

A low-poly, blue-toned iceberg floats on a dark blue sea. The top of the iceberg is visible above the waterline, while the much larger, jagged base is submerged below. The sky is a light, hazy blue.

Data sharing ethics & culture, and how NorDataNet services help

Øystein Godøy



Outline

- Data sharing ethics, certainly before publishing
- Data Life Cycle and its relation to the scientific workflow, revisited from a scientists point of view
- Data sharing in a cultural perspective and relations to the scientific workflow
- NorDataNet service overview

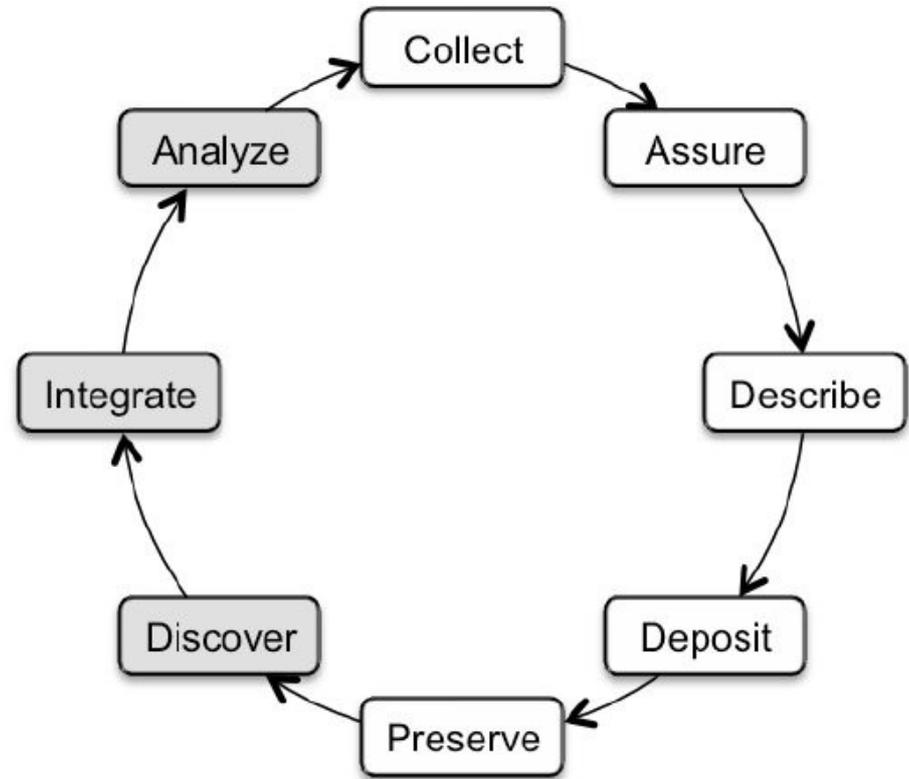
Data sharing ethics prior to publishing

- Identify the PI(s) of a dataset of interest
- Present your intentions to the PI(s) of a dataset of interest
- Make an agreement on how to use the data
- Acknowledge and respect agreements made
 - Shared data must only be used for the agreed purpose.
 - Use in other form than agreed upon has to be negotiated with the data PI(s).
 - If a publication or presentation is planned, clarify the mutual expectations on (co-)authorship.
- Never share unpublished data
- Do not present or publish without consent
- Publish according to international standard
 - ICMJ “Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals”.
- Acknowledge the project

Always cite data used!!

Data Life Cycle and its relation to the researchers workflow

- Generate or collect data
- Quality control and document data
- Analyse data
- Prepare and publish data through a mandated archive
 - Remember to get a DOI!!!
- Write a data paper
 - Cite your datasets using the provided DOI
- Write your scientific paper
 - Cite your data paper describing the context of the data



