



Technical documentation

Acronyms and Concepts

Versions

Version	Date	Comment	Responsible
0.4	2022-02-25	Added use metadata	Øystein Godøy
0.3	2020-03-15	Cleaning of old comments and minor modifications.	Øystein Godøy
0.2	2017-03-17	Comments from Markus.	Markus Fiebig Øystein Godøy
0.1	2017-02-03	First draft based on a similar document for NorDataNet.	Øystein Godøy

This work is released under the Creative Commons Attribution 4.0 License. To view a copy of the license, visit <https://creativecommons.org/licenses/by/4.0/>.



Table of Contents

- 1. Introduction 3
- 2. Scope of document 4
- 3. Applicable documents 4
- 4. Acronyms..... 4
- 5. Concepts..... 6
 - 5.1. Machine interface 6
 - 5.2. Dataset 6
 - 5.3. Data Discovery 6
 - 5.4. Discovery metadata 6
 - 5.5. Discovery service..... 7
 - 5.6. Download service 7
 - 5.7. Metadata 7
 - 5.8. Product..... 7
 - 5.9. Product cart..... 7
 - 5.10. Product cart item..... 7
 - 5.11. Product information..... 7
 - 5.12. Service 7
 - 5.13. Spatial data 8
 - 5.14. Spatial data set..... 8
 - 5.15. Use metadata 8

1. Introduction

Environmental and climate changes are currently observed at a global scale and in particular in the Arctic. In order to give better estimates of the future changes, the Arctic has to be monitored and analysed by a multi-disciplinary observation system which is suited to validate and gradually improve Earth System Models. The best chance to achieve significant results within a relatively short time frame is found in regions with a large natural climate gradient, and where processes sensitive to the expected changes are particularly important.

Svalbard and the surrounding ocean areas fulfil all these criteria: Svalbard is located in a region with a very large climate gradient, being alternately influenced by cold central Arctic or mild marine climate conditions at time scales of weeks to years. It is also located in the region with the strongest inflow and outflow processes between the Arctic and lower-latitude oceans. In addition, Svalbard is the only region in the world (and has the facilities) where one can study and quantify one of the remaining unknowns in the climate puzzle: the extraterrestrial and especially solar influence on climate.

The vision for the Svalbard Integrated Arctic Earth Observing System (SIOS) is to be a regional observational system for long term acquisition and proliferation of fundamental knowledge on global environmental change (GEC) within an Earth System Science (ESS) perspective in and around Svalbard. SIOS will systematically develop and implement methods for how observational networks are to be construed and thus become a leader regarding observational systems in the Arctic and Polar regions. The SIOS Data Management System (SDMS) Data Portal is the entry point to SIOS datasets. It offers a web interface that contains information about datasets (metadata). These metadata are harvested on a regular basis from data centres contributing to SIOS. These data centres manage the data on behalf of the owners/providers of the data.

A major innovative element of SIOS will be the Knowledge Centre (KC), which will facilitate interaction between observation, modelling and process research, strategic processes, a service point to user communities and a platform for data handling and utilization. The SIOS Data Management System (SDMS) will be a functionality enabling component of the Knowledge Centre, supporting data submission, discovery, access, use and preservation of SIOS relevant data sets ([\[RD-1\]](#), [\[RD-2\]](#)).

For the SDMS the term “data set” is defined in line with the INSPIRE Directive as “an identifiable collection of spatial data”, i.e. a collection of data that has a reference by name or coordinates to a geographic location or area, and which in addition have a designated start and end time. A data set can contain observations (remote or in situ), derived quantities (from either of these two types of data sources) or forecasts of future states of environmental parameters. The data values can be located at a single point, along a line or transect, in a regular or irregular grid, and be captured or estimated at one or more altitudes/depths. A data set can be stored on paper, in files (one or more), or in a database, and is often accompanied by descriptions (metadata) of its content. The purpose of the SDMS is to the SIOS user community with a single entry point to relevant datasets. It is however not the purpose to centralise all data, but rather to integrate existing and future data centres contributing to SIOS using machine readable interfaces to metadata and data.

The first version of this document is based on a similar document developed for NorDataNet.

2. Scope of document

This document describes various acronyms and concepts needed to establish a documentation framework for the definition of the SIOS Data Management System (SDMS). Wherever possible, the document refers to online documentation for further reading. A brief knowledge of the concepts identified in this document is required to follow the documentation system SIOS is using.

3. Applicable documents

[RD-1] [SDMS Interoperability Guidelines Document](#)

[RD-2] SIOS Operations Manual (Not Public available)

4. Acronyms

Acronym	Description	Resources
CSW	Catalogue Service for the Web. This is a protocol that is used for distributed search in metadata and to exchange metadata.	http://en.wikipedia.org/wiki/Catalog_Service_for_the_Web
CF	Climate and Forecast Metadata Conventions This is a metadata framework focusing on use metadata. It is primarily used to describe the detailed contents of a data file (standardised names for variables, units etc) as well structures that allow generic software to automatically read, understand and analyse/visualise the data.	http://en.wikipedia.org/wiki/Climate_and_Forecast_Metadata_Conventions , http://cf-pcmdi.llnl.gov/ http://cfconventions.org/index.html
DOI	Digital Object Identifier	https://www.doi.org/ https://en.wikipedia.org/wiki/DOI
FTP	File Transfer Protocol	http://en.wikipedia.org/wiki/Ftp
GML	Geography Markup Language	http://en.wikipedia.org/wiki/Geography_Markup_Language
HTTP	HyperText Transfer Protocol	http://en.wikipedia.org/wiki/Http
KML	Keyhole Markup Language	http://en.wikipedia.org/wiki/KML
NetCDF	Network Common Data Form This is a file format that is extensively used for gridded data, but that recently has been more and more used for observations as well. It shares many features with databases and is closely linked to a structural and semantic framework named CF.	http://en.wikipedia.org/wiki/Netcdf , http://www.unidata.ucar.edu/software/netcdf/

