

Norwegian Marine Data Centre



Seamless access to marine data

Terry Hannant
terry.hannant@imr.no



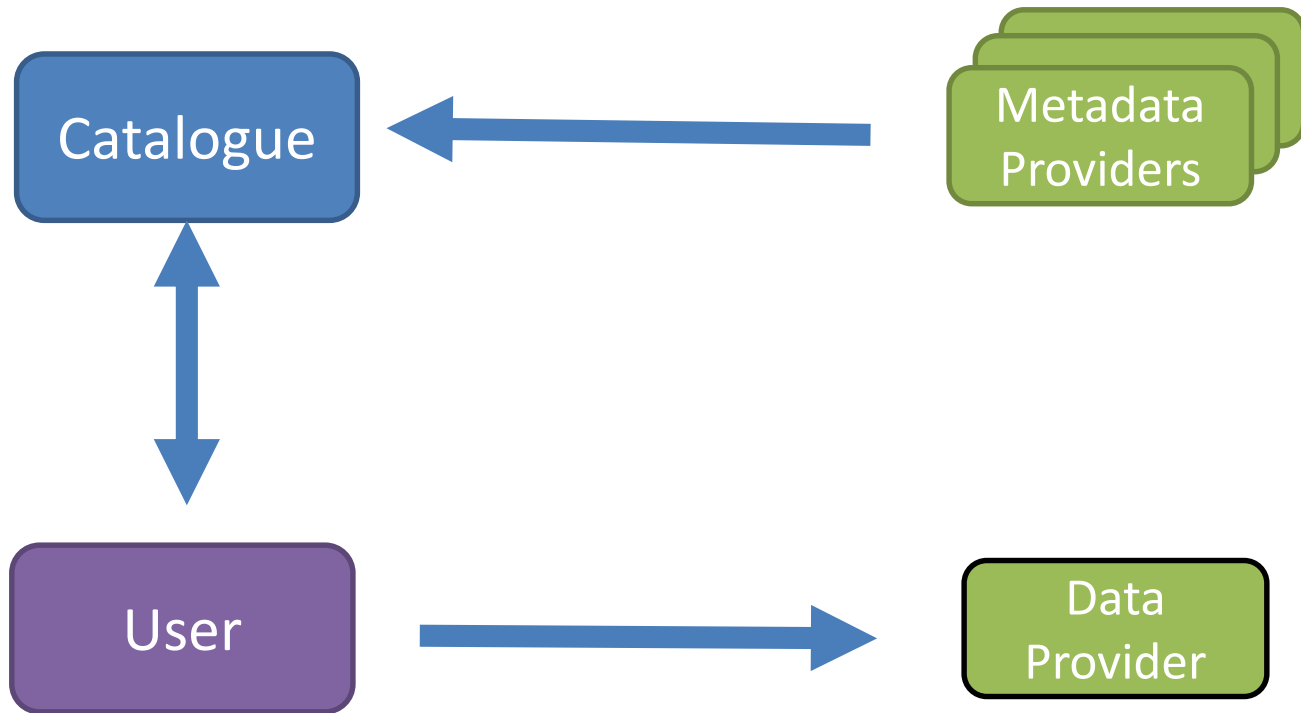
Project overview



- 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



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



UiO : Universitetet i Oslo



INSTITUTE OF MARINE RESEARCH
HAVFORSKNINGSINSTITUTTET



NINA
Norwegian Institute for Nature Research



GEOLOGICAL
SURVEY OF
NORWAY
- NGU -



FFI Forsvarets
forskningsinstitutt
ARTSDATABANKEN
Nasjonal kunnskapskilde for biologisk mangfold



