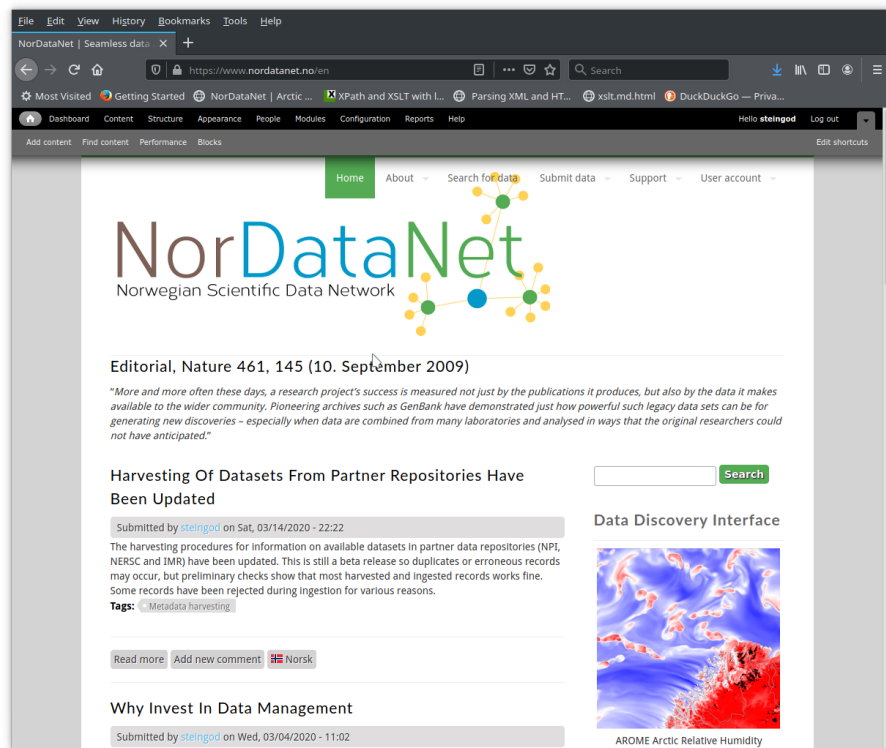


An illustration of a blue iceberg floating in a dark blue sea under a light blue sky. The iceberg is rendered in a low-poly, geometric style. The visible tip is above the water line, while the much larger, jagged base is submerged. The text is overlaid on the image.

Norwegian Scientific Data Network and linkages

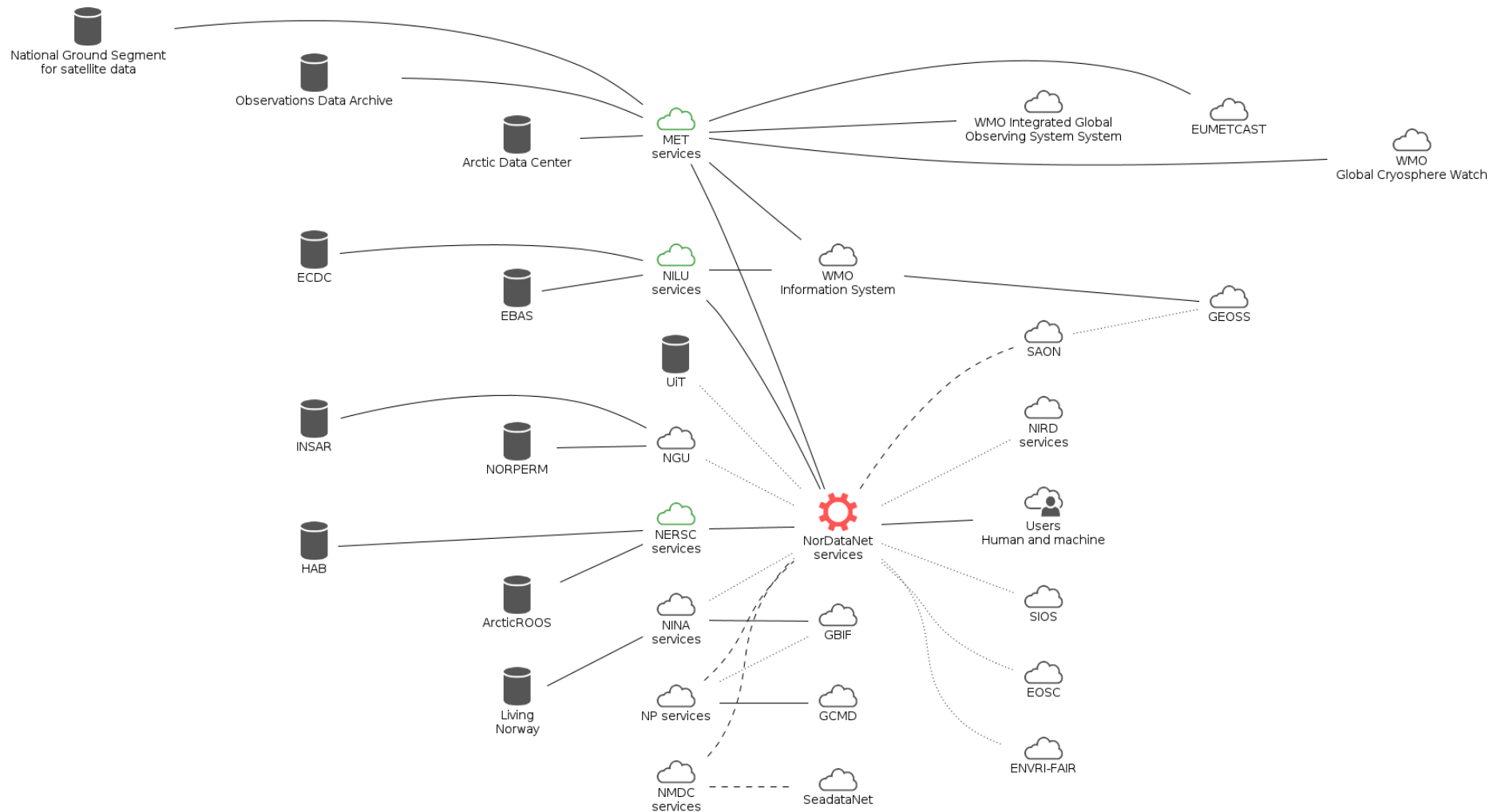
Øystein Godøy, Stein Tronstad, Torill Hamre,
Markus Fiebig, Helge Sagen and Lara Ferrighi

