

Norwegian Marine Data Centre



Seamless access to marine data

Terry Hannant
terry.hannant@imr.no



Project overview



- Main objective:
 - Serve the marine science community with seamless access to marine datasets covering waters of Norwegian interests.
- Website: <http://nmdc.no/>
- Coordinator:
 - Institute of Marine Research
- Partners:
 - 16 national partners



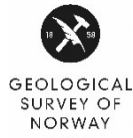
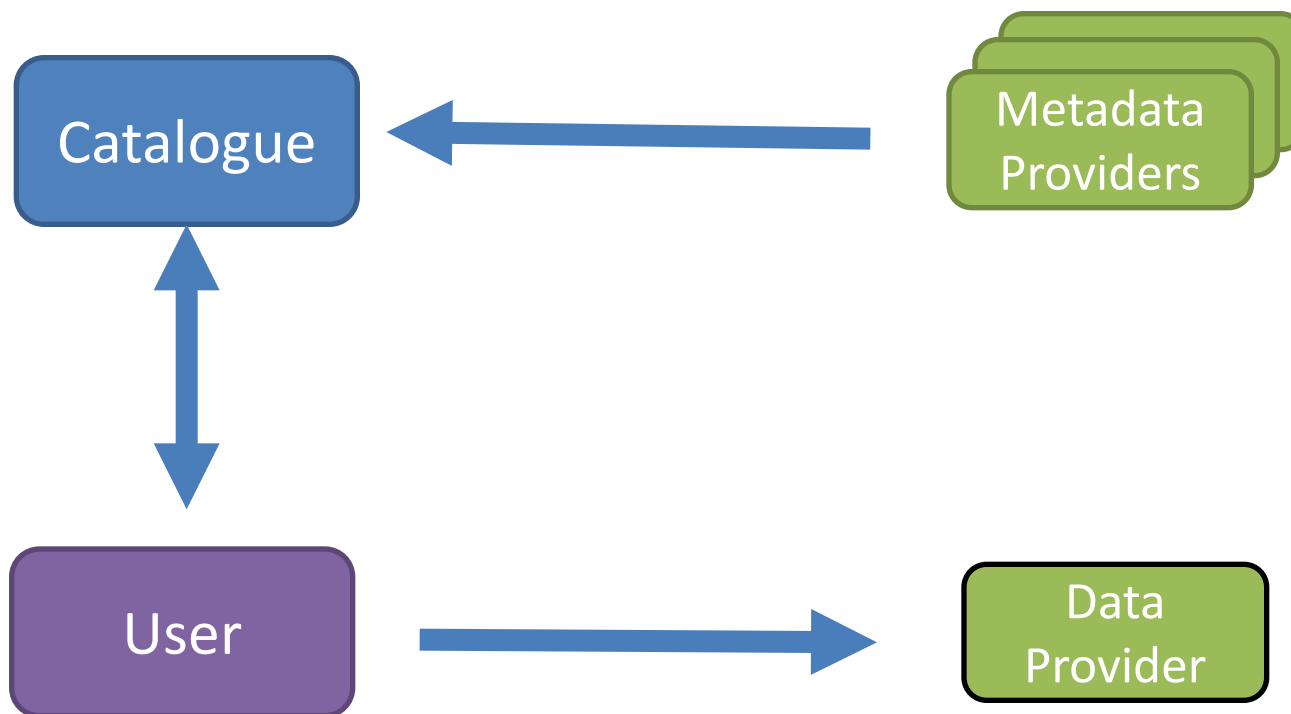
Project overview



- 8 Work packages:
- Project administration
 - System design
 - Data format and flow
 - Data storage
 - Data catalogue
 - Applications
 - Data availability
 - Dissemination and Sustainability



Project overview



Components



- Centralized metadata catalogue
 - All metadata are searchable from the same place
- Distributed metadata providers
 - Each partner provides metadata through their own infrastructure or via another partners infrastructure
- Distributed data providers
 - Each partner can use its own infrastructure



Components



- Public REST api for 3rd party users
 - Accessible on the central node
- Data catalogue web application
 - Shows all the metadata in the centralized catalogue
 - Possible to search and download data



