



## Arctic Data Centre

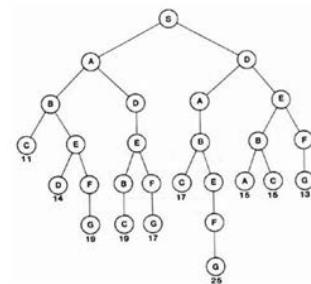
Bridging scientific and operational data streams

Øystein Godøy and Bard Saadatnejad



**network**

From <http://n-e-r-v-o-u-s.com/blog/?p=26>

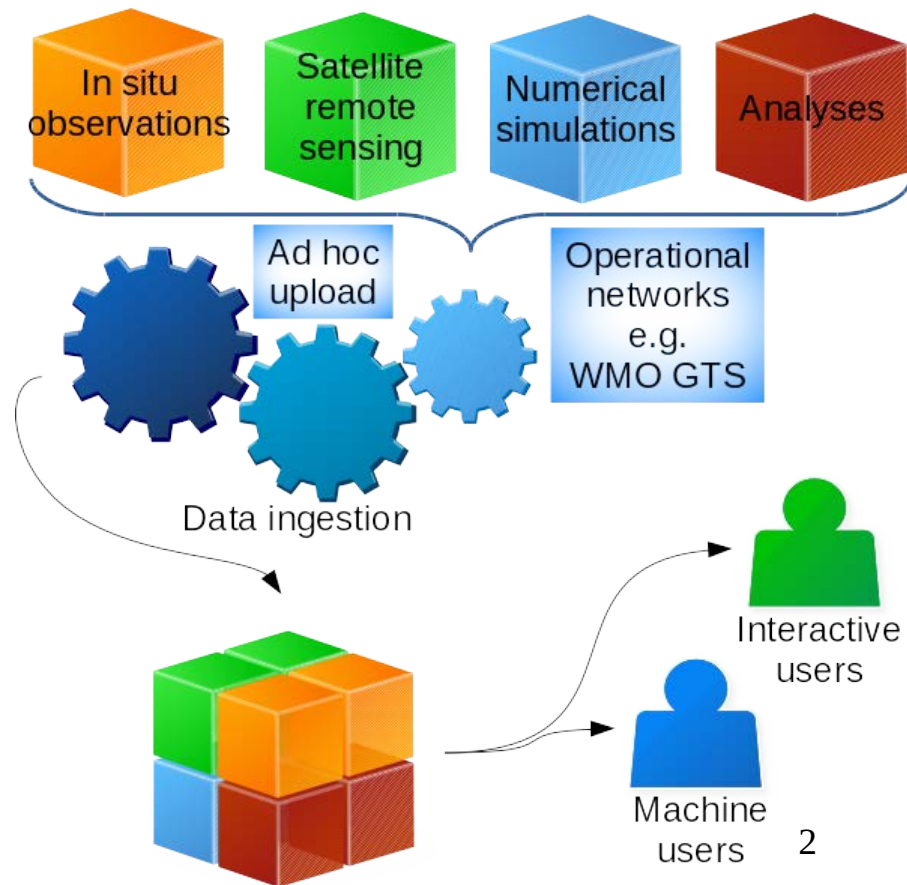


**hierarchy**



# Background

- A legacy of IPY
  - International Operational Data Coordination through IPYCOORD
  - National Data Coordination through DOKIPY
- WMO Information System
  - Data Collection and Production Centre





# Supporting projects

- WMO
  - Global Cryosphere Watch
  - Year of Polar Prediction
- Research Council of Norway
  - Norwegian Satellite Earth Observation Database for Marine and Polar Research (NORMAP)
  - Norwegian Marine Data Centre (NMDC)
  - Norwegian Scientific Data Network (NorDataNet)
  - Svalbard Integrated Arctic Earth Observing System (SIOS KC)
  - Pilot project for geoscientific data management (GeoAccessNO)
- EU FP6/FP7/H2020
  - DAMOCLES
  - ACCESS
  - APPLICATE
  - INTERACT
- EUMETSAT
  - Ocean and Sea Ice SAF
- Copernicus
  - Copernicus Marine Environmental Monitoring Service
- ESA/Norwegian Space Centre
  - CryoClim
  - National Ground Segment for Copernicus data