Norwegian Institute for Water Research



uni Research

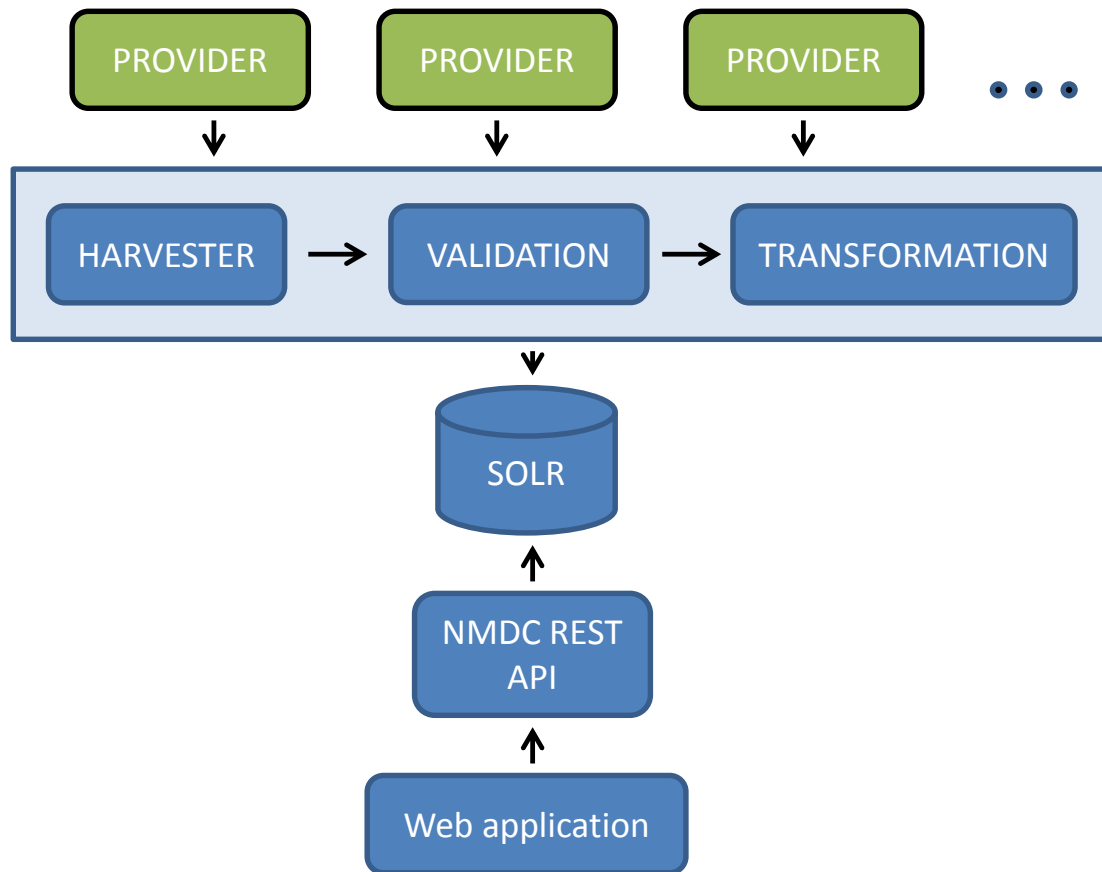
Data flow



Metadata provider

Central node

Application



Metadata harvester



- Periodic retrieval via OAI-PMH protocol
- MetaData retrieved and verified using either ISO 19139 or GCMD DIF
- Harvested metadata are transformed to internal NMDC format



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
 - Science keywords
 - Institution
 - Geospatial (Simple)/Temporal



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



The screenshot shows the NMDC data catalogue application interface. The browser address bar displays `http://prod1.nmdc.no/UserInterface/#!/`. The page features a header with the NMDC logo and a search bar labeled "Scientific_Keyword (138)". On the left, a navigation menu lists categories such as "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)", and "OCEAN TEMPERATURE (1)". Below this is a "Provider (17)" section listing institutions like "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)", "Uio: Universitetet i Oslo (1)", "NOF-NINA-DN (1)", "Norwegian Hydrographic Service (1)", "Norwegian Ornithological Society (1)", and "University of Bergen (1)".

The main content area displays "Selected facets: 133 data sets found" and lists several data sets with expandable icons:

- 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

On the right side, there are filter sections for "Geographical coverage" and "Temporal coverage". The "Geographical coverage" section includes a "Filter by map bounding box" checkbox, an "Operation" dropdown set to "Is within", and a map of the North Atlantic region. The "Temporal coverage" section includes a "Filter by temporal coverage" checkbox, an "Operation" dropdown set to "Is within", and date input fields for "From" (1980-01-01) and "To" (2016-05-13). At the bottom of the main content area, there is a pagination control showing page 1 of 2.



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



UiO : Universitetet i Oslo

cmr Computing



INSTITUTE OF MARINE RESEARCH
HAVFORSKNINGSINSTITUTTET



FFI Forsvarets
forskningsinstitutt
ARTSDATABANKEN
Nasjonal kunnskapskilde for biologisk mangfold



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



Thank you for our attention