Norwegian Scientific Data Network - NorDataNet

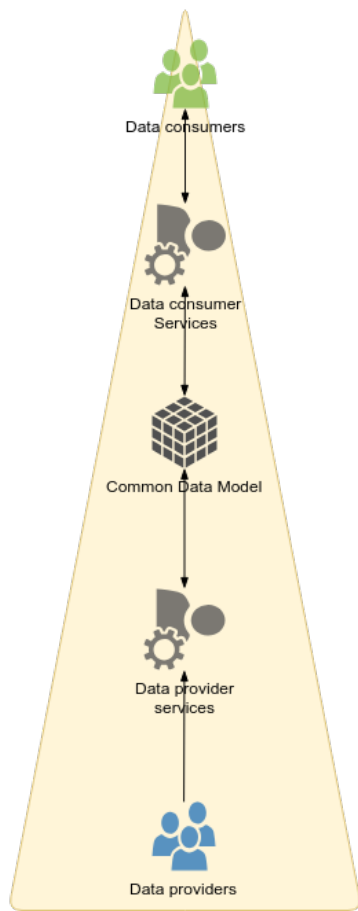


- A distributed data management network
 - Building on existing data centres nationally and the legacy of IPY
- Originally focused on geoscientific data
 - But requested by funder to think interdisciplinary
- Development funded by the Research Council of Norway (RCN) through the infrastructure programme
- Being part of a family of national e-infrastructures focusing on geoscientific data, in particular related to
 - Norwegian Satellite Earth Observation Database for Marine and Polar Research (NORMAP)
 - Norwegian Marine Data Centre (NMDC)
 - Svalbard Integrated Arctic Earth Observing System (SIOS)
- But also being linked with
 - Norwegian Infrastructure for Research data (NIRD)
 - Living Norway/GBIF-Norway
- Target community is scientists

NorDataNet concept outline



Approach



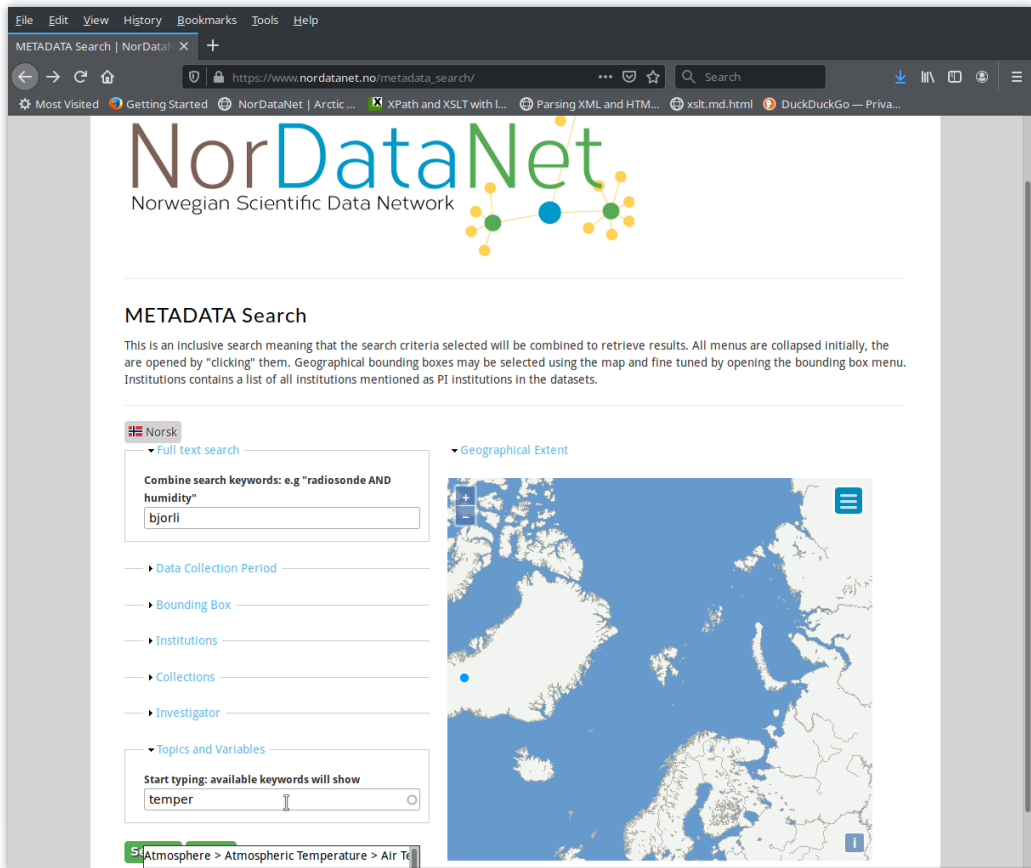
- Discovery metadata are harvested into a unified catalogue using
 - OAI-PMH, OGC CSW, (OpenSearch – only testing)
 - Standardised interfaces and documentation is a prerequisite
- Moving towards standardised data hosted by core partners
 - Machine actionable data is required to establish cost efficient user oriented services
 - Strong focus on NetCDF-CF wherever possible to achieve FAIR data
 - Most widely used FAIR compliant data encoding approach used within geoscience across communities etc.