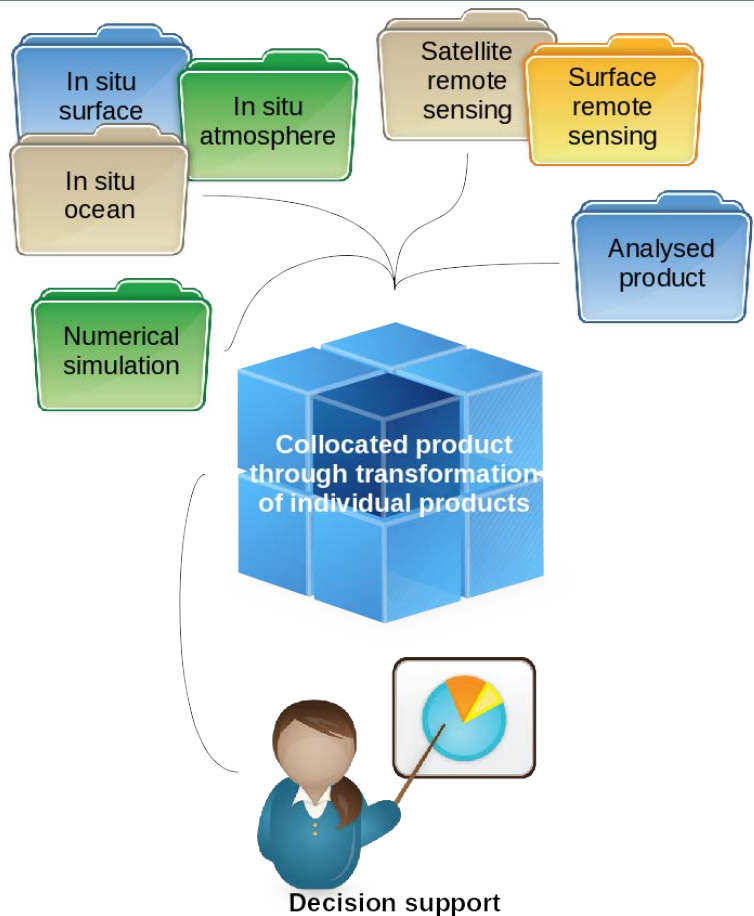
# Relation between ADC and projects

- ADC is the long term data management commitment focusing on both scientific and operational data streams of the Norwegian Meteorological Institute
- All data hosted in project specific portals will be included in ADC
- Functionality targeting the scientific user community is normally developed through dedicated projects





# Approach (1)



- Dataset oriented
  - Metadata driven
- Open data space
  - Higher order services offered when the data space can be constrained
- Net centric
  - Linkages with other data centres is vital
  - Implies brokering of metadata and data
- Interdisciplinary
  - Dataset agnostic in the open data space



# Approach (2)

- Technical development through METNO Scientific Information System (METSIS)
- Not a necessarily a distribution system
- Rather a toolbox
  - To set up distribution services for projects and the general public
  - Easily configurable frontend to credit various projects
  - Sharing backend elements for cost effective implementation, operation and maintenance
- Utilisation of open source software where possible
- Releasing software developed as open source code (GPL/LGPL)

The screenshot shows the Arctic Data Centre website interface. The header includes the Arctic Data Centre logo and the text 'Hosted by the Norwegian Meteorological Institute'. Below the header, there are navigation links for 'Browse data repository', 'ISO23950/SRU Search', and 'Metadata Catalogue Search'. The main content area displays search results for 'ROMS-Arctic20km-analysis'. A table lists the dataset name, institutions, areas, source, and a detailed summary. The summary describes the model's resolution, data coverage, and its application in emergency models.

