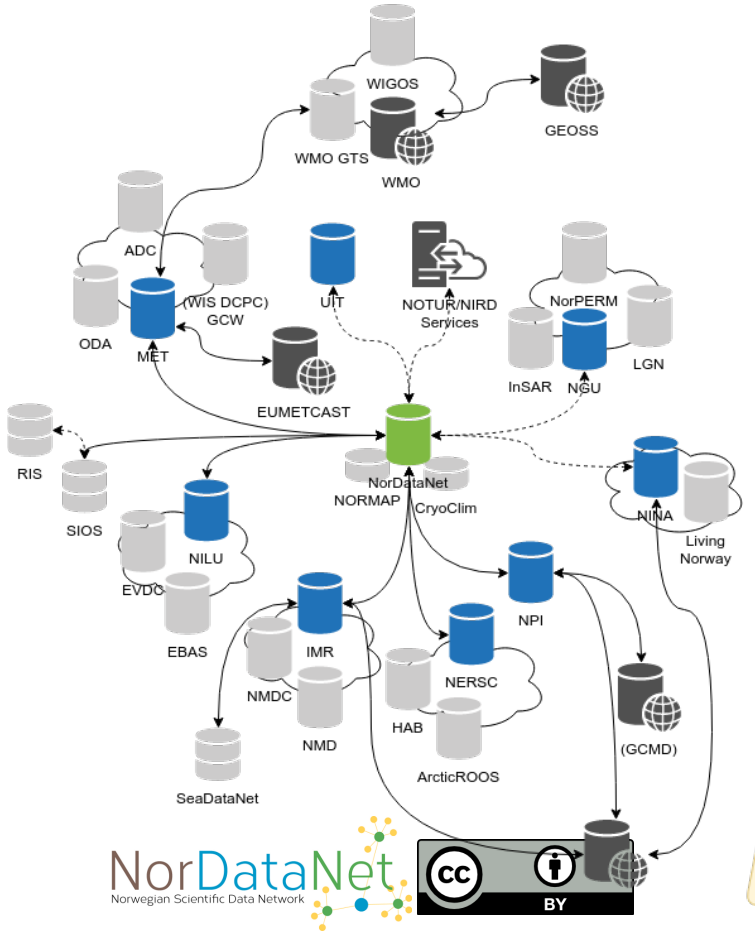
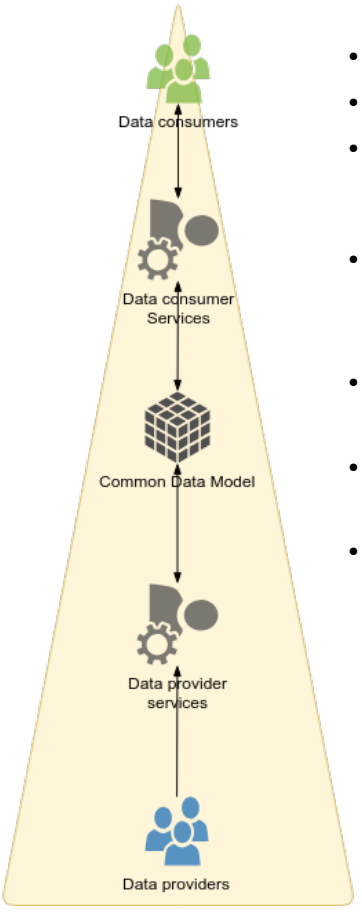
Norwegian Scientific Data Network
Services for providers and consumers

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

Norwegian Scientific Data Network - NorDataNet





- A distributed data management network
- Development funded by the Research Council of Norway (RCN)
- Activities are supporting other efforts like Svalbard Integrated Arctic Earth Observing System, WMO Global Cryospghere Watch, Year of Polar Prediction, ...
- Building on existing data centres nationally and the legacy of IPY
  - Originally focused on geoscientific data
  - But requested by RCN to think interdisciplinary
- Adding previous efforts like NORMAP and CryoClim as collections
  - Preparing integrations with the National Ground Segment for Satellite data
- Discovery metadata are harvested into a unified catalogue using
  - OAI-PMH, OGC CSW, (OpenSearch – only testing)
- Moving towards standardised data hosted by core partners
  - Actionable data as basis for user oriented services
  - Externally harvested information may be of any kind
  - Strong focus on NetCDF-CF wherever possible to achieve FAIR data

