

5.3 *Data must be provided in a timely manner*

Data must be published as soon as practicable following collection, in near real-time if possible, unless access is limited on the basis of §5.1, or if necessary to support reasonable ethical, cultural, legal, operational (including data processing, quality control, and documentation), or scholarly requirements. Such data embargoes should be applied only for good cause and for the shortest time feasible to allow for good data processing practices and to respect the scientific endeavours of data creators. For SCAR data, a maximum embargo limit should be provided, along with documented reasons for the embargoed status.

5.4 *FAIR Principles should be applied to the greatest extent practicable*

To the greatest extent practicable, data should be made findable, accessible, interoperable, and reusable¹, subject to §5.1.

5.5 *All data must be accompanied by a complete set of metadata*

Data collected as part of SCAR programmes and projects must be accompanied by complete metadata, compliant with DIF or ISO-19115 standards. Complete metadata must contain sufficient information to understand, access, and replicate the data set to a level of quality, accuracy, and precision specified in the metadata. Metadata elements should provide a clear description of the data; their provenance, the data structure; calibrations; and methods, including units, associated errors, or other limitations where possible. Whenever possible, data and metadata should include spatial and

¹ Wilkinson *et al.* 2016. The FAIR guiding principles for scientific data management and stewardship. <https://doi.org/10.1038/sdata.2016.18>

temporal attribution, in order to enhance their interoperability and suitability for broader use. The metadata must also contain information to support appropriate attribution. Metadata should be public, even if the associated data cannot be made public.

5.6 Data should have persistent and globally unique identifiers

SCAR data should be assigned persistent and globally unique identifiers (PID) that remain linked to the data through republication or data aggregation processes, to facilitate unequivocal identification, attribution, data citation, provenance tracking, linking data with scientific results, and tracking distribution and impact of data collections. Examples of relevant PIDs include Digital Object Identifiers (DOIs) for unambiguously linking to data objects, or Open Researcher and Contributor Identifiers (ORCIDs) for unambiguously linking to individual researchers.

5.7 Data must be labelled as reusable

Data collected by SCAR programmes and projects must be labelled as reusable by attaching a rights waiver, a public domain statement, or an internationally recognised data licence to the data. Data licences should be non-restrictive, specifying that the data may be re-used, with no requirement more onerous than an acknowledgement of the data's source – for example, the Creative Commons Open Attribution Licence (CC-BY). Where possible, the rights waiver or licence should be assigned by the owner or source of the data, and these parties should be identified in accompanying metadata. Data providers are responsible for ensuring that data submitted to NADCs are suitable for publication in accordance with their specified license.

5.8 Data sources should be attributable and attributed

Users of SCAR data should formally acknowledge data authors and sources².

5.9 Data must be appropriately preserved for the long term

National Antarctic Data Centres must ensure that SCAR data held by them are preserved in such a manner that they are enduring, resilient to corruption or loss, and remain accessible via the Antarctic Metadata Directory.

5.10 Data management and long-term curation must be planned and resourced

All approved SCAR science projects will be required to develop a Data Management Plan. The Data Management Plan should demonstrate that the data will retain its value over the long-term, and the data will be curated effectively in a trusted manner³. Data Management Plans will be reviewed by

² Data Citation Synthesis Group 2014. Joint Declaration of Data Citation Principles.
<https://doi.org/10.25490/a97f-egyk>

³ Lin, D., Crabtree, J., Dillo, I. *et al.* 2020. The TRUST principles for digital repositories.
<https://doi.org/10.1038/s41597-020-0486-7>

the host National Antarctic Data Centre proposed in the Data Management Plan, or by SCADM in cases where the choice of a host is unclear.

5.11 Quality assurance and fitness-for-purpose

Data providers are responsible for any quality assurance and quality control required to meet community standards. Data users are responsible for ensuring that the data they use is fit-for-purpose.

5.12 Reporting

SCAR NADCs will deliver an annual report to SCADM, providing a summary of their SCAR data collections.

6. Related Legislation and Documents

[The Antarctic Treaty 1959](#)

[International Polar Year 2007-2008 Data Policy](#)

[The FAIR guiding principles for scientific data management and stewardship](#)

[Creative Commons Attribution 4.0 International Public License](#)

[Alignment of Polar Data Policies – Recommended Principles](#)

[Data Citation Synthesis Group: Joint Declaration of Data Citation Principles](#)

[The TRUST principles for digital repositories](#)

7. Feedback

Questions and feedback relating to this document can be sent to data@antarctica.gov.au.

8. Approval and Review Details

Approval and Review	Details
Approval Authority	SCAR Directors
Administrator	SCAR Secretariat

Approval and Amendment History	Details
Original Approval Authority and Date	SCAR Directors, 14 May 2021
Amendment Authority and Date	[Relevant approval authority, DD MMM YYYY]
Notes	[If relevant, add notes to alert readers about the document/s this one replaces, e.g. This document consolidates and replaces X Policy, Y Procedure and Z Procedure.]

Appendix – List of Acronyms

ATCM	Antarctic Treaty Consultative Meeting
CC-BY	Creative Commons Open Attribution Licence
DIF	Data Interchange Format
DOI / doi	Digital Object Identifier
FAIR	Data principles of F indability, A ccessibility, I nteroperability, and R eusability
GEO	Group on Earth Observations
IASC	International Arctic Science Committee
ICSU	International Council for Science
IOC	Intergovernmental Oceanographic Commission
ISC	International Science Council
ISO	International Organization for Standardization
ISSC	International Social Science Council
NADC	National Antarctic Data Centre
OECD	Organisation for Economic Co-operation and Development
ORCID	Open Researcher and Contributor Identifiers
PID	Persistent Identifier
TRUST	Data principles of T ransparency, R esponsibility, U ser focus, S ustainability and T echnology
UNESCO	United Nations Educational, Scientific and Cultural Organization
UN-GGIM	United Nations Committee of Experts on Global Geospatial Information Management
WMO	World Meteorological Organisation