Acronym	Description	Resources
	Open Archives Initiative - Protocol for Metadata Harvesting	
OGC	Open Geospatial Consortium	http://www.opengeospatial.org/
OPeNDAP	Open-source Project for a Network Data Access Protocol	http://en.wikipedia.org/wiki/OPeNDAP , http://www.opendap.org/
REST	Representational State Transfer	http://en.wikipedia.org/wiki/Representational_state_transfer
SKOS	Simple Knowledge Organisation System	https://www.w3.org/2004/02/skos/ https://en.wikipedia.org/wiki/Simple_Knowledge_Organization_System
SOA	Service Oriented Architecture	http://en.wikipedia.org/wiki/Service-oriented_architecture
SOAP	Simple Object Access Protocol	http://en.wikipedia.org/wiki/SOAP
THREDDS	Thematic Realtime Environmental Distributed Data Services	http://www.unidata.ucar.edu/projects/THREDDS/ , http://www.unidata.ucar.edu/software/tds/
URI	Uniform Resource Identifier	http://en.wikipedia.org/wiki/Uniform_resource_identifier
URL	Uniform Resource Locator	http://en.wikipedia.org/wiki/Url
WCS	Web Coverage Service	http://en.wikipedia.org/wiki/Web_Coverage_Service
WFS	Web Feature Service	http://en.wikipedia.org/wiki/Web_Feature_Service
WIGOS	WMO Integrated Global Observing System	https://www.wmo.int/pages/prog/www/wigos/index_en.html
WIS	WMO Information System	http://www.wmo.int/pages/prog/www/WIS/
WMDS	WIGOS Metadata Standard	Not updated as of 2017-02-03, new public version expected April 2017.
WMS	Web Map Service	http://en.wikipedia.org/wiki/Web_Map_Service
WMTS	Web Map Tile Service	http://en.wikipedia.org/wiki/Web_Map_Tile_Service
WPS	Web Processing Service	http://en.wikipedia.org/wiki/Web_Processing_Service
XML	Extensible Markup Language	http://en.wikipedia.org/wiki/Xml

5. Concepts

5.1. Machine interface

For computers to communicate without human intervention it is necessary to have both well defined formats for exchange of messages (protocols) and semantic frameworks that clearly define meaning of various concepts. The combination of protocols and semantic frameworks defines a machine interface.

5.2. Dataset

A data set is

- a. a collection of measurements that are acquired by one or more instruments within a well defined and known geographical area and time period, or
- b. a collection of analysed or simulated data that are generated by a specific processing chain, and that have a well defined and known geographical area and time period,
- c. that are described by an informative and (preferably)
 - i. unique title,
 - ii. a version number (alphanumeric),
 - iii. an abstract summarising the content, marked up with keywords from designated vocabularies,
 - iv. and that have a unique identifier.

Examples of data sets include

- a. all measurements collected during a field experiment (e.g. scientific cruise) or seismic campaign,
- b. all measurements collected by a meteorological station or weather ship, drifting of fixed buoy, glider or cluster of acoustical cabled moorings, as well as biological or physical samples
- c. all time steps of a numerical simulation, being a weather prediction model, a climate model or something else,
- d. all time steps in an analysed satellite remote sensing product

All datasets in SIOS are spatial datasets(see [Section 5.14](#)) in the sense that they have a temporal and geographical reference.

5.3. Data Discovery

This is the process of finding the relevant datasets (data and products) based upon filtering of search criteria (temporal, spatial, scientific). The process of finding out how to use the data found is data mining.

5.4. Discovery metadata

Discovery or descriptive metadata are typically used to find relevant datasets or resources. This kind of metadata sometimes overlap with use metadata, but are not necessarily useful to fully understand the

nature of a dataset. It represents the *who, what, when, where, why* and how of the resource.

5.5. Discovery service

Distinct part of the functionality that is provided by an entity through interfaces for the inquiry of the nature and content of a spatial resource.

5.6. Download service

Distinct part of the functionality that provides access to the full extent of geographic and thematic information in a data set. By access to the full extent is meant access to features described with feature attributes and real world coordinates. A download service may includes elements of a feature type service, a coordinate conversion/transformation service and schema transformation service.

5.7. Metadata

Metadata are data about data. A metadata record is a file of information, usually presented as an XML document, which captures the basic characteristics of a data or information resource. Metadata are often divided in Discovery metadata (see [Section 5.4](#)) and use metadata (see [Section 5.15](#)) where the first describes overall characteristics of the dataset while the latter is required to actually use the dataset without previously knowing the contents or speaking to the data provider. See <http://www.fgdc.gov/metadata> for details.

5.8. Product

A spatial data set with associated metadata.

5.9. Product cart

A collection of product cart items.

5.10. Product cart item

Either a single product or a composite collection of products.

5.11. Product information

Information related to a product. This information is typically too bulky to include in the metadata, such as written procedures on how to collect data, verbose explanations on models used to interpolate data between measurements, etc.

5.12. Service

Distinct part of the functionality that is provided by an entity through interfaces (ISO19119). In computing terms, a service is an application that provides information and/or functionality to other applications.

Services are typically non-human-interactive applications that run on servers and interact with applications via an interface.

5.13. Spatial data

Any data with a direct or indirect reference to a specific location or geographic area. (INSPIRE Directive)

5.14. Spatial data set

An identifiable collection of spatial data. (INSPIRE Directive)

5.15. Use metadata

Use metadata are metadata that are necessary to fully understand and utilise the data they describe. It includes information like standardised names on variables, how missing values are captured, what the units of the variable is, map projections used etc.