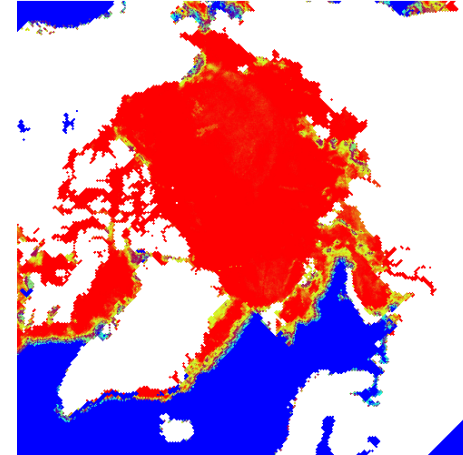
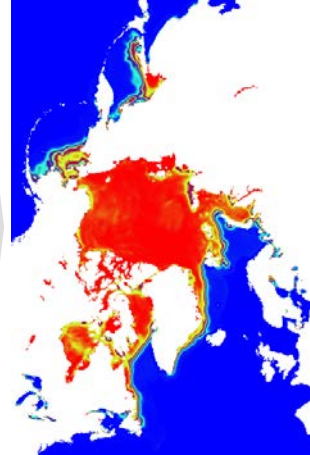
Dataset name	Institutions	Areas	Source	Summary
ROMS-Arctic20km-analysis	Norwegian Meteorological Institute (met.no)	Nordic Seas	Model run	This ocean model is operated at 20km resolution covering the Nordic Seas and the Arctic Ocean. This specific dataset provides the daily analysis from the operational model. Only the analysis is provided for historical periods, the daily forecast with 1 hour resolution is provided as a separate dataset. Currently the WMS presentation of this dataset is not supporting the 3D nature. A numerical model is applied to describe the dynamics of the oceans, such as sea level variations (tides and storm surge), movements in the water column (currents) and the salinity and temperature. To simulate the ocean, a 3-D grid is applied with different sizes, i.e., small grids for fine scale or detailed calculations, and larger or coarser grids to cover larger areas (and depth). The model runs on a supercomputer, and provides forecasts of sea level, currents, salinity and temperature for a time-range between 66 (2.75 days) and 240 hours (10 days). The model is run operationally, i.e. in a "24/7/365" environment to provide a 99.5% stability on a yearly basis. Currents from the model is further applied in emergency-models that simulates pathways of oil slicks and drifting objects (Search And Rescue). The ocean model used is the Regional Ocean Modeling System (ROMS). This is a three-dimensional, free-surface, terrain-following numerical model that solve the Reynolds-averaged Navier-Stokes equations using the hydrostatic and Boussinesq assumptions (Haidvogel et al., 2008). Haidvogel, D.B., H. Arango, W.P. Budgell, B.D. Cornuelle, E. Curchiser, E. Di Lorenzo, K. Fennel, W.R. Geyer, A. J. Hermann, L. Lanerolle, J. Levin, J. C. McWilliams, A. J. Miller, A. M. Moore, T. M. Powell, A. F. Shchepetkin, C. R. Sherwood, R. P. Signell, J. C. Warner, and J. Wilkin, Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System, JOURNAL OF COMPUTATIONAL PHYSICS, 227, 3695–3624, 2008.
ROMS-Arctic20km-forecast	Norwegian Meteorological Institute	Nordic Seas	Model run	This ocean model is operated at 20km resolution covering the Nordic Seas and the Arctic Ocean. This specific dataset provides the hourly forecast fields from the operational model. For historical purposes, the daily analysis is provided as



Transformations allow users to do comparisons of products and to extract tailored products for their specific need

**Search results**

Dataset name	Institutions	Area	Operational status	Abstract
seaice-01n	NORwegian Meteorological Institute	Arctic Ocean	Operational	Monthly sea ice concentration estimated from satellite data within the framework of EUMETSAT Ocean and Sea Ice SAF
seaice-02n	NORwegian Meteorological Institute	Arctic Ocean	Operational	Monthly aggregated sea ice concentration in the Arctic Ocean derived with a homogeneity algorithm from successive satellite data (OSIRIS, AMSR-E)
seaice-03n	NORwegian Meteorological Institute	Arctic Ocean	Operational	Monthly aggregated sea ice concentration in the Arctic Ocean derived with a homogeneity algorithm from successive satellite data (OSIRIS, AMSR-E)