The FAIR Guiding Principles for scientific data management and stewardship

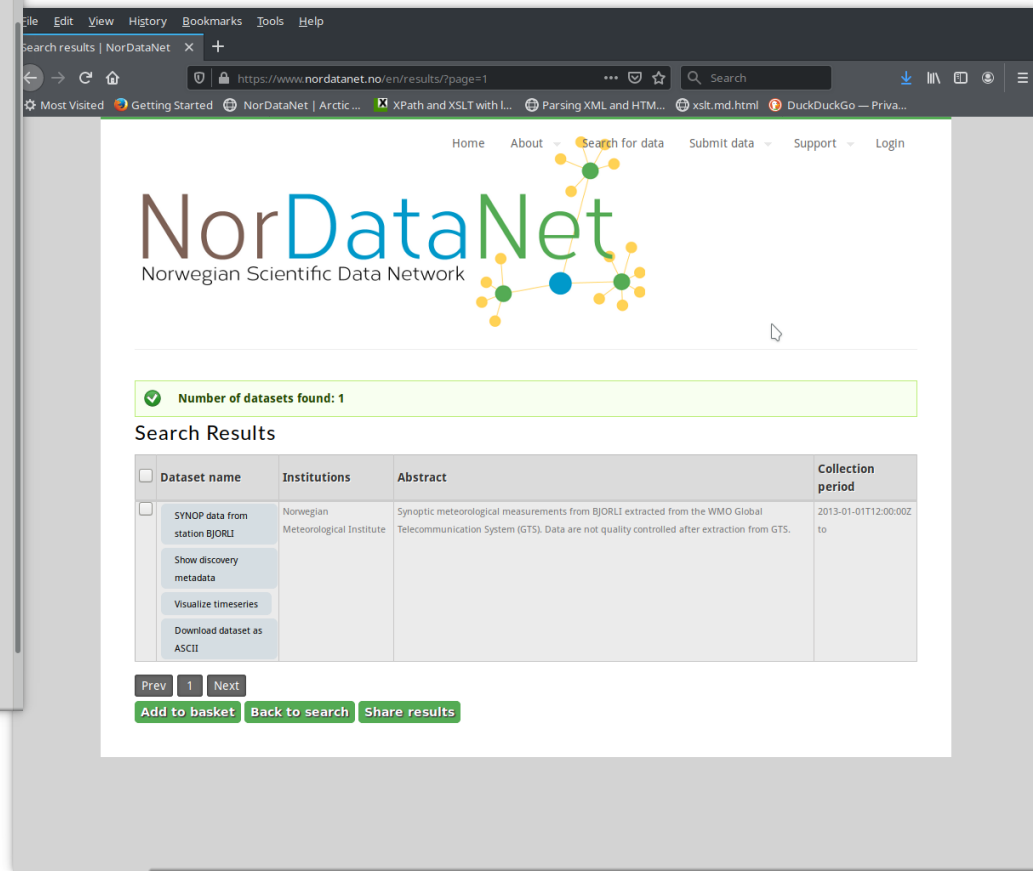
- To be Findable:
 - F1. (meta)data are assigned a **globally unique and persistent identifier**
 - F2. **data are described with rich metadata** (defined by R1 below)
 - F3. metadata clearly and explicitly include the identifier of the data it describes
 - F4. **(meta)data are registered or indexed in a searchable resource**
- To be Accessible:
 - A1. (meta)data are retrievable by their identifier using a **standardized communications protocol**
 - A1.1 the protocol is open, free, and universally implementable
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
 - A2. metadata are accessible, even when the data are no longer available
- To be Interoperable:
 - I1. **(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.**
 - I2. **(meta)data use vocabularies** that follow FAIR principles
 - I3. (meta)data include qualified references to other (meta)data
- To be Reusable:
 - R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 - R1.1. **(meta)data are released with a clear and accessible data usage license**
 - R1.2. (meta)data are associated with detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards

Types of metadata for datasets

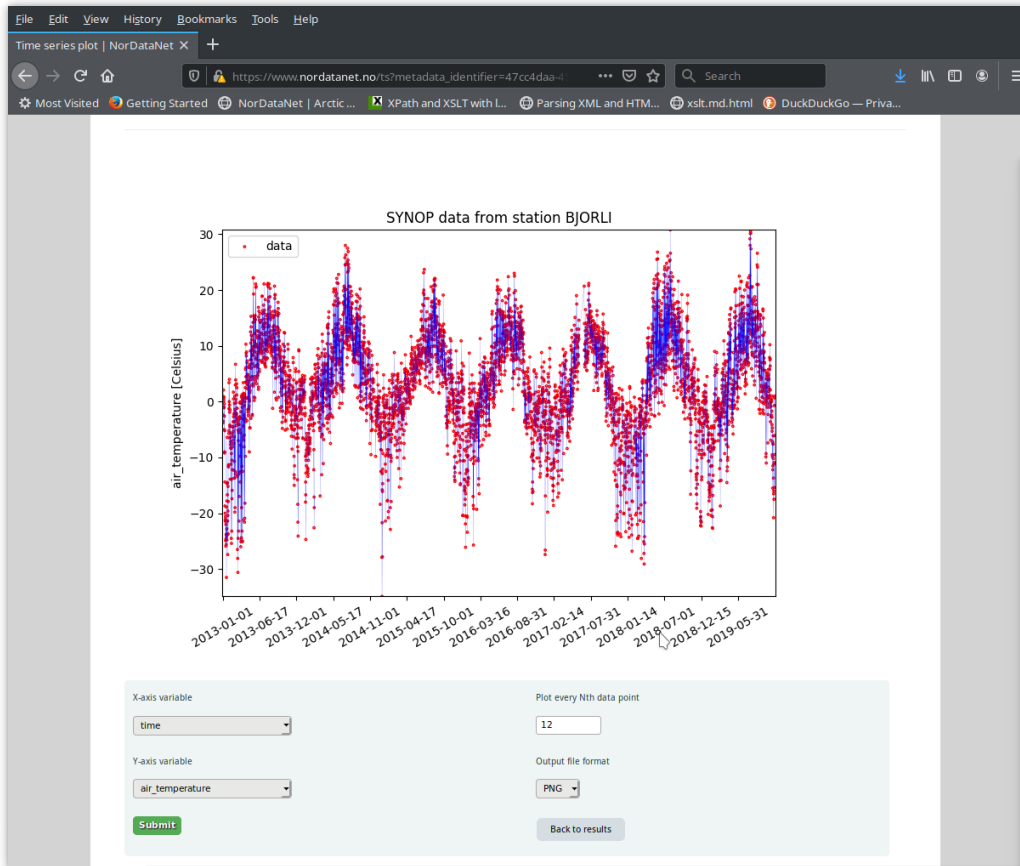
Type	Purpose	Description	Examples
Discovery metadata	Used to find relevant data	Discovery metadata are also called index metadata and are a digital version of the library index card. It describes who did what, where and when, how to access data and potential constraints on the data. It shall also link to further information on the data like site metadata.	ISO19115 (converted to MMD) GCMD DIF (converted to MMD) ACDD (converted to MMD) MMD (used by NorDataNet)
Use metadata	Used to understand data found	Use metadata are describing the actual content of a dataset and how it is encoded. The purpose is to enable the user to understand the data without any further communication. It describes content of variables using standardised vocabularies, units of variable, encoding of missing values, map projections etc.	Climate and Forecast Convention BUFR GRIB DwCA
Configuration metadata	Used to tune portal services for datasets for users.	Configuration metadata are used to improve the services offered through a portal to the user community. This can be e.g. how to best visualise a product.	MMD (used by NorDataNet)
Site metadata	Used to understand data found	Site metadata are used to describe the context of observational data. It describes the location of an observation, the instrumentation, procedures etc. To a certain extent it overlaps with discovery metadata, but more so it really extends discovery metadata. Site metadata can be used for observation network design.	WIGOS OGC O&M



Screenshot examples of existing functionality - search



Visualisation and transformation of timeseries



ASCII data download | NorDataNet

https://www.nordatanet.no/csv?metadata_identifier=47cc4daa-4

<input type="checkbox"/>	Standard name	Units
<input type="checkbox"/>	air_pressure	hPa
<input type="checkbox"/>	air_pressure_at_sea_level	hPa
<input type="checkbox"/>	air_temperature	Celsius
<input type="checkbox"/>	dew_point_temperature	Celsius
<input type="checkbox"/>	precipitation_amount	kg m-2
<input type="checkbox"/>	relative_humidity	percent
<input type="checkbox"/>	sea_surface_temperature	Celsius
<input type="checkbox"/>	thickness_of_snowfall_amount	cm
<input type="checkbox"/>	time	days since 1970-01-01 00:00:00
<input type="checkbox"/>	wind_from_direction	degree
<input type="checkbox"/>	wind_speed	m s-1
<input type="checkbox"/>	wind_speed_of_gust	m s-1