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

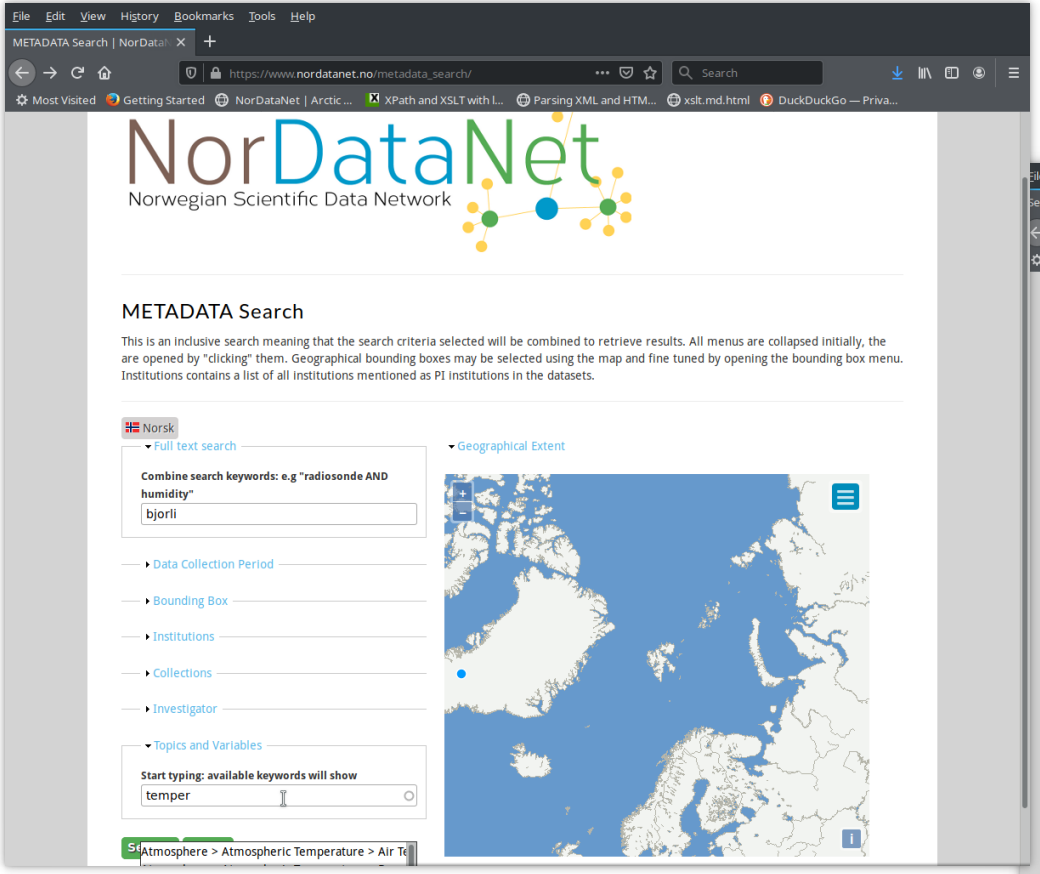
Overview

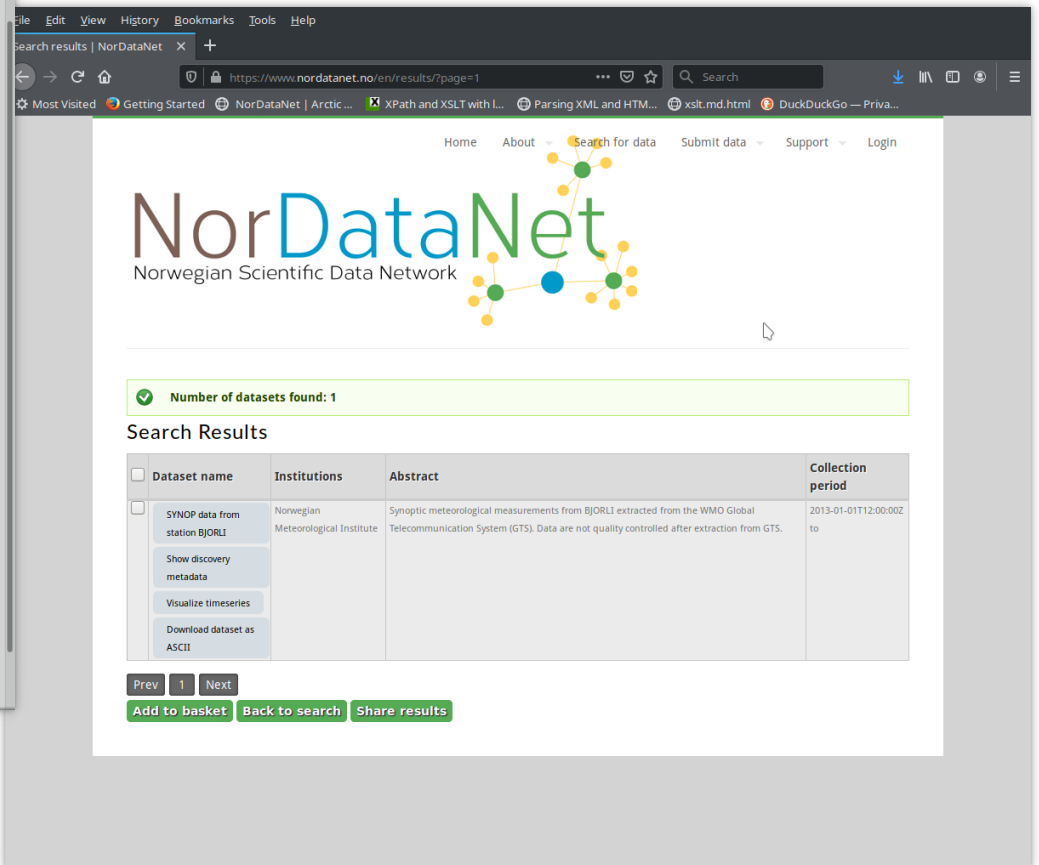
- Trying to identify which services to offer to users (data providers and data consumers)
  - Simplifying their work
  - Showing the benefit of structured data management
  - But slow uptake in user communities
    - Need to interact better with user communities
  - Exploring interactions between research and operational services
- Primary focus on Findable, Accessible, Interoperable, Reuseable
  - Data for services, easy integration in VREs
- Modular approach
  - Ideally connecting functional components with OGC WPS

- Currently offering
  - A unified search interface
  - On the fly visualisation and transformation of data depending on FAIRness and protocols
    - Using OPeNDAP, OGC WMS and OGC WPS
  - 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 Norwegian Infrastructure for Research Data
    - Joint with Nansen Legacy and SIOS InfraNOR
  - Improving visualisation and transformation of NIRD hosted datasets
  - Improving data documentation support to data providers
  - Improving visualisation services