Data sharing in a cultural perspective and relations to the researchers workflow

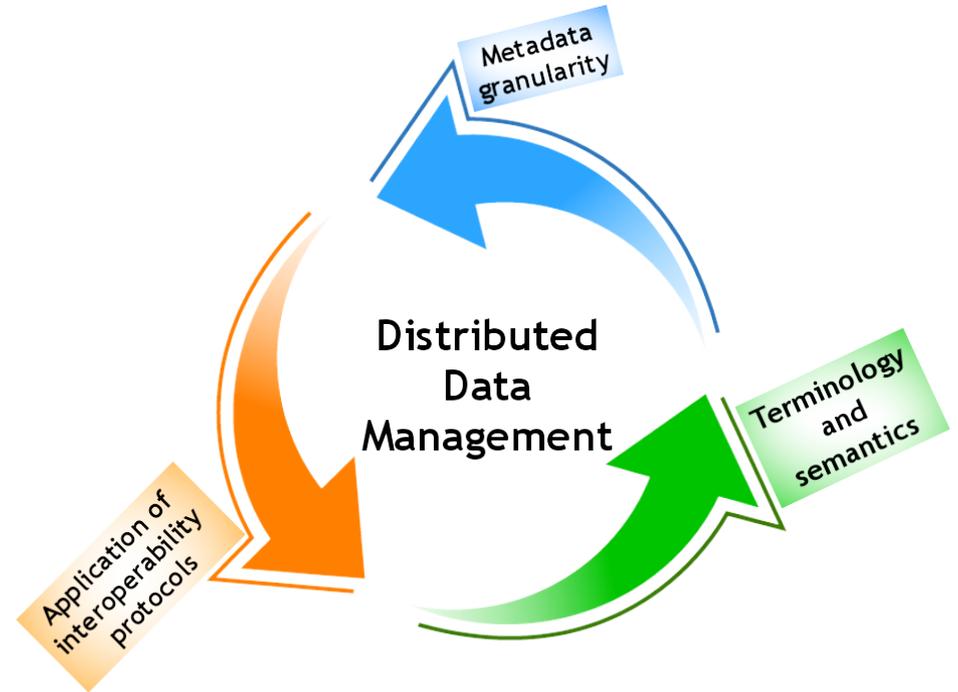
- Within some communities data sharing is the normal situation
 - In particular within remote sensing communities
- Other communities are more restrictive and consider data a personal item
 - In particular those “suffering” to get data
- Ownership to data are regulated through research grants contracts
 - Usually between the funding agency and the institution of the researcher
- While researchers often only use their own data, the intention of stakeholders (ministries and funding agencies) are to maximise the usage of data
 - Increased value for money
- Research programmes are increasingly focused on interdisciplinary science and societal benefit (impact) of science

Good data sharing practises

- Reduces the amount of duplicate data records
- Reduces the amount of duplicate field activities collecting data
- Reduces the environmental footprint of science
- Maintains the visibility of institutions and individuals
- Supports the intention of ministries and funding agencies

Bottlenecks

- Individual thresholds for sharing data are too high
 - Training, toolboxes and web services reducing the threshold is required
- Education lacks focus on data management
 - Equally important as measurement practise
 - Need to get this into the education
- Data managers and scientists has to work together, identifying gaps and closing them
 - Communication and willingness are essential
- Roles of scientists and data centres are changing, but mutually depending of each other



NorDataNet User Services

- Need to identify which services to offer to users (data providers and data consumers)
 - Simplifying their work
 - Showing the benefit of structured data management
 - Slow uptake in user communities
 - Need to interact better with user communities
 - Exploring interactions between research and operations
- Primary focus on Findable, Accessible, Interoperable, Reuseable
 - Data for services
- Modular approach
- Currently offering
 - A unified search interface
 - On the fly visualisation and transformation of data depending on FAIRness and protocols
 - Compliance checker for NetCDF-CF
 - Joint with SIOS and Nansen Legacy
 - Conversion tool for NetCDF-CF
 - Joint with SIOS and NMDC
- Ongoing work
 - Integration with NIRD
 - Joint with Nansen Legacy and SIOS InfraNOR
 - Improving visualisation and transformation
 - Improving data documentation support to data providers



METADATA Search

View Edit Track Translate Grant Node export

This is an inclusive search meaning that the search criteria selected will be combined to retrieve results. All menus are collapsed initially, they are opened by "clicking" them. Geographical bounding boxes may be selected using the map and fine tuned by opening the bounding box menu. Institutions contains a list of all institutions mentioned as PI institutions in the datasets.

Norsk
Full text search

Combine search keywords: e.g "radiosonde AND humidity"
bjorli

- Data Collection Period
- Bounding Box
- Institutions
- Collections
- Investigator
- Topics and Variables

Search Reset Basket (1)

Empty basket



SYNOP data from station BJORLI

Additional info

Title: SYNOP data from station BJORLI

Abstract: Synoptic meteorological measurements from BJORLI extracted from the WMO Global Telecommunication System (GTS). Data are not quality controlled after extraction from GTS.