Output format: CSV

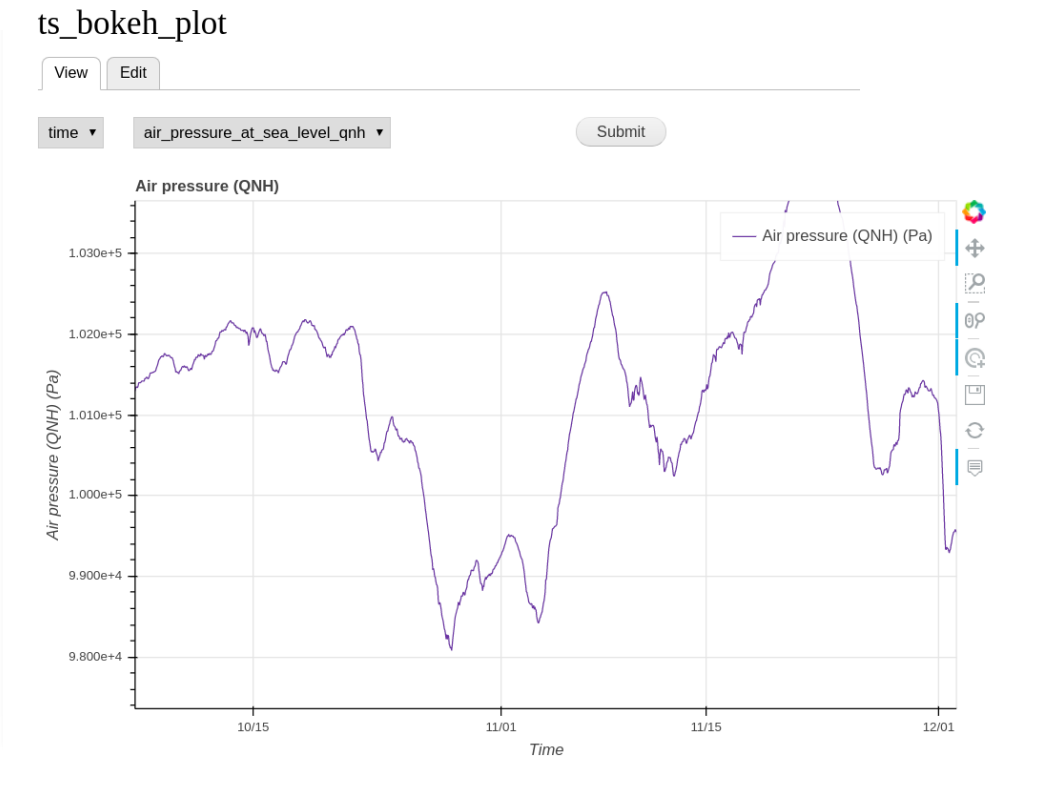
Submit

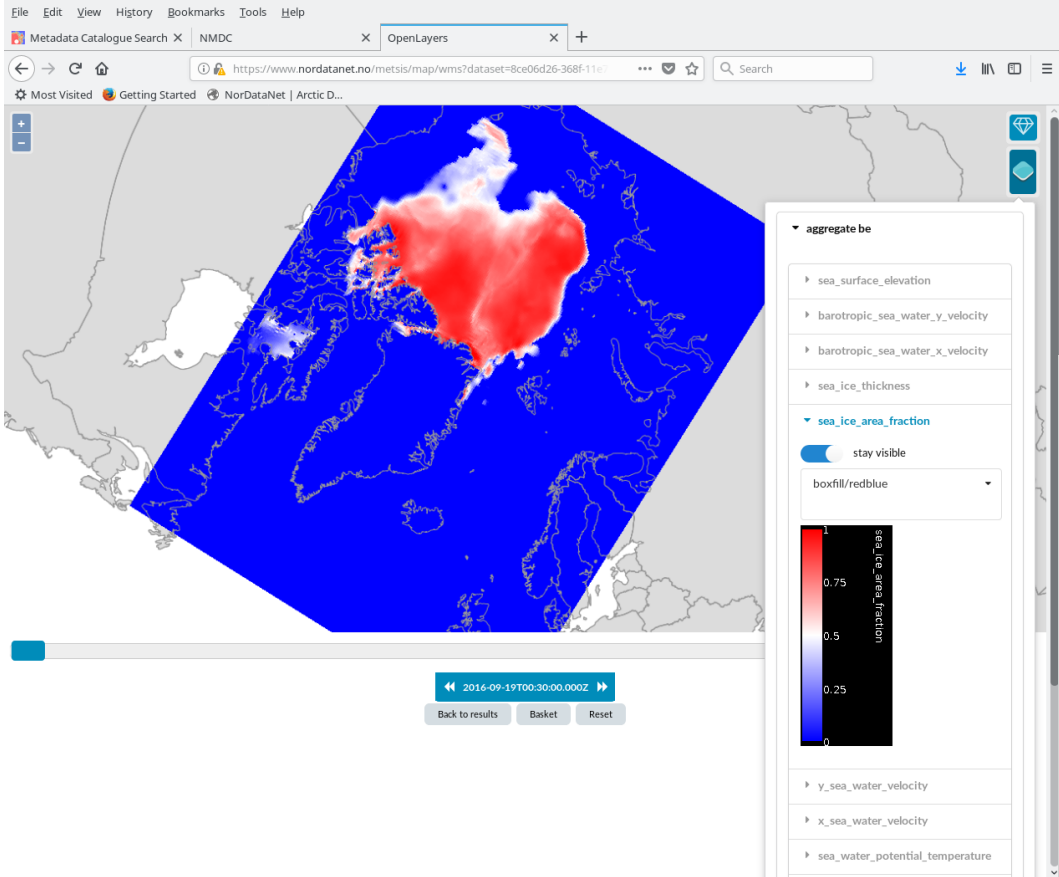
Back to results

Time Series plotting – new dynamic presentation

<input type="checkbox"/>	Visualize timeseries Download dataset as ASCII		
<input type="checkbox"/>	Observations from Kongsøya Download data View metadata Visualize timeseries Download dataset as ASCII	Norwegian Meteorological Institute	Quality controlled timeseries from climate consistent following a nu
<input type="checkbox"/>	Observations from Karl XII-Øya	Norwegian Meteorological Institute	Quality controlled timeseries from climate consistent following a nu

Upgrade and extension: Work in progress!