NorDataNet
Norwegian Scientific Data Network

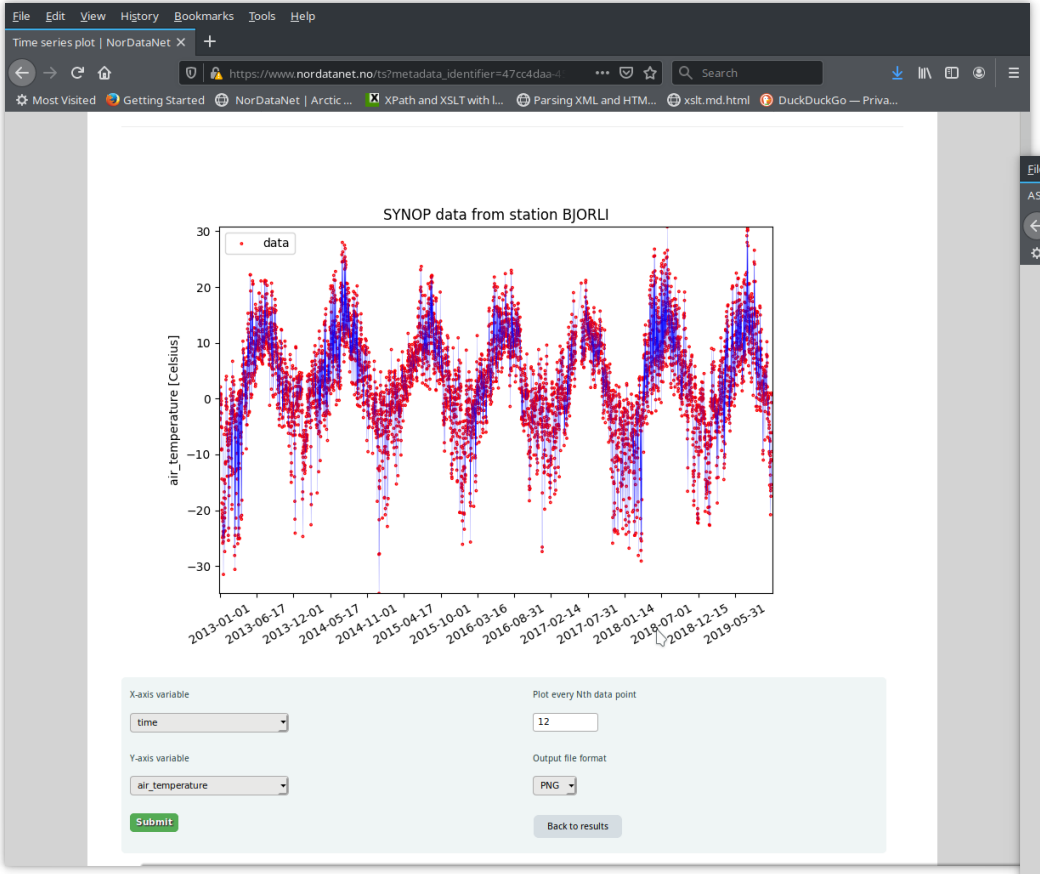
Screenshot examples of existing functionality - search