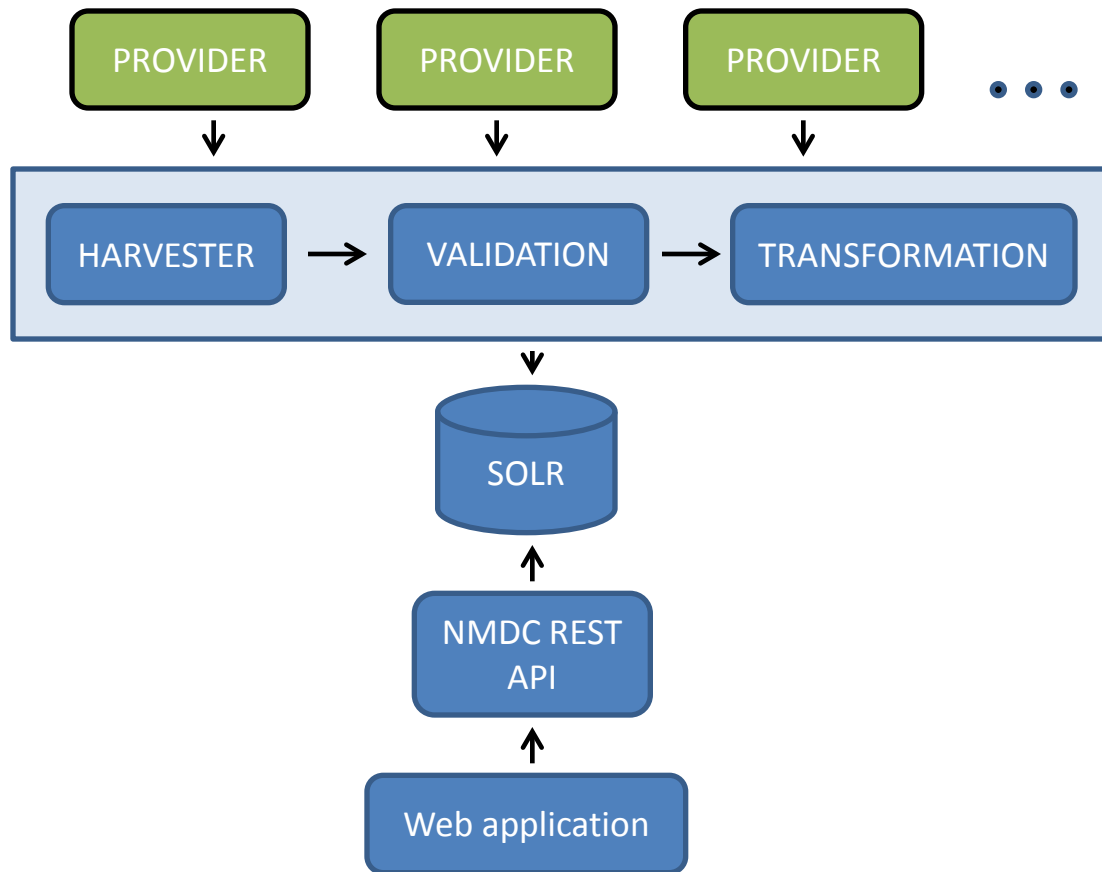
Data flow



Metadata provider

Central node

Application



Metadata harvester

- Periodic retrieval via OAI-PMH protocol
- Formats retrieved ISO 19139/GCMD DIF
Metadata can be provided in either schema
- Harvested metadata are transformed to internal NMDC format
- Metadata providers are notified if metadata fails validation against ISO 19139 or GCMD DIF schemas.

Harvester implementation



- Apache Camel and ActiveMQ
- Processing performed using Enterprise integration patterns
- Separate processes for each operation

Metadata search engine

Apache Solr

- Open source search platform
- Indexes metadata for efficient search
- Free text indexing
- Facet indexing



UiO : Universitetet i Oslo

cmr Computing



INSTITUTE OF MARINE RESEARCH
HAVFORSKNINGSINSTITUTTET



FFI Forsvarets
forskningsinstitutt

ARTSDATABANKEN
Nasjonal kunnskapskilde for biologisk mangfold



Search criteria



Search by:

- Free text
- Science keyword (GCMD Keywords)
- Institution
- Geographical: Bounding Box or Polygon
- Time



Public NMDC API

- REST API
- JSON response
- getFacets,search,getMetadataDetails
- Used by Data catalogue application

URL: http://prod1.nmdc.no/UserInterface/metadata_api

Data download



- Each metadata record contains one or more URLs to data
- Links can be to provider landing page or direct data download
- Links can be access via data catalogue GUI or public api



Data catalogue application



Scientific_Keyword (138)