Visualisation and transformation of gridded products

Select temporal extent

Start date

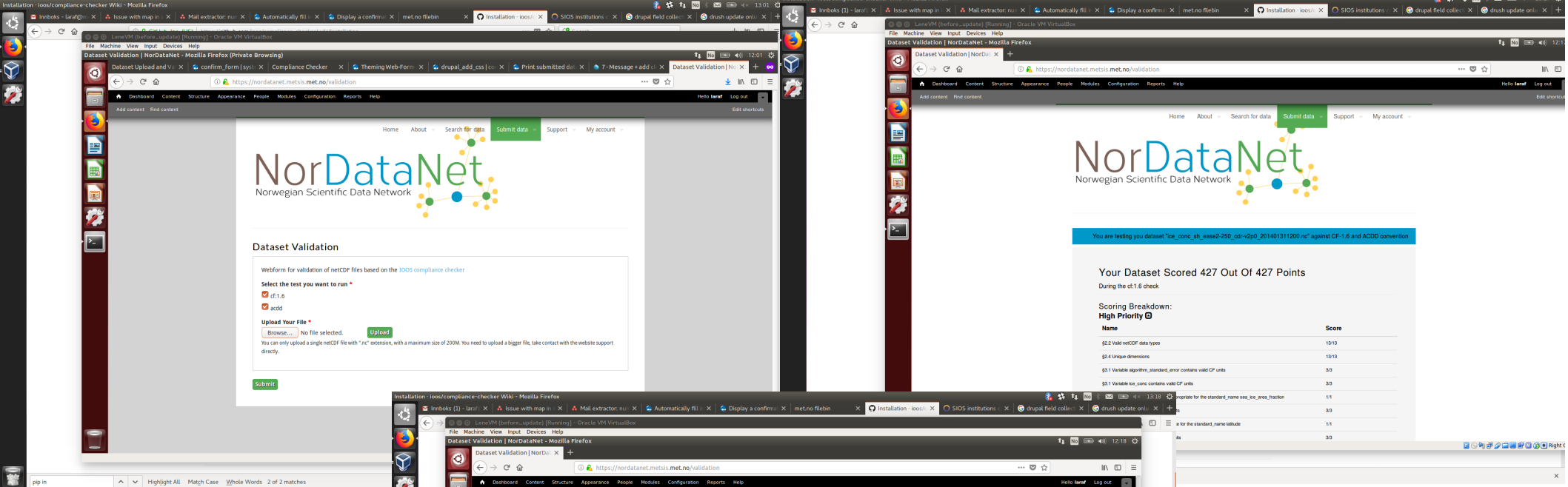
Stop date

Select variables

<input type="checkbox"/>	Name	Standard name	Long name	Units
<input type="checkbox"/>	h	sea_floor_depth_below_sea_level	sea_floor_depth_below_sea_level	meter
<input type="checkbox"/>	latitude	latitude	latitude	degree_north
<input type="checkbox"/>	longitude	longitude	longitude	degree_east
<input type="checkbox"/>	mask		mask on RHO-points	
<input type="checkbox"/>	aice	sea_ice_area_fraction	fraction of cell covered by ice	
<input type="checkbox"/>	hice	sea_ice_thickness	average ice thickness in cell	meter
<input type="checkbox"/>	salinity	sea_water_salinity	salinity	1e-3
<input type="checkbox"/>	temperature	sea_water_potential_temperature	Sea water potential temperature	Celsius
<input type="checkbox"/>	u	x_sea_water_velocity	Sea water x velocity	meter second-1
<input type="checkbox"/>	ubar	barotropic_sea_water_x_velocity	Barotropic sea water x velocity	meter second-1
<input type="checkbox"/>	v	y_sea_water_velocity	Sea water y velocity	meter second-1
<input type="checkbox"/>	vbar	barotropic_sea_water_y_velocity	Barotropic sea water y velocity	meter second-1
<input type="checkbox"/>	zeta	sea_surface_elevation	Sea surface height above geoid	meter

Select map projection

x-axis from:



Compliance checker for NetCDF-CF and ACDD attributes



Rosetta

Select Observation Platform

- ▶ Upload File
- ▶ Specify Header Lines
- ▶ Specify Delimiters
- ▶ Specify Variable Attributes
- ▶ Specify Site Specific Information
- ▶ Specify General Information
- ▶ Download Converted File

Next

Select Observation Platform



☐ Single Station or Tower (timeSeries)



☐ Moored Buoy (profile)



☐ Radiosonde (trajectory)



☐ Wind Profiler (profile)



☐ Aircraft (trajectory)



☐ Single CTD/XBT cast (profile)



Questions or comments about Rosetta can be sent to: support-rosetta@unidata.ucar.edu, or rosetta@nersc.no for this specific instance.