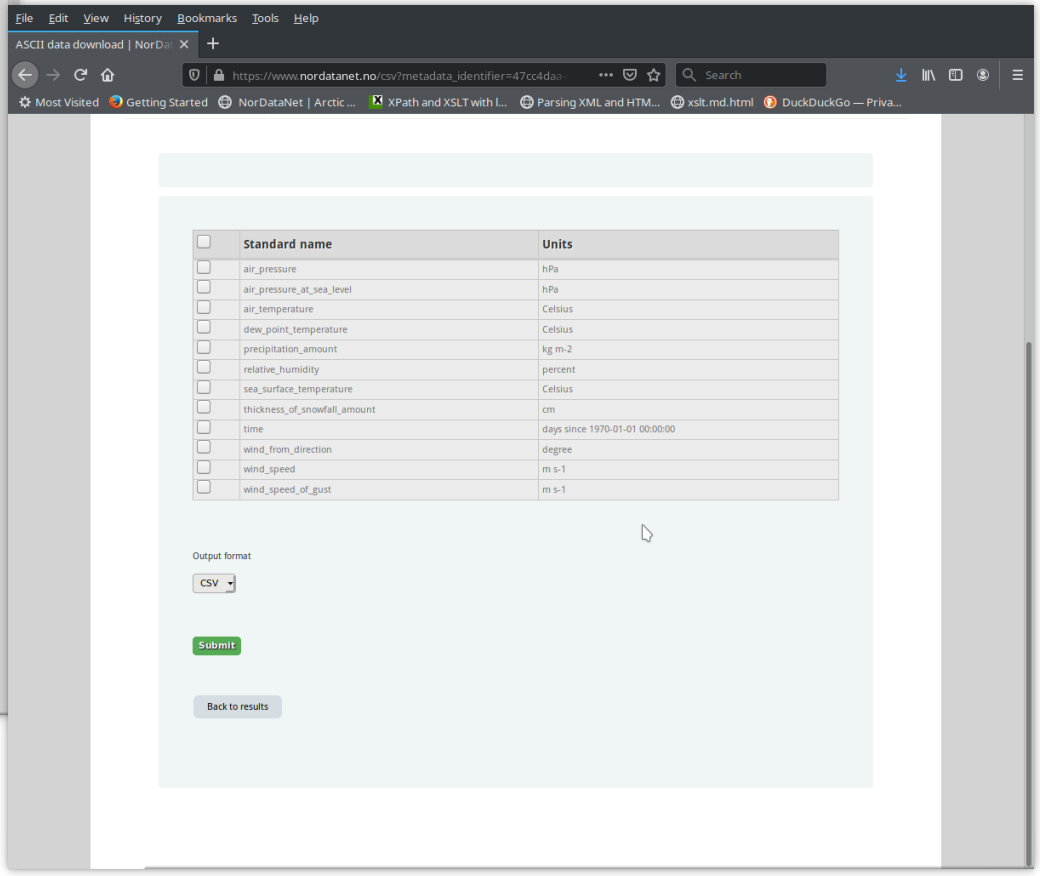




NorDataNet
Norwegian Scientific Data Network

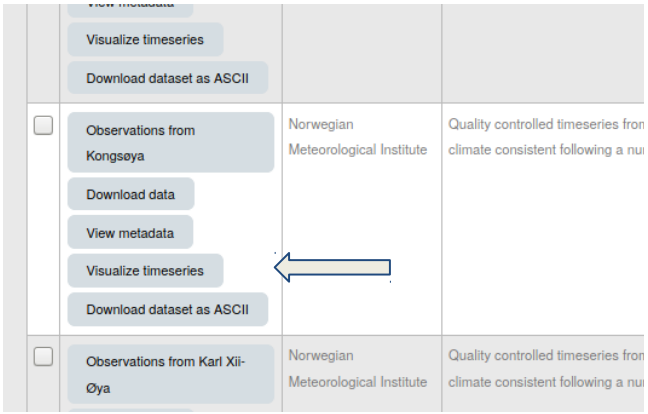
Visualisation and transformation of timeseries



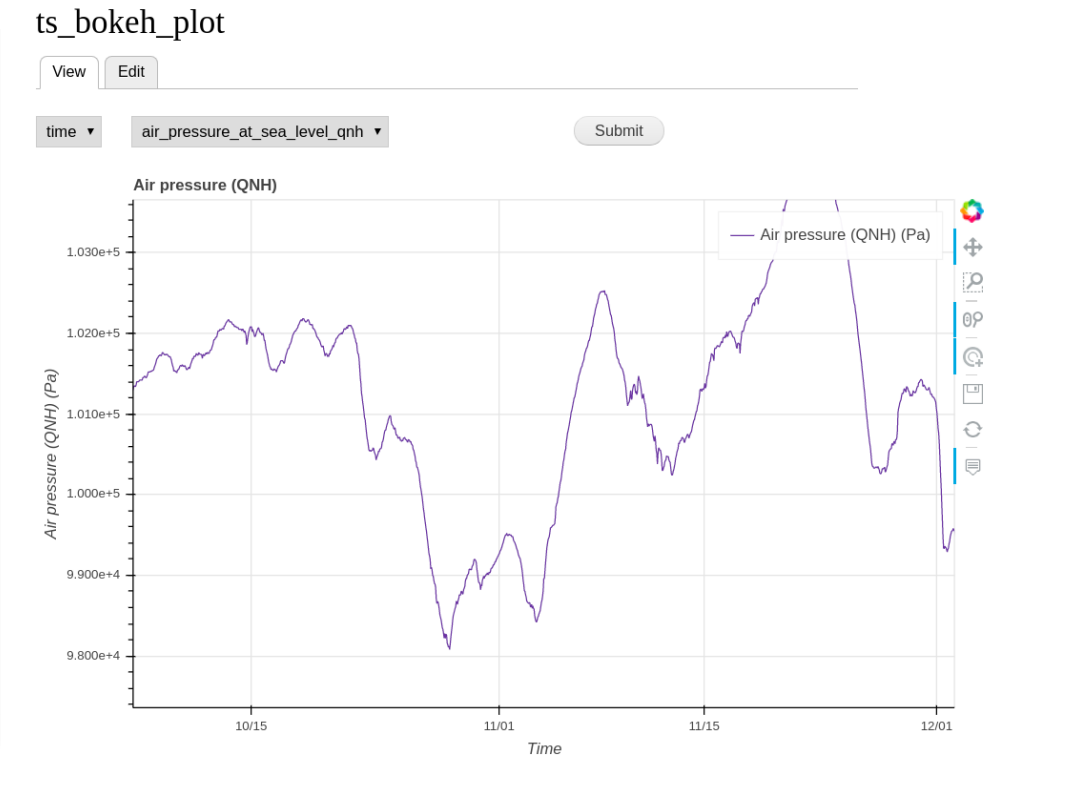


NorDataNet
Norwegian Scientific Data Network

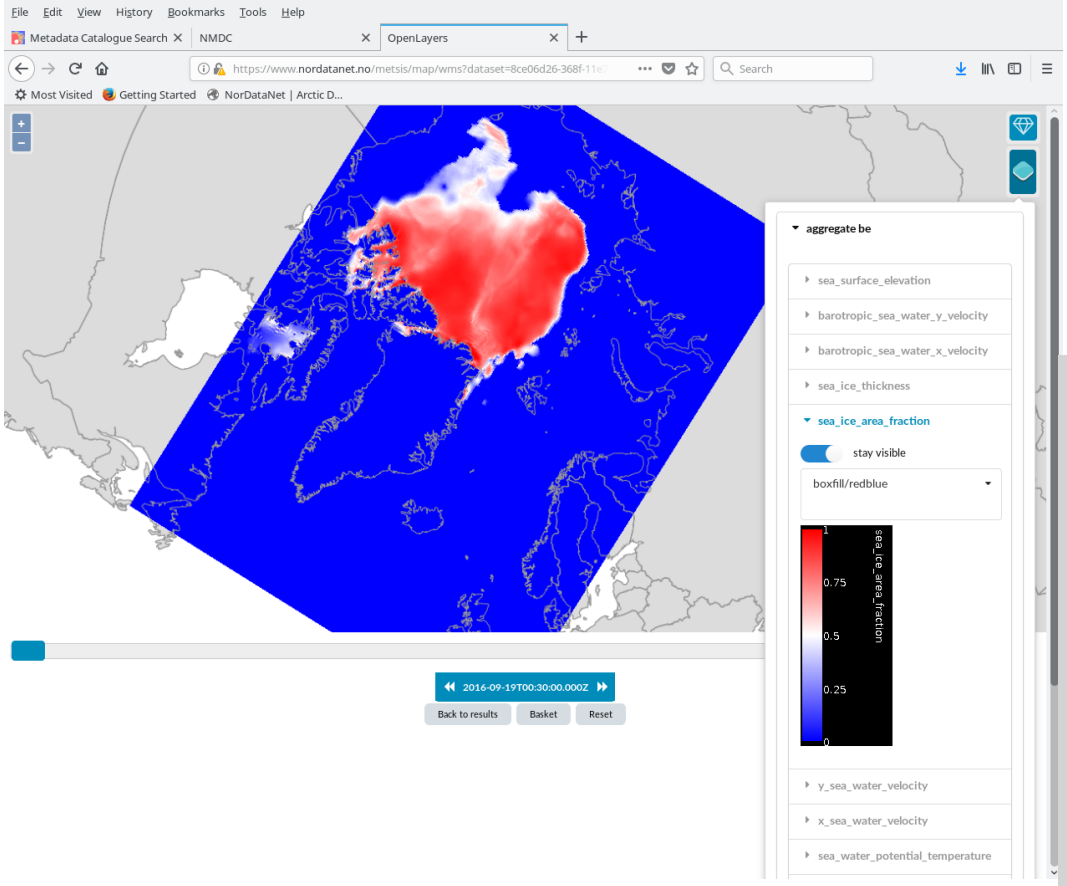
Time Series plotting – new dynamic presentation

