

INTERACT: FAIR Data from Cold Region Research Stations

A Data Management Challenge

Radosavljevic, Boris¹; Øystein Godøy²; Margareta Johansson³; Elmer Topp-Jørgensen⁴; Morten Rasch⁵; Boris Biskaborn¹
¹Alfred Wegener Institute for Polar and Marine Research, Potsdam, Germany; ²Norwegian Meteorological Institute, Oslo, Norway; ³Lund University, Lund, Sweden; ⁴Aarhus University, Denmark; ⁵University of Copenhagen, Copenhagen, Denmark

What is INTERACT?

INTERACT (International Network for Terrestrial Research and Monitoring in the Arctic) is a EU Horizon 2020 funded infrastructure project seeking to provide a geographically comprehensive infrastructure for arctic and high altitude research stations. The overall goals of the project include facilitating the identification of environmental and ecological change, the understanding of change and prediction of future changes.

INTERACT Data

Currently, 82 sites in arctic and northern alpine areas are part of the INTERACT network and collect data across various scientific disciplines, e.g. geo-sciences (including the atmosphere and cryosphere), hydrology, biology, ecology, and to some extent anthropology. These data are generated as a result of monitoring activities or short term projects.

Data Management Goal

One of the major tasks in the project is to create a coordinated and unified data management approach. We present the plan to carry out this objective by focusing on four principles: **Findability, Accessibility, Interoperability, and Reusability** (FAIR). This approach would optimize potential future reuse, sharing, and guarantee data and metadata stewardship and preservation.