- EARTH SCIENCE (130)
 - OCEANS (92)
 - MISSING-TOPIC (45)
 - ATMOSPHERE (18)
 - BIOSPHERE (15)
 - BIOLOGICAL CLASSIFICATION (14)
 - EARTH SCIENCE (4)
 - GEOGRAPHIC REGION (4)
 - NOT AVAILABLE (4)
 - SOLID EARTH (4)
 - ACOBAR (3)
 - ACOUSTIC (3)
 - ACOUSTIC TECHNOLOGY FOR OBSERVING THE INTERIOR OF THE ARCTIC OCEAN (3)
 - CLIMATE INDICATORS (3)
 - HUMAN DIMENSIONS (3)
 - VELOCITY (3)
 - PALEOCLIMATE (2)
 - AGRICULTURE (1)
 - LAND SURFACE (1)
 - OCEAN TEMPERATURE (1)

Provider (17)

- Institute of Marine Research, Bergen, Norway (24)
- Geological Survey of Norway, Norway (18)
- Norwegian Meteorological Institute, Norway (11)
- Norwegian Institute for Water Research (10)
- Norwegian Polar Institute (NO/NPI) (9)
- Norwegian Institute for Nature Research (5)
- Nansen Environmental and Remote Sensing Center (4)
- Norwegian Defence Research Establishment (FFI) (3)
- Akvaplan-Niva (2)
- (1)
- NOF-NINA-DN (1)
- Norwegian Hydrographic Service (1)
- Norwegian Ornithological Society (1)
- University of Bergen (1)

Selected facets:
133 data sets found

- N-ICE2015 on-ice based conductivity-temperature-depth (CTD) deep profiles
- Otter in coastal areas - Carcasses
- Hydrographical data from the Outer Oslofjord (CTD) 2001-2013
- MAREANO - Base-line mapping of hyperbenthic crustacea fauna obtained with RP-sledge
- Seabed sediments (grain size), regional mapping (scale 1:50 000 to 1:250 000)
- N-ICE2015 data on oxygen isotope ratios in seawater and sea ice
- N-ICE2015 data on salinity in sea ice, snow and seawater
- N-ICE2015 ship-based conductivity-temperature-depth (CTD) data
- Nutrients in the Inner Oslofjord, 1990-2012
- N-ICE2015 on-ice based conductivity-temperature-depth (CTD) from shallow profiles

Geographical coverage

Filter by map bounding box

Operation Is within

Temporal coverage

Filter by temporal coverage

Operation Is within

From 1980-01-01

To 2016-05-13



Data catalogue application



A digital temperature archive for the Norwegian Sea

PersonLastName	[Ottersen]
Entry_ID	imr_1718
location_rpt	POLYGON((20 60,-20 60,-20 80,20 80,20 60))
Data_Summary	A digital temperature archive for the Norwegian Sea compiled from 59496 temperature profiles is presented. It consists of interpolated temperature fields for the quarters January-March, April-June, July-September, and October-December for each of the years 1990-2007. The archive spans the area 20° W - 20° E, 60° - 80° N with a spatial resolution of 1/2 degree longitude * 1/3 degree latitude and there are 28 depth levels from 0 to 500 m. Two different versions of the archive were produced; one based only on actual data, the second including climatological values where data are missing.
PersonFirstName	[Geir]
Stop_Date	Sun Dec 30 01:00:00 CET 2007
Scientific_Keyword	[EARTH SCIENCE>OCEANS>OCEAN TEMPERATURE>WATER TEMPERATURE]
Provider	[Institute of Marine Research, Bergen, Norway]
Entry_Title	A digital temperature archive for the Norwegian Sea
Data_URL_Type	[GET DATA GET DATA VIEW RELATED INFORMATION, GET DATA GET DATA VIEW RELATED INFORMATION, GET DATA GET DATA VIEW RELATED INFORMATION]
version	1547211282890883072
Data_URL	ftp://ftp.nmdc.no/nmdc/IMR/templatlas/do_odv.zip ftp://ftp.nmdc.no/nmdc/IMR/templatlas/ac_odv.zip ftp://ftp.nmdc.no/nmdc/IMR/templatlas/fh_2009_06.pdf
Start_Date	Sun Dec 31 01:00:00 CET 1989



Challenges



- Encouraging metadata provider to supply meaningful metadata attributes
- Diverse institutions with varying expectations to agree on common approach
- Ensuring each data provider has a data policy for sharing data
 - NLOD (Norwegian)
 - CC BY 4.0
- Changing existing mindset to encourage data sharing



Future direction



- Verification of dataset links
- Collect selected datasets into single download file.
- Support subsetting of datasets for download for large datasets
- Provide permalinks for search results
- Provide Digital Object Identifier (DOI)



Questions