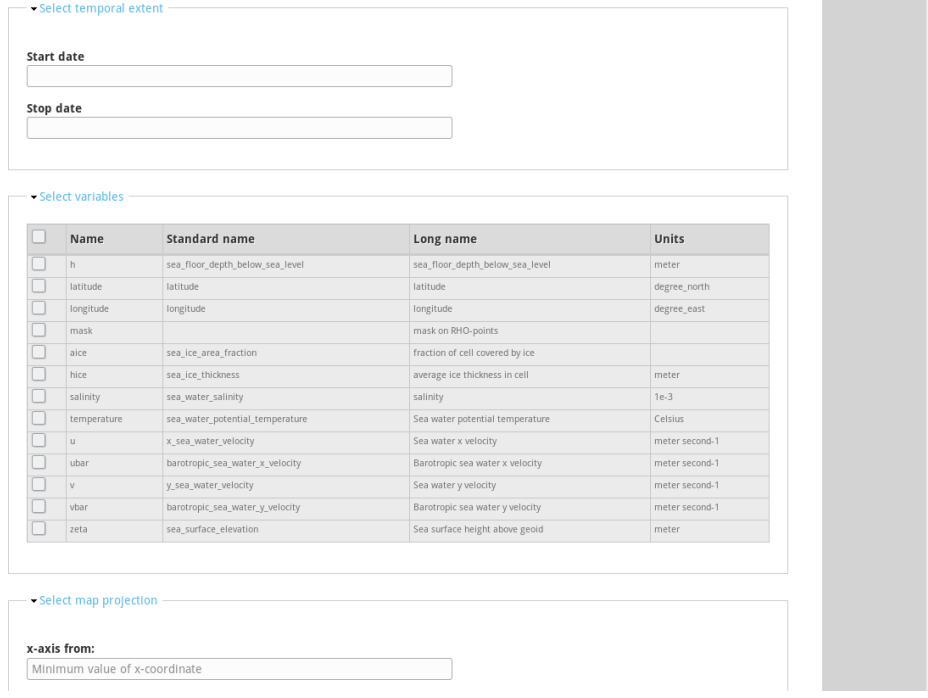


Upgrade and extension: Work in progress!