Version : 0.5.2-NERSC

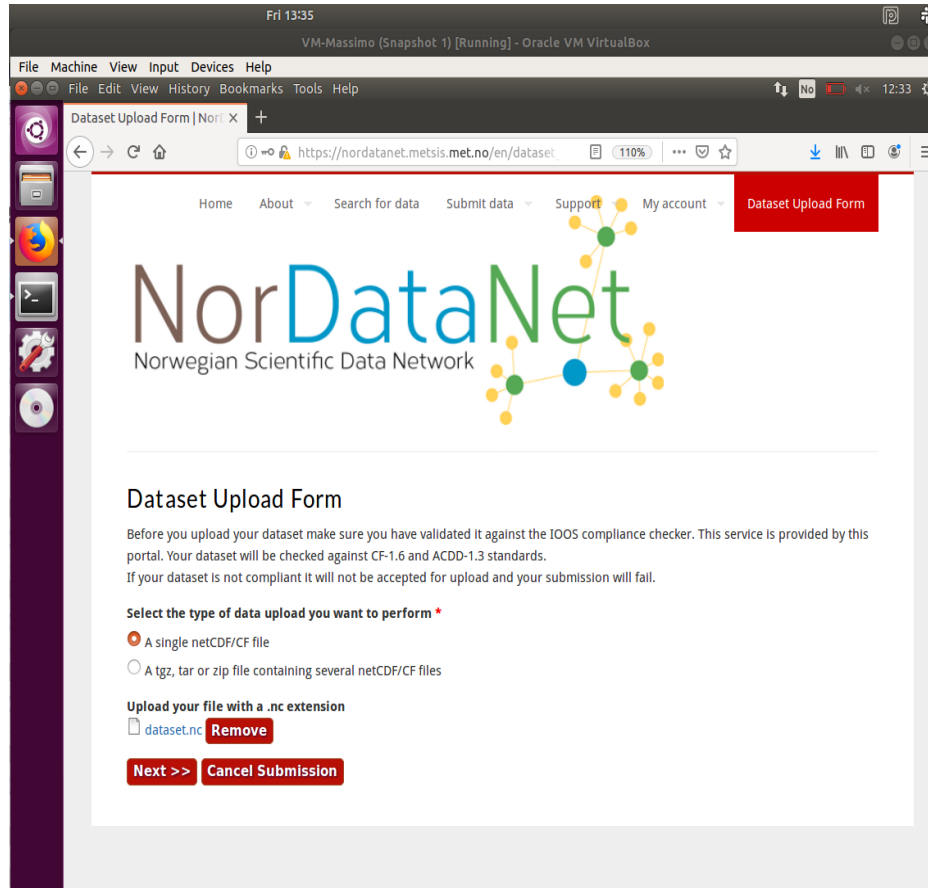
Build Date: Thu Feb 06 10:22:54 CET 2020



unidata

providing innovative data services and tools to transform the conduct of geoscience

Upload Interface to expose data from NIRD archive (1)



- The user has the possibility to upload a single dataset in NetCDF/CF format.
- A group of files (tgz/zip file) containing NetCDF/CF files.
- The NetCDF files must follow the CF/ACDD conventions in order to be able to extract the correct metadata from the dataset.
- Validation tools for NetCDF/CF are integrated in the work flow.

Upload Interface to expose data from NIRD archive (2)

VM-Massimo (Snapshot 1) [Running] - Oracle VM VirtualBox

Dataset Upload Form | NorDataNet - Mozilla Firefox

https://nordatnet.metsis.met.no/en/dataset_upload

Home About Search for data Submit data Support My account Dataset Upload Form

NorDataNet
Norwegian Scientific Data Network

✓ Your dataset dataset.nc is compliant with CF and ACDD standards. The submission can now proceed.

Dataset Upload Form

Your uploaded file has the metadata as reported in the following table. Please make sure they are correct before confirming your submission. If the metadata are not correct, cancel your submission, correct your information and proceed with a new submission.

Metadata for: "Dataset.nc"

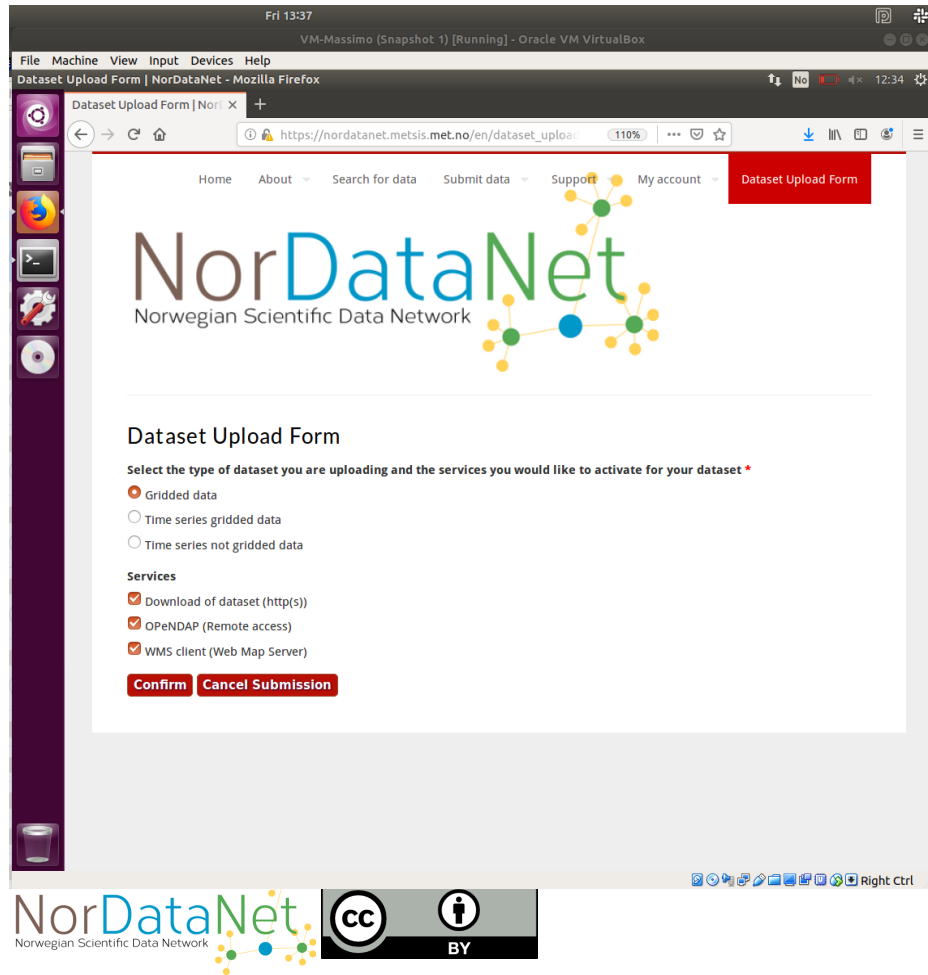
Metadata Key	Metadata value
metadata_identifier	Not available
title	Title of my Dataset
abstract	Observations of temperature in tall masts made by Kjeller Vindteknikk for the E39 bridge project in Western Norway
metadata_status	Active
dataset_production_status	Not available
collection	ADC
last_metadata_update	2019-10-02
temporal_extent_start_date	2019-10-01
temporal_extent_end_date	2019-10-09
iso_topic_category	Not available