Institutions: Norwegian Meteorological Institute

PI: Øystein Godøy

Additional Metadata

Data Access

Visualize

Download datasets as ASCII

File Edit View History Bookmarks Tools Help

CHES Course: An Intro x 20200512_NorDataN x Search results | NorData x NCO netCDF Operato x NASA GISS: Panoply x Submit data as NetCDF/...

https://www.nordatanet.no/en/results/?page=1

Dashboard Content Structure Appearance People Modules Configuration Reports Help Hello steingod Log out



SYNOP data from station BJORLI Additional Info

Title: SYNOP data from station BJORLI
Abstract: Synoptic meteorological measurements from BJORLI extracted from the WMO Global Telecommunication System (GTS). Data are not quality controlled after extraction from GTS.
Institutions: Norwegian Meteorological Institute
PI: Øystein Godøy

Additional Metadata Data Access Visualize Download dataset as ASCII

<input type="checkbox"/> Dataset name	Institutions	Abstract	Collection period
<input type="checkbox"/> SYNOP data from	Norwegian	Synoptic meteorological measurements from BJORLI extracted from the WMO Global	2013-01-01T00:00:00Z

File Edit View History Bookmarks Tools Help

CHES Course: An Intro x 20200512_NorDataN x Search results | NorData x NCO netCDF Operato x NASA GISS: Panoply x Submit data as NetCDF/...

https://www.nordatanet.no/en/results/?page=1

Dashboard Content Structure Appearance People Modules Configuration Reports Help Hello steingod Log out



Choose a variable

- TA
- TD
- TAX_12
- TAN_12
- TW
- PR
- PO
- PP
- AA
- RR_1
- RR_6
- RR_12
- RR_24
- SA
- EE
- FF
- DD
- FG_010
- FG

SYNOP data from station BJORLI Additional Info

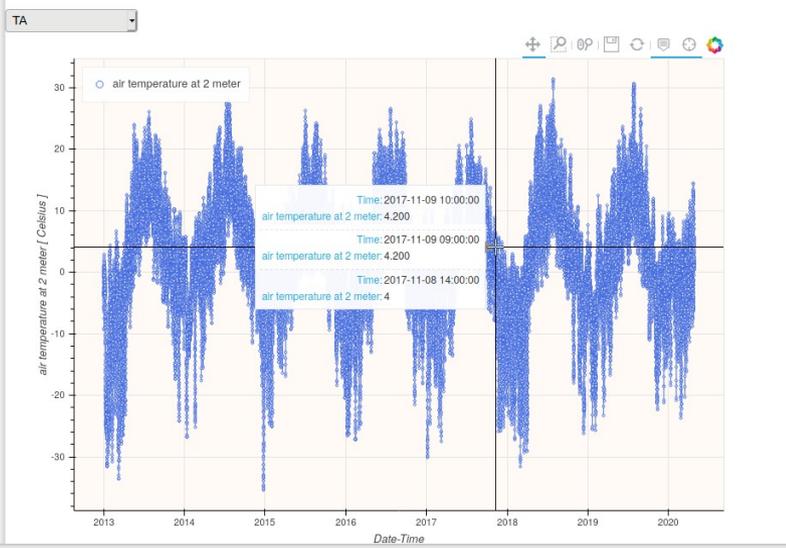
Title: SYNOP data from station BJORLI
Abstract: Synoptic meteorological measurements from BJORLI extracted from the WMO Global Telecommunication System (GTS). Data are not quality controlled after extraction from GTS.
Institutions: Norwegian Meteorological Institute

Data Access Visualize Download dataset as ASCII

<input type="checkbox"/> Dataset name	Institutions	Abstract	Collection period
<input type="checkbox"/> SYNOP data from	Norwegian	Synoptic meteorological measurements from BJORLI extracted from the WMO Global	2013-01-01T00:00:00Z