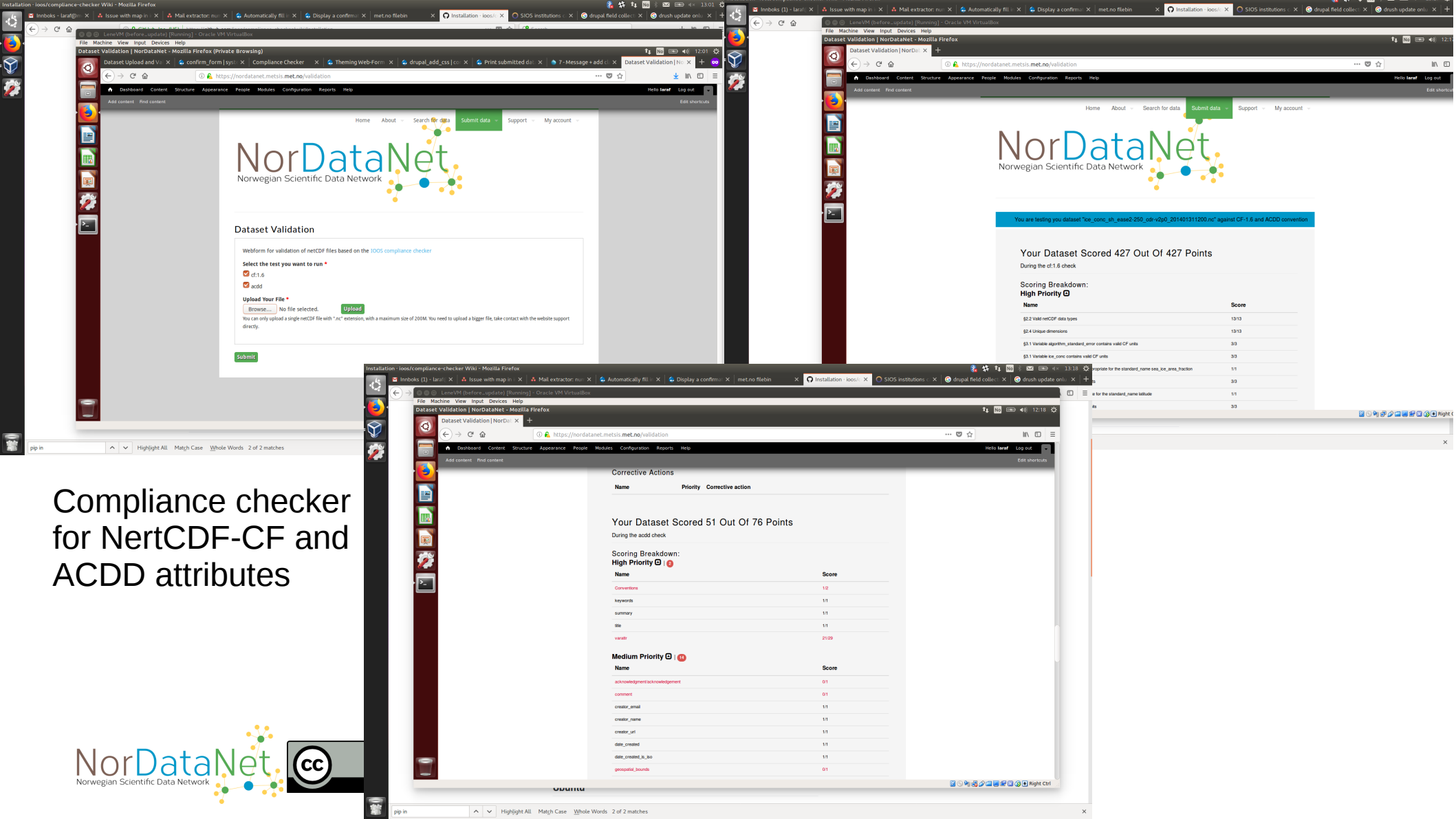


Visualisation and transformation of gridded products



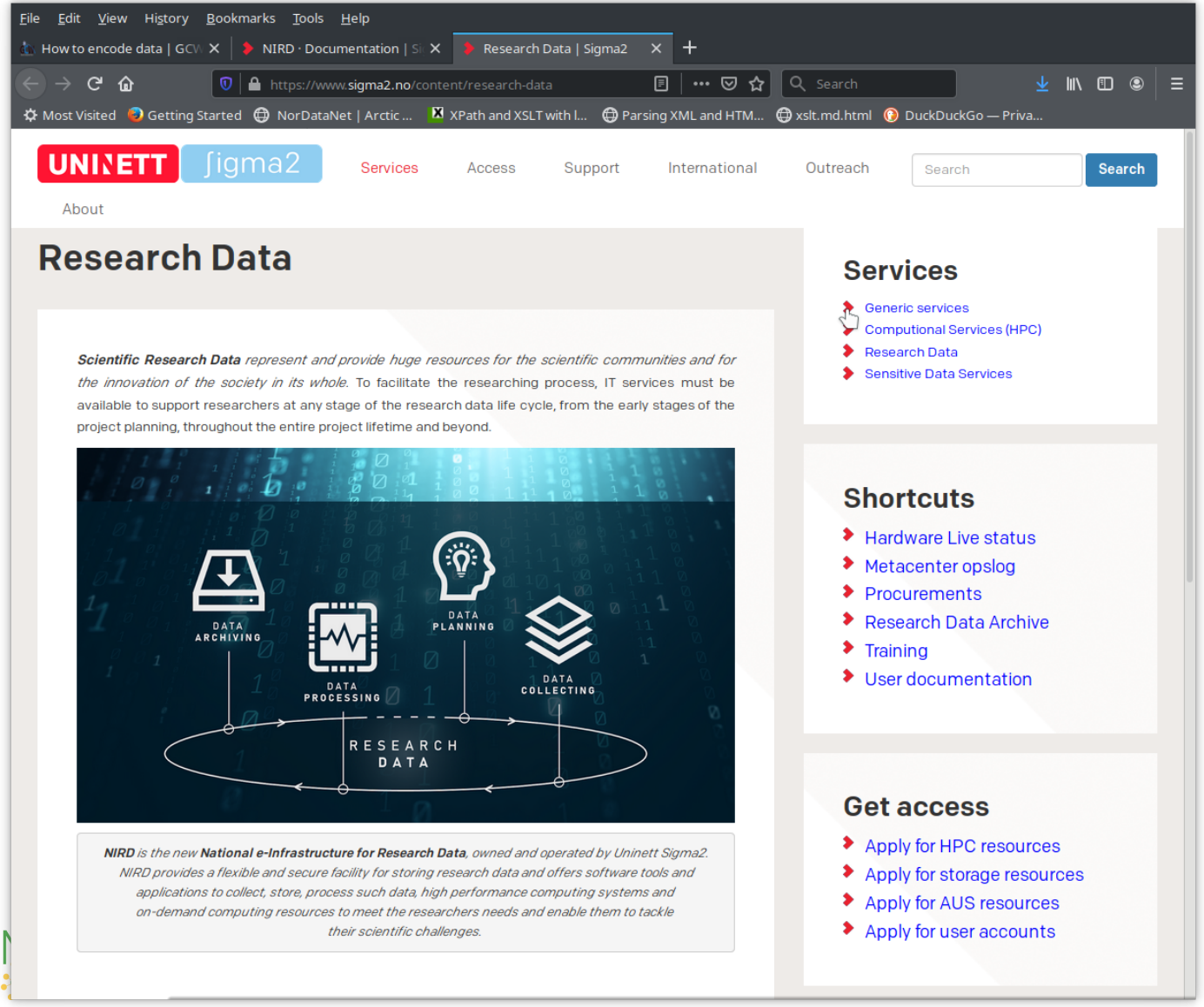


NorDataNet
Norwegian Scientific Data Network



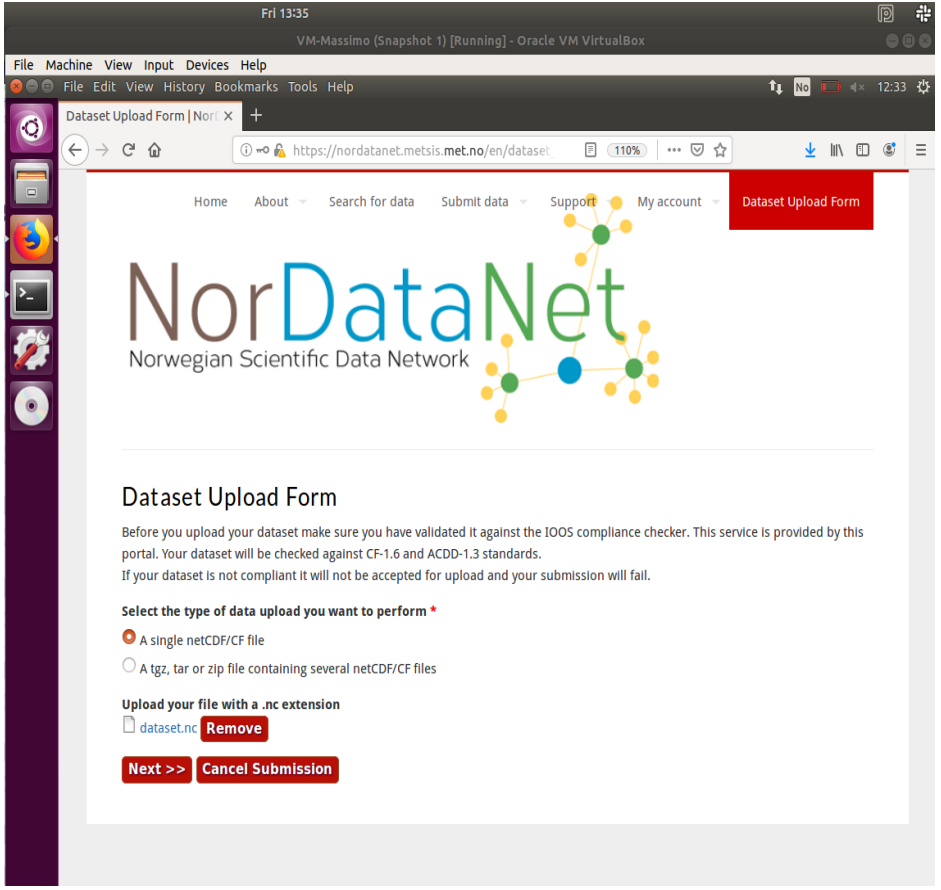
Compliance checker for NertCDF-CF and ACDD attributes