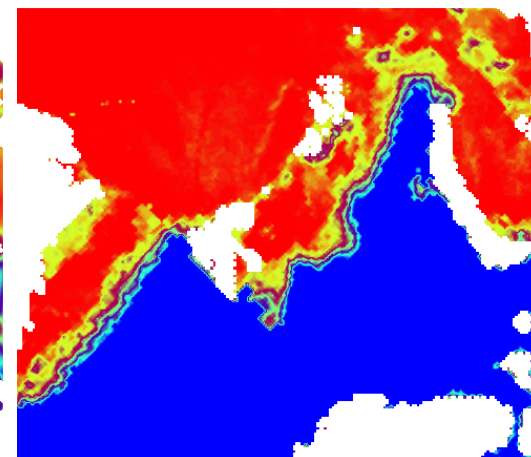
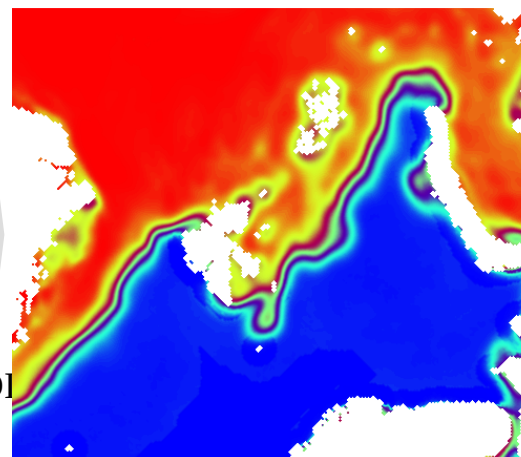


**Transformation request**

**Area**  
Specify geographical area of desired output data (in degrees, relative to zero meridian/equator):  
 Select region:

**Time**  
Specify time period of desired output data:  
 Select time range: from  to

Download	Name	Description	Standard name	Long name
<input type="checkbox"/>	Sea_ice			
<input type="checkbox"/>	Sea	longitude	longitude coordinate	lon
<input type="checkbox"/>	Sea	latitude	latitude coordinate	lat
<input type="checkbox"/>	Sea	time	time	t
<input type="checkbox"/>	Sea	time	time	t





# User interaction

- Primarily through supporting projects
  - Presentations at general assemblies through science oriented EU projects
  - User workshops through national e-infrastructure projects
- User oriented articles in the national e-infrastructure magazine META
- Questionnaires
- Polls and issue tracking for feature requests
- Much the same people responding regardless of method
- Increasing awareness in the scientific community as funding agencies are starting to require data deposition





# Lessons learned

- Flexible and adaptable user interfaces are vital
  - to establish a good dialogue with users
- Basket/Cart is required
  - to move from open to closed data space enabling higher order services
- Data streams instead of file transfers reduce complexity and cost
  - Reducing housekeeping requirements
  - Especially important for data brokering
- Interdisciplinary translation of documentation structure is “easy”
  - but translation of content is difficult
- Machine readable controlled vocabularies are needed to bridge communities
  - Blocking issues are more cultural than technical
- The requirement for use metadata is underestimated
  - Standardised data documentation as in NetCDF/CF and conveyed through OPeNDAP is highly valuable for automated interaction between data centres
- Configuration metadata are required in distributed data management
  - Should preferably follow index metadata



# Future plans

- Currently Developing a new human frontend
  - Standard CMS using Drupal 7
    - functionality as web services
  - More flexible to serve user requirements and editorial content
- Currently also modularising software
  - To ensure flexibility and sustainability when responding to user requirements
- Using OGC WPS as communication layer
- Work flow management
  - Linked to HPC services
- Extended semantic translations
- Efficient utilisation of DOI in KPI reporting
- Integration with the national high performance data archive
  - To support interdisciplinary use and simplify integration with national HPC services