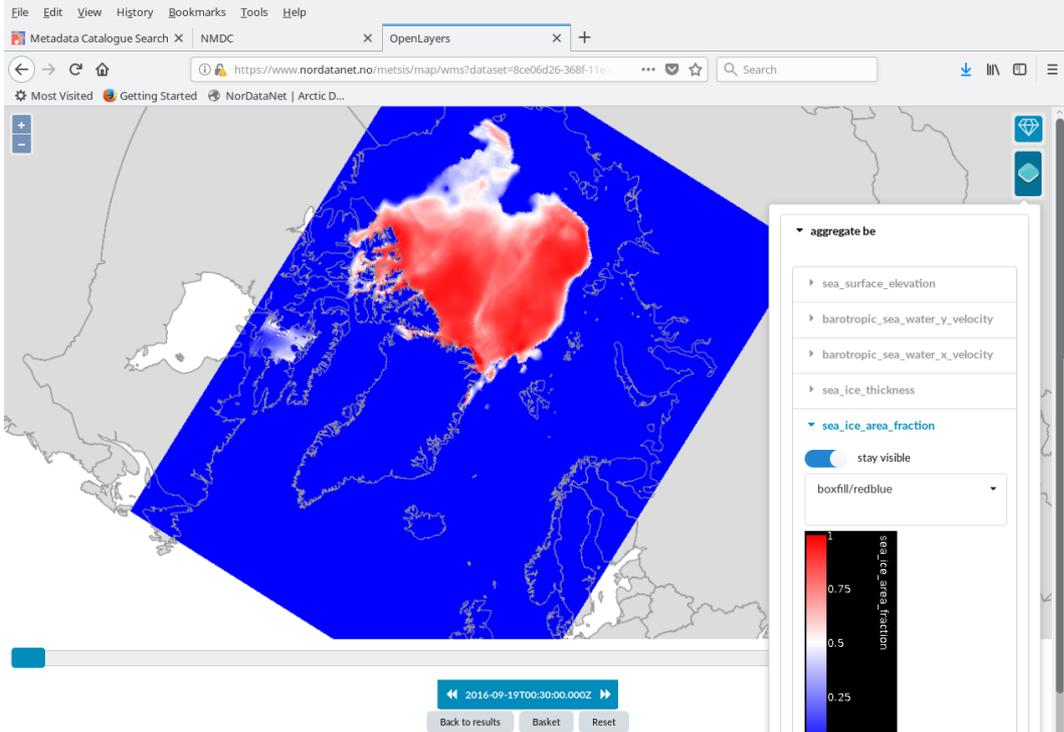
Title: SYNOP data from station BJORLI
 Abstract: Synoptic meteorological measurements from BJORLI extracted from the WMO Global Telecommunication System (GTS). Data are not quality controlled after extraction from GTS.
 Institutions: Norwegian Meteorological Institute
 PI: Øystein Godøy

Additional Metadata Data Access Visualize Download dataset as ASCII



	Variable
<input type="checkbox"/>	TA
<input type="checkbox"/>	TW
<input type="checkbox"/>	PR
<input type="checkbox"/>	PO
<input type="checkbox"/>	PP
<input type="checkbox"/>	AA
<input type="checkbox"/>	RR_1
<input type="checkbox"/>	RR_6
<input type="checkbox"/>	RR_12
<input type="checkbox"/>	RR_24
<input type="checkbox"/>	SA
<input type="checkbox"/>	EE
<input type="checkbox"/>	FF
<input type="checkbox"/>	DD
<input type="checkbox"/>	FG_010
<input type="checkbox"/>	FG
<input type="checkbox"/>	VV
<input type="checkbox"/>	NN
<input type="checkbox"/>	HL
<input type="checkbox"/>	CL
<input type="checkbox"/>	CM
<input type="checkbox"/>	CH
<input type="checkbox"/>	UU
<input type="checkbox"/>	WW
<input type="checkbox"/>	W1
<input type="checkbox"/>	W2
<input type="checkbox"/>	HW
<input type="checkbox"/>	PW
<input type="checkbox"/>	HWA
<input type="checkbox"/>	PWA