NorDataNet
Norwegian Scientific Data Network



Integration with national storage resources, using a service oriented approach

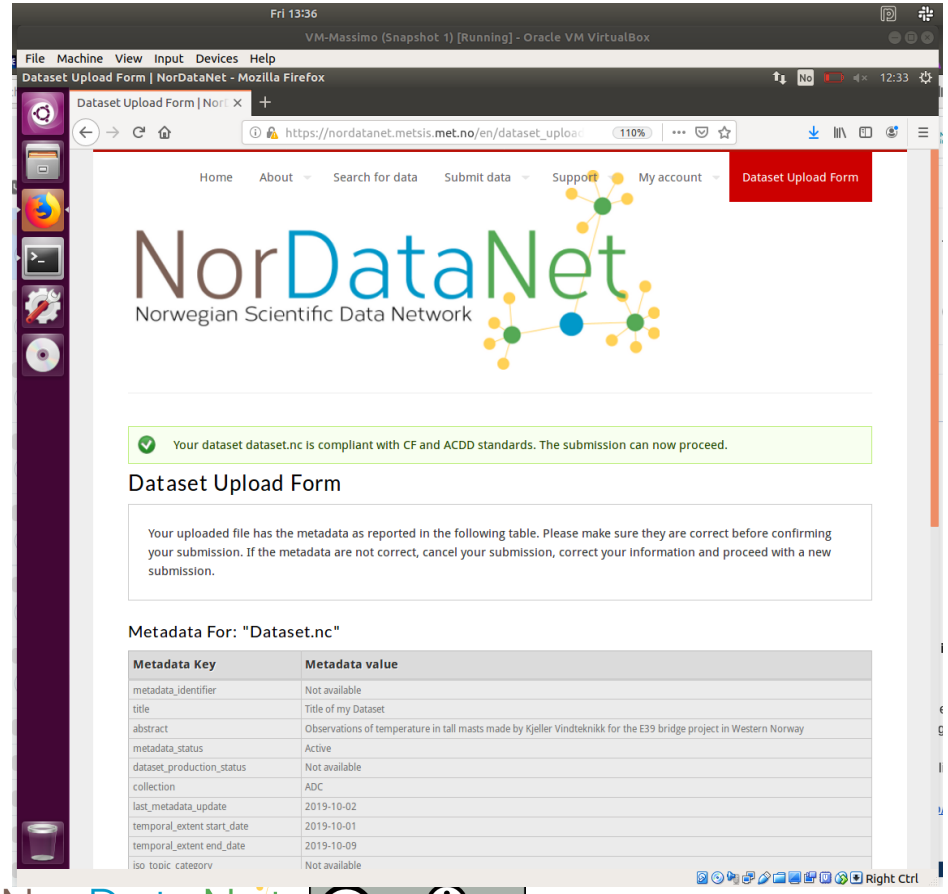
Upload Interface to expose data from NIRD archive (1)



- 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.

