



Digital Antarctica

C. Requirements

May 2022

Contents

Executive Summary	3
Requirements	3
Functional Requirements	4
FR1 - Digital Antarctica shall enable the access of in-scope data across multiple data sources	4
FR2 - Similar data made available via Digital Antarctica shall be presented in a standardised format	4
FR3 - Digital Antarctica shall provide all available provenance details for any data	4
Non-Functional Requirements	5
NR1 - Digital Antarctica shall enable interoperability of data across multiple sources	5
NR2 - Digital Antarctica shall provide guidance on standards	5
NR3 - Where practicable, Digital Antarctica shall investigate and re-use existing standards and services	5
NR4 - Digital Antarctica data items shall be presented with no loss of information or precision	5
NR5 - Digital Antarctica shall not compromise an organisation's data policy	6
NR6 - Data that cannot be shared to the level of Digital Antarctica standards should still be available via Digital Antarctica	6

Version information

Version	Description	Author	Date
1.0	Initial release for DARG review	Rob Jennings	20/4/2022
2.0	Update after DARG review	Rob Jennings	4/5/2022

© Copyright Australian Antarctic Program Partnership 2022



This work by the Australian Antarctic Program Partnership is licensed under a Creative Commons Attribution (CC BY) 4.0 International License.

Details of the licence are available at: <https://creativecommons.org/licenses/by/4.0/>

Executive Summary

This document lists the functional and non-functional requirements for *Digital Antarctica* as gathered over the course of the *Digital Antarctica* analysis. There are nine top level requirements identified, four of which have one or more sub-requirements.

The document describes the requirements that any solution for *Digital Antarctica* must meet. These requirements briefly define how data should be shared by participant organisations and the standards that will be required to enable that sharing.

Requirements

Requirements listed in the table below are split into two categories, functional and non-functional.

- Functional requirements are requirements that will impact the functionality of a potential solution. Each requirement should be clear and testable, and should relate to a specific operation of any potential solution.
- Non-functional requirements are requirements that can reflect how the potential solution performs, and may reflect the effect a potential solution should have. As with functional requirements, they should be clear and testable, although the tests may be broader in scope.

Note that requirements are not design specifications. They do not define a solution or functionality, but instead define what a potential solution needs to offer or achieve.

Further explanation of each top-level requirement follows the table:

Req. ID	Requirement
<i>Functional Requirements</i>	
FR1	<i>Digital Antarctica</i> shall enable the access of in-scope data across multiple data sources.
FR2	Similar data made available via <i>Digital Antarctica</i> shall be presented in a standardised format.
FR2.1	Similar data shall be presented in a standardised structure.
FR2.2	<i>Digital Antarctica</i> shall enable standardised nomenclature for similar values.
FR2.3	<i>Digital Antarctica</i> shall enable standardised formats for similar values.
FR3	<i>Digital Antarctica</i> shall provide all available provenance details for any data.
<i>Non-functional Requirements</i>	
NR1	<i>Digital Antarctica</i> shall enable interoperability of data across multiple sources.
NR1.1	<i>Digital Antarctica</i> shall enable data analytics to be performed using data from multiple data sources.
NR1.2	<i>Digital Antarctica</i> shall enable search of data across multiple data sources.
NR2	<i>Digital Antarctica</i> shall provide guidance on standards.
NR3	Where practicable, <i>Digital Antarctica</i> shall investigate and re-use existing standards and services.
NR4	<i>Digital Antarctica</i> data items shall be presented with no loss of information or precision.
NR4.1	If data presented via <i>Digital Antarctica</i> is altered to meet the requirements of <i>Digital Antarctica</i> , a clear understanding of the data processing will be provided.
NR5	<i>Digital Antarctica</i> shall not compromise an organisation's existing data policy.
NR5.1	Data made available via <i>Digital Antarctica</i> will should be appropriately licensed for re-use
NR6	Data that cannot be shared to the level of <i>Digital Antarctica</i> standards should still be available via <i>Digital Antarctica</i>

Functional Requirements

FR1 - Digital Antarctica shall enable the access of in-scope data across multiple data sources

Digital Antarctica will ensure Australian Antarctic and Southern Ocean research data from participating data sources shall be available to users to view, download, and to incorporate directly into analytical tools.

FR2 - Similar data made available via Digital Antarctica shall be presented in a standardised format

To facilitate interoperability, similar data from any source shall be presented with standardised structure, nomenclature, and format.

FR2.1 - Similar data shall be presented in a standardised structure.

For the purposes of this requirement, similarity will be determined using a combination of one or more attributes including, but not limited to, data type, research category, and data format. For example, sea surface temperature data in a database may be presented differently to sea surface temperature data in a netCDF file. Similarly, a database of sea surface temperature data may present differently from a database of penguin habitat data.