- CSV
- netcdf
- CSV



Select temporal extent

Start date

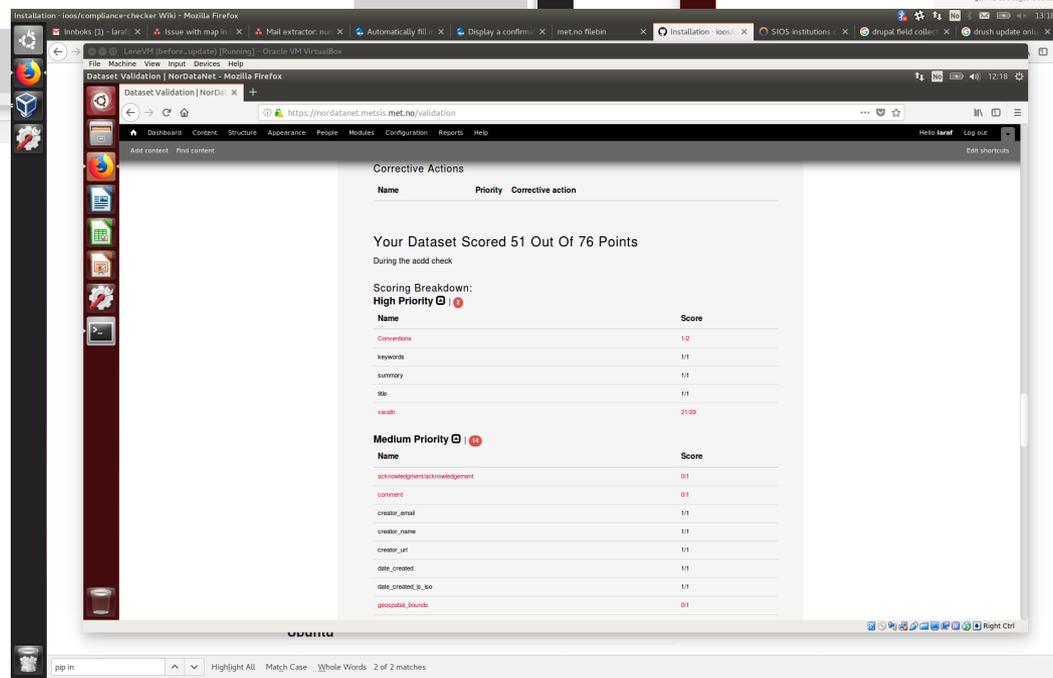
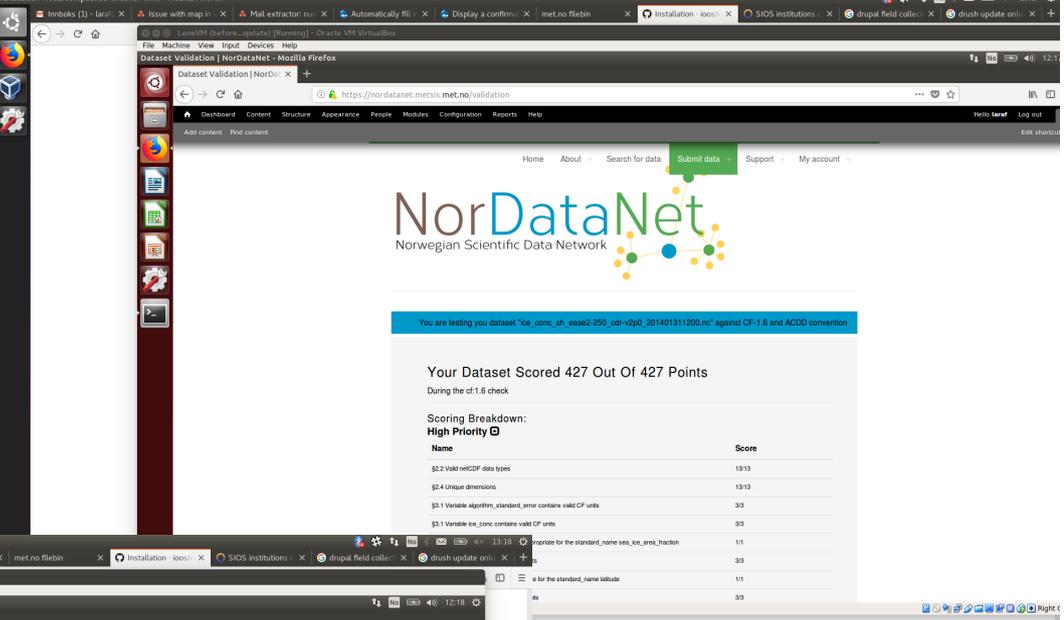
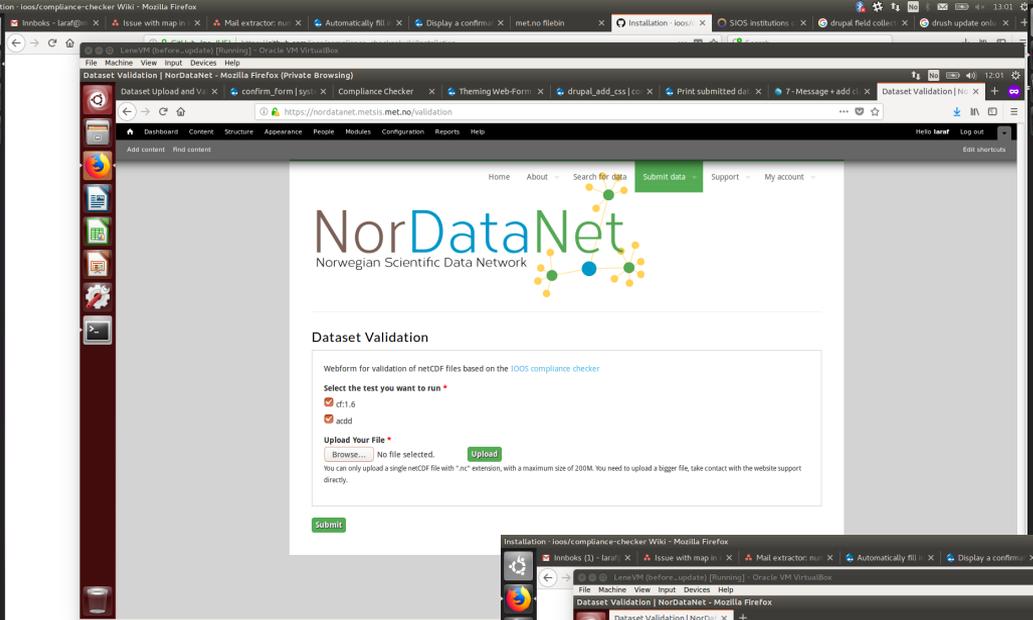
Stop date