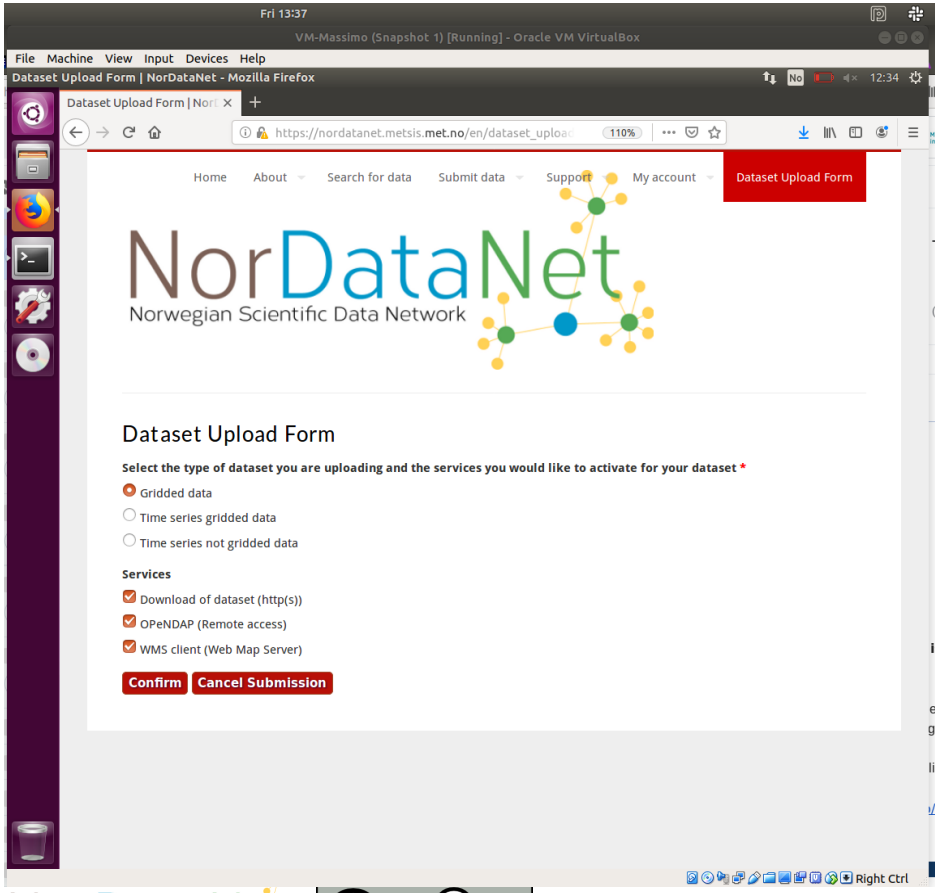
NorDataNet
Norwegian Scientific Data Network

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.

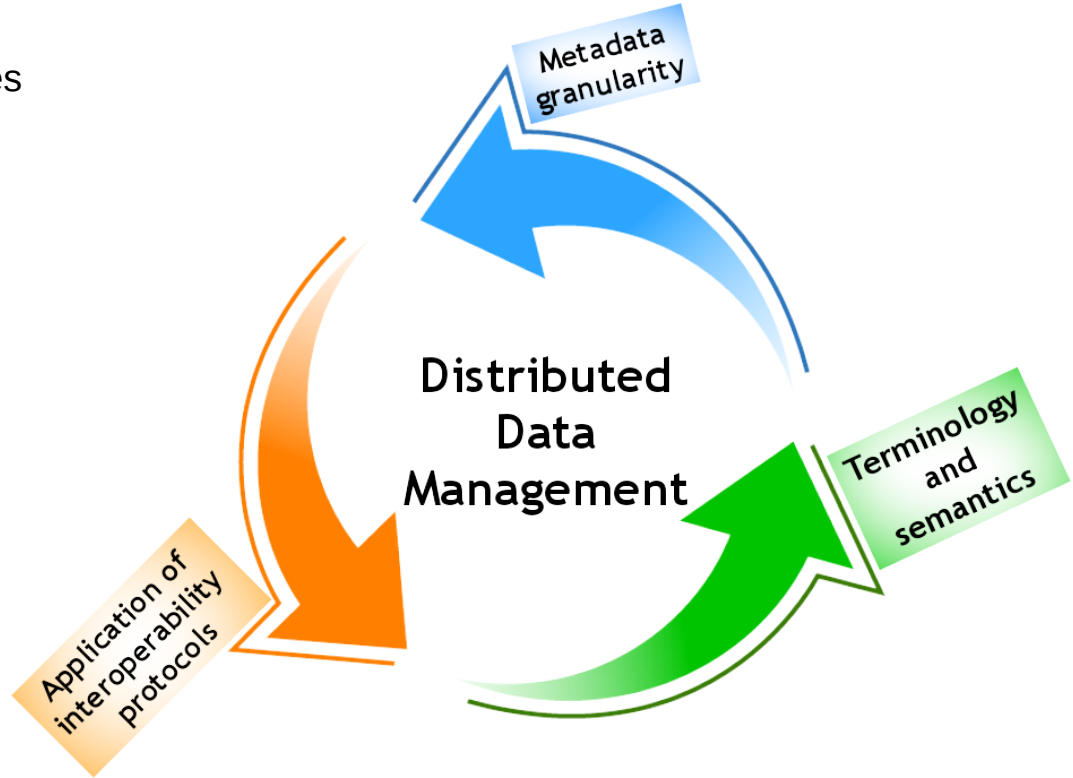
Upload Interface to expose data from NIRD archive (3)



- 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)

Bottlenecks

- Interoperability at the data level
  - Need to engage data providers and data centres (varying degree of adherence to standards)
  - Need tools helping them
- Application of controlled vocabularies and proper identification of which controlled vocabularies
- Interaction with data providers in academia
  - Understanding emerging requirements from funding agencies on free and open data
- The Catch 22 of users not finding, then not submitting data resulting in no data found
- Provision of useful tools for the user community – hard to get feedback



NorDataNet
Norwegian Scientific Data Network