Data of various formats will have a specific structure. For example, databases may have a hierarchical parent/child structure, while XML files may have a structure that conforms to XML standards. Data made available via *Digital Antarctica* will, where practicable, be presented in standardised formats, based on its data type, to enable ease of integration.

FR2.2 - Digital Antarctica shall enable standardised nomenclature for similar values.

Datasets from multiple sources may have similar or identical values within them. Most data defining a spatial area will, for example, have latitude and longitude values defined. However, those values may have different labels within datasets. Latitude may be labelled “Latitude” or “Lat” and Longitude may be labelled “Longitude”, “Long”, or “Lon”. It is a requirement of *Digital Antarctica* that, where practicable, fields with the same content are presented with the same labels. For example, all fields containing a longitudinal value will have the label “Longitude”.

FR2.3 - Digital Antarctica shall enable standardised formats for similar values.

Data with like values should, where practicable, be presented in the same format. Fields containing text, for example, should have the same string properties. Fields containing numbers should have the same numerical properties (e.g. precision). This requirement should not affect the accuracy of the data being presented. For example, a number stored in the data should not, under normal circumstances, be rounded to a certain number of decimals to meet this requirement. Rather the requirement should be met in such a way that such an operation is unnecessary.

FR3 - Digital Antarctica shall provide all available provenance details for any data

Any user who has retrieved data via *Digital Antarctica* must be able to reference the primary location of those data. *Digital Antarctica* will not be considered a data source for the purposes of publications.

This applies to any data item, whether a full dataset or a subset of data, made available via *Digital Antarctica*.

If provenance details beyond that source are available, they should also be made available to a *Digital Antarctica* user.

Non-Functional Requirements

NR1 - Digital Antarctica shall enable interoperability of data across multiple sources

NR1.1- Digital Antarctica shall enable data analytics to be performed using data from multiple data sources.

NR1.2 - Digital Antarctica shall enable search of data across multiple data sources.

With commonality across attributes (for example, keywords, values, labels etc.), data from multiple sources should be able to be brought together with minimal effort.

By meeting requirements FR1 and FR2, data made available via *Digital Antarctica* shall be presented in such a way as to enable interoperability of data across multiple sources. This will facilitate the data's use in analytical tools, and allow searches for data and metadata to be performed across multiple data sources.

NR2 - Digital Antarctica shall provide guidance on standards

The documentation of the standards defined in FR2, and the expertise gained in generating meeting requirement FR1 and FR2, may be used by any organisation wishing to standardise their data services, either for incorporation into *Digital Antarctica* or as a best-practice exercise.

NR3 - Where practicable, Digital Antarctica shall investigate and re-use existing standards and services

Standards already exist for much of the data stored and shared by Australian Antarctic data centres. Existing standards should be tested against defined criteria such as:

- Current usage
- Applicability
- Maintainability
- Ease of use

For example, if an existing standard is already in use by many organisations, meets the requirements of *Digital Antarctica*, is simple to implement, and can be easily maintained then it should be used in favour of creating a new standard.

NR4 - Digital Antarctica data items shall be presented with no loss of information or precision.

While *Digital Antarctica* will facilitate data interoperability, and may enable data users to find and retrieve subsets of data, any data returned by *Digital Antarctica* should be preserved, accurately reflecting the source data. While data may be transformed to meet the interoperability standards of *Digital Antarctica*, it should not be altered to the point where its original value is obfuscated.

A dataset may, for example, consist of 1,000 rows of CTD data (conductivity, temperature, and depth), which was originally measured in degrees Fahrenheit, with each column representing a parameter of the CTD measurement, and each row representing a particular sample. A user of *Digital Antarctica* may request only the temperature data, and only from the first 300 rows. The resulting dataset will be a subset of the original, and each data item in that grid may display data in degrees Celsius, but will have enough precision to enable re-conversion back to Fahrenheit if necessary. It will otherwise be identical to its source record.

NR4.1 - If data presented via Digital Antarctica is altered to meet the requirements of Digital Antarctica, a clear understanding of the data processing will be provided.

The nature of any change to any data item in *Digital Antarctica* (for example, format conversion) will be communicated to the user. This may be represented in the metadata, in the data itself, or in some other way.

NR5 - Digital Antarctica shall not compromise an organisation's data policy

Any *Digital Antarctica* solution shall perform within the parameters of each organisation's policy and architecture in relation to elements such as security and performance. Organisations may, however, choose to update their policy to accommodate *Digital Antarctica* implementation.

NR5.1 - Data made available via Digital Antarctica will should be appropriately licensed for re-use

As per the *Digital Antarctica* scope, all data available through *Digital Antarctica* shall have an appropriate licence which makes the data available for re-use.

NR6 - Data that cannot be shared to the level of Digital Antarctica standards should still be available via Digital Antarctica

If data cannot be shared to the standards defined by *Digital Antarctica* (for example, it cannot be grouped with other similar data, or cannot be transformed or presented in a standardised format), but is still otherwise available digitally, it should be made available via *Digital Antarctica*.