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

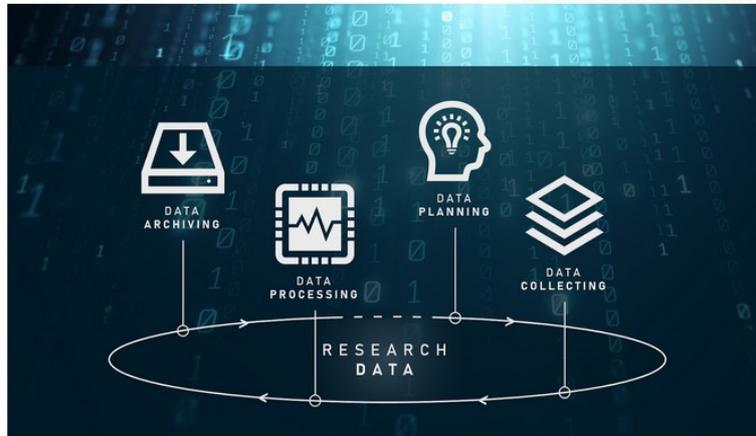
x-axis from:



About

Research Data

Scientific Research Data represent and provide huge resources for the scientific communities and for the innovation of the society in its whole. To facilitate the researching process, IT services must be available to support researchers at any stage of the research data life cycle, from the early stages of the project planning, throughout the entire project lifetime and beyond.



NIRD is the new **National e-Infrastructure for Research Data**, owned and operated by Uninett Sigma2. NIRD provides a flexible and secure facility for storing research data and offers software tools and applications to collect, store, process such data, high performance computing systems and on-demand computing resources to meet the researchers needs and enable them to tackle their scientific challenges.

