

Arctic Data Centre, bridging scientific and operational data streams

Øystein Godøy, Norwegian Meteorological Institute (Norway), o.godoy@met.no

Bard Saadatnejad, Norwegian Meteorological Institute (Norway), bards@met.no

Background

The Arctic Data Centre (ADC) is a Data Collection and Production Centre within the WMO Information System (WIS). During the International Polar Year (IPY), the Norwegian Meteorological Institute (METNO) was coordinating IPY operational data streams and Norwegian IPY data management. This bridging of scientific and operational communities has since been a primary focus for the activities undertaken within the context of ADC. It is the end point for project specific data collections hosted by METNO, whether funded by EU (e.g. Damocles and ACCESS) or Research Council of Norway (e.g. DOKIPY).

Content

ADC contains both operational data generated by the Norwegian Meteorological Institute as well as research data generated in research projects nationally or through EU projects. These data range from in situ measurements, through satellite remote sensing products to numerical simulations of the ocean and the atmosphere. Much of the data hosted by the Norwegian Meteorological Institute are not yet included in the catalogue, but there is an ongoing effort to properly document and prepare data for sharing through this catalogue interface as well as to integrate legacy systems.

Functionalities supported

ADC supports metadata exchange (harvest and export), visualisation of data (utilising OGC WMS and OPeNDAP), transformation of data (remapping, reformatting, subsetting utilising OPeNDAP access), automated and ad hoc data upload, file format compliance checking, and subscription to datasets.

A strong focus for ADC technological development has been to standardise ad hoc data delivery in self describing file formats like NetCDF following the Climate and Forecast convention wherever possible. To help users submit properly formatted data, conversion tools, templates and guidance documentation have been developed. This ensures cost effective data management for projects while retaining requirements for long term data preservation through a high degree of automation.

Interoperability

The ADC is developed to integrate with other data management systems. It is an approved WIS DCPC implying that metadata are prepared for exchange with WIS centres. Data hosted locally are primarily served using a THREDDS data server offering access through HTTP and OPeNDAP as well as OGC WMS for gridded data. Metadata are exposed and can be harvested using OAI-PMH serving GCMD DIF and ISO19115. Translations of harvested metadata are done using XSLT. This also supports translation of content through mapping of controlled vocabularies used by some of the external partners.

Technological development of ADC is done in synergy with developments for WMO Global Cryosphere Watch, ESA/NSC CryoClim, Norwegian Satellite Earth Observation Database for Marine and Polar Research (NORMAP), Norwegian Marine Data Centre (NMDC) and the emerging Svalbard Integrated Arctic Earth Observing System (SIOS) guided by INSPIRE, WIS, GEOSS and SAON/IASC Arctic Data Committee technical requirements where relevant and appropriate.

ADC serves data to the Norwegian Marine Data Centre through exposure of relevant metadata using OAI-PMH and data using OPeNDAP.

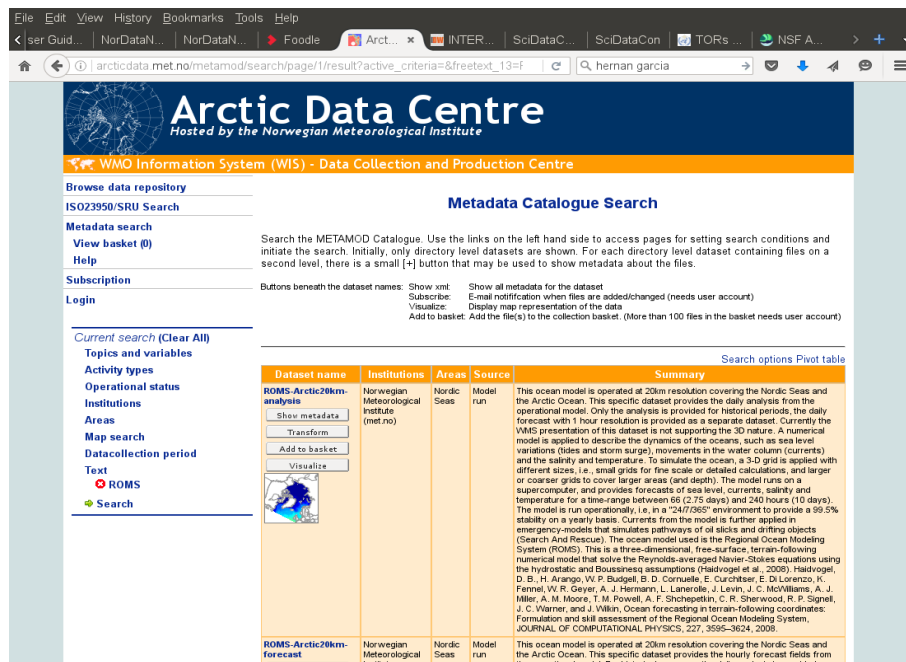


Figure 1: Screenshot of the current web frontend to the arctic data centre.

Future plans

ADC is currently undergoing a transition to a new technological framework to ensure improved flexibility and modularity. The background for this transition was input from user communities requiring increased flexibility when configuring the user interface. ADC has been supporting a number of scientific projects, e.g. EU FP6 and FP7 projects DAMOCLES and ACCESS, and Research Council of Norway (RCN) projects NORMAP and NMDC. User feedback through these projects and IPY, which except for NMDC all used the same software basis, has been consistent. What the development team thought were important, were not necessarily what the users wanted. Furthermore, an internal evaluation of the current software system revealed insufficient scalability and flexibility, or in other words, the system was unable to serve continuously changing external requirements. In order to achieve greater flexibility and scalability, the new platform places emphasis on standard interfaces using standard protocols between its constituent modules. The constituent modules are therefore loosely coupled through HTTP and protocols based on HTTP. This ensures easier integration or replacement of software components as well as the ability to spread the load on multiple servers if required. Continuing this process, user feedback will be collected through web based tools (issue tracking for feature requests, polls, etc.) as well as through workshops. The intention is to continuously show the users the status of feature requests. Interaction with the user community will be coordinated by two RCN projects focusing on data management and user interaction. These are the Norwegian Scientific Data Network which is an infrastructure project, and GeoAccessNO which is a pilot project evaluating geoscientific data management.