NorDataNet
Norwegian Scientific Data Network

CC BY

- When submitting the dataset, it will be checked with respect to be CF/ACDD conventions.
- If these tests are passed, metadata are extracted and prepared to be sent to NIRD.
- The user can also check that the metadata extracted are correct and if not cancel the submission.

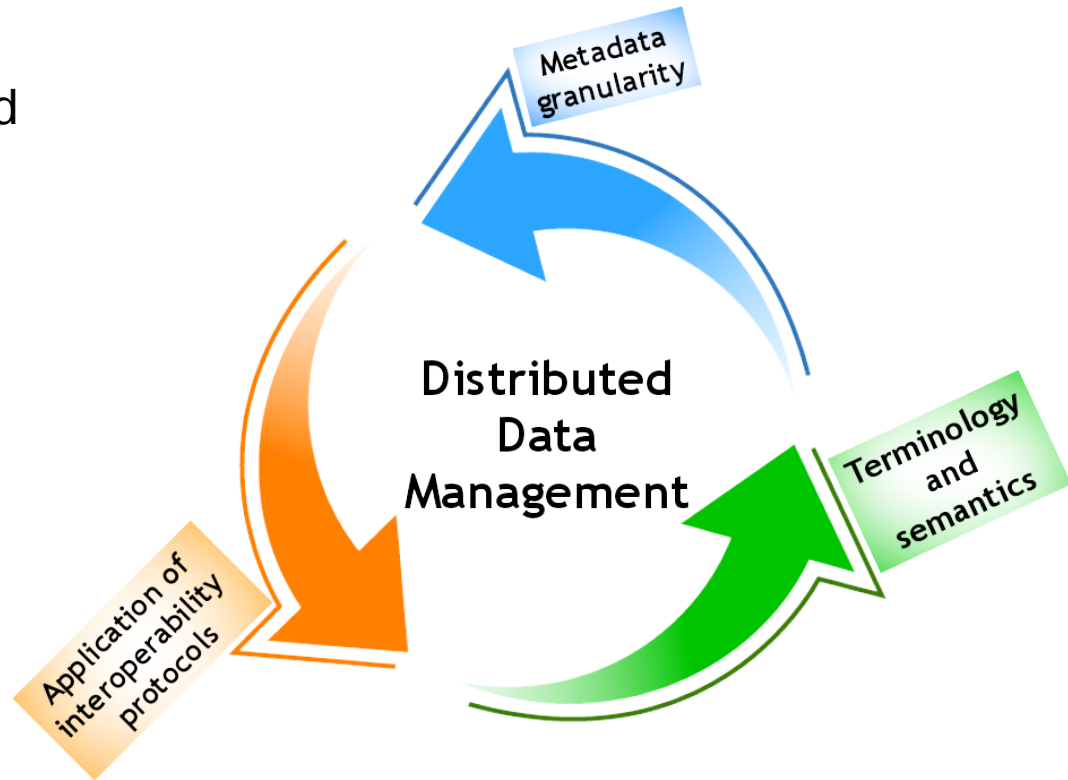
Upload Interface to expose data from NIRD archive (3)



- Depending on the type of data submitted the user can decide with type of services needed by NIRD to expose the dataset
- Data download (HTTP)
- OPeNDAP (remote access)
- OGC WMS (map visualization)

Bottlenecks/speed bums

- Interoperability at the data level
 - Need to engage data providers and data centres (varying degree of adherence to standards)
 - The easy solution is NetCDF-CF with ACDD served through OPeNDAP (and WMS)
 - But remember proper configuration
- Application of controlled vocabularies and proper identification of which controlled vocabularies



Status and plans

- Working integration at
 - Discovery level
 - MET, NMDC, NP, NERSC, (NILU, NIRD)
 - NSIDC, NIPR/ADS, CHINARE, PANGAEA, WGMS, PPDC, BAS, (PDC, CCADI, CNR, EUMETSAT, NINA)
 - Data level
 - MET, NERSC, (NMDC, NILU)
- New search/results interface under development
- Underlying information model has been upgraded, need to implement this in services
 - Interlinking with e.g. description of observation facilities
- Visualisation services are being reimplementation
- Basket service under reimplementation
- Need to finalise the integration with NIRD
- Checking MeteorIO as web service
- Machine readable exposure service for discovery metadata under reimplementation
- Further reading - <https://www.nordatanet.no/>