Services

- Generic services
- Computational Services (HPC)
- Research Data
- Sensitive Data Services

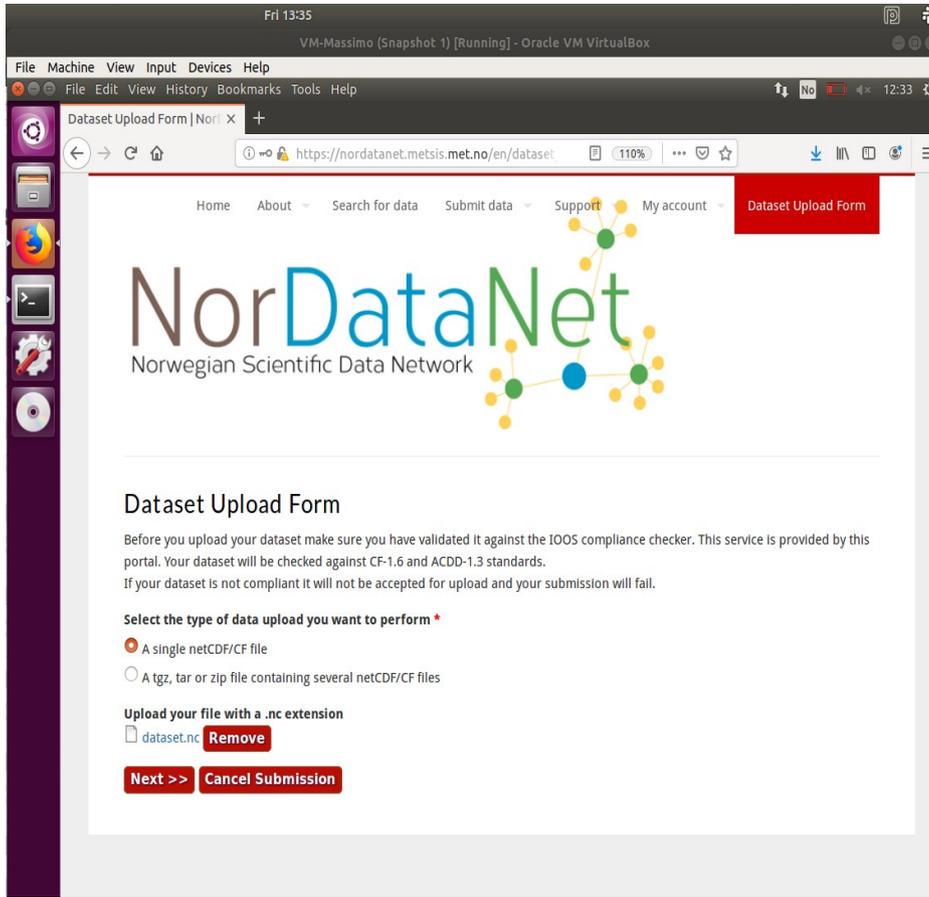
Shortcuts

- Hardware Live status
- Metacenter opslog
- Procurements
- Research Data Archive
- Training
- User documentation

Get access

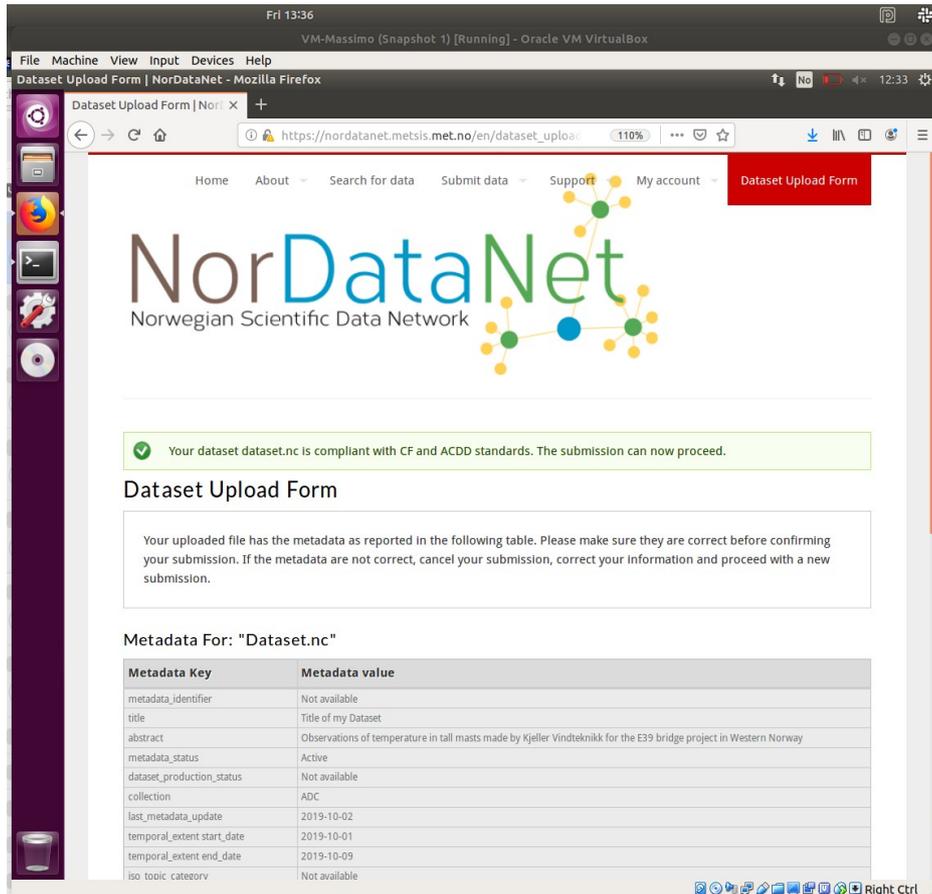
- Apply for HPC resources
- Apply for storage resources
- Apply for AUS resources
- Apply for user accounts

Upload Interface to expose data from NIRD archive (1)



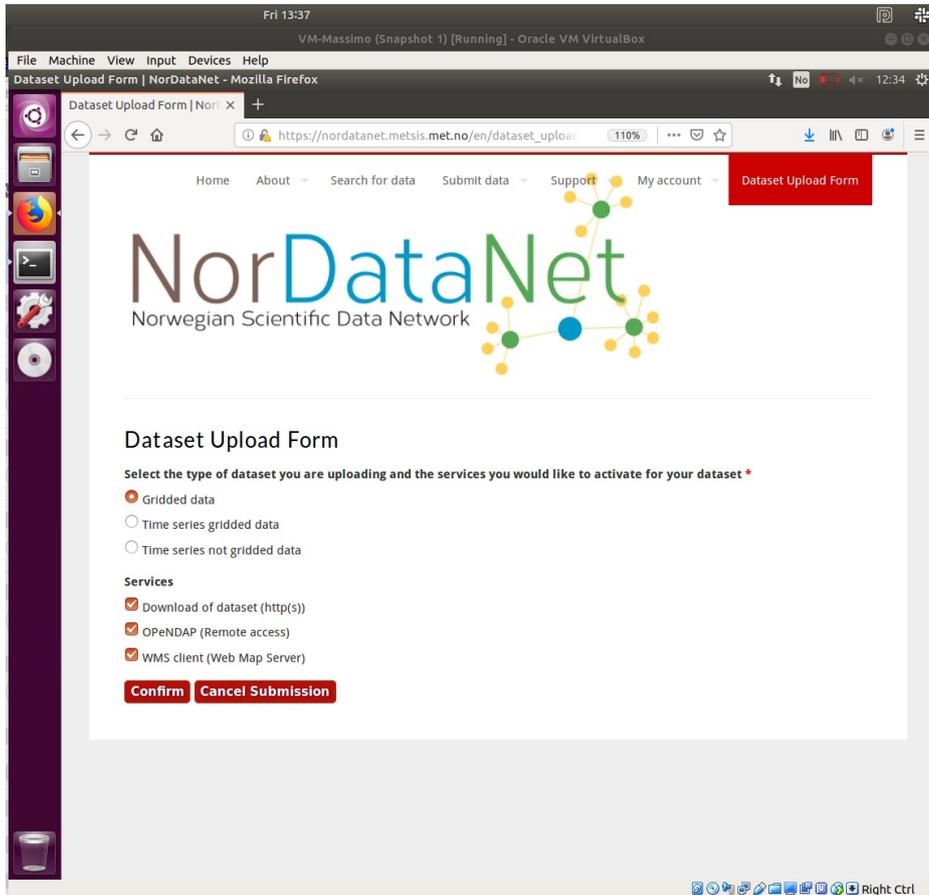
- The user has the possibility to upload a single dataset in NetCDF/CF format.
- A group of files (tgz/zip file) containing NetCDF/CF files.
- The NetCDF files must follow the CF/ACDD conventions in order to be able to extract the correct metadata from the dataset.
- Validation tools for NetCDF/CF are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

Upload Interface to expose data from NIRD archive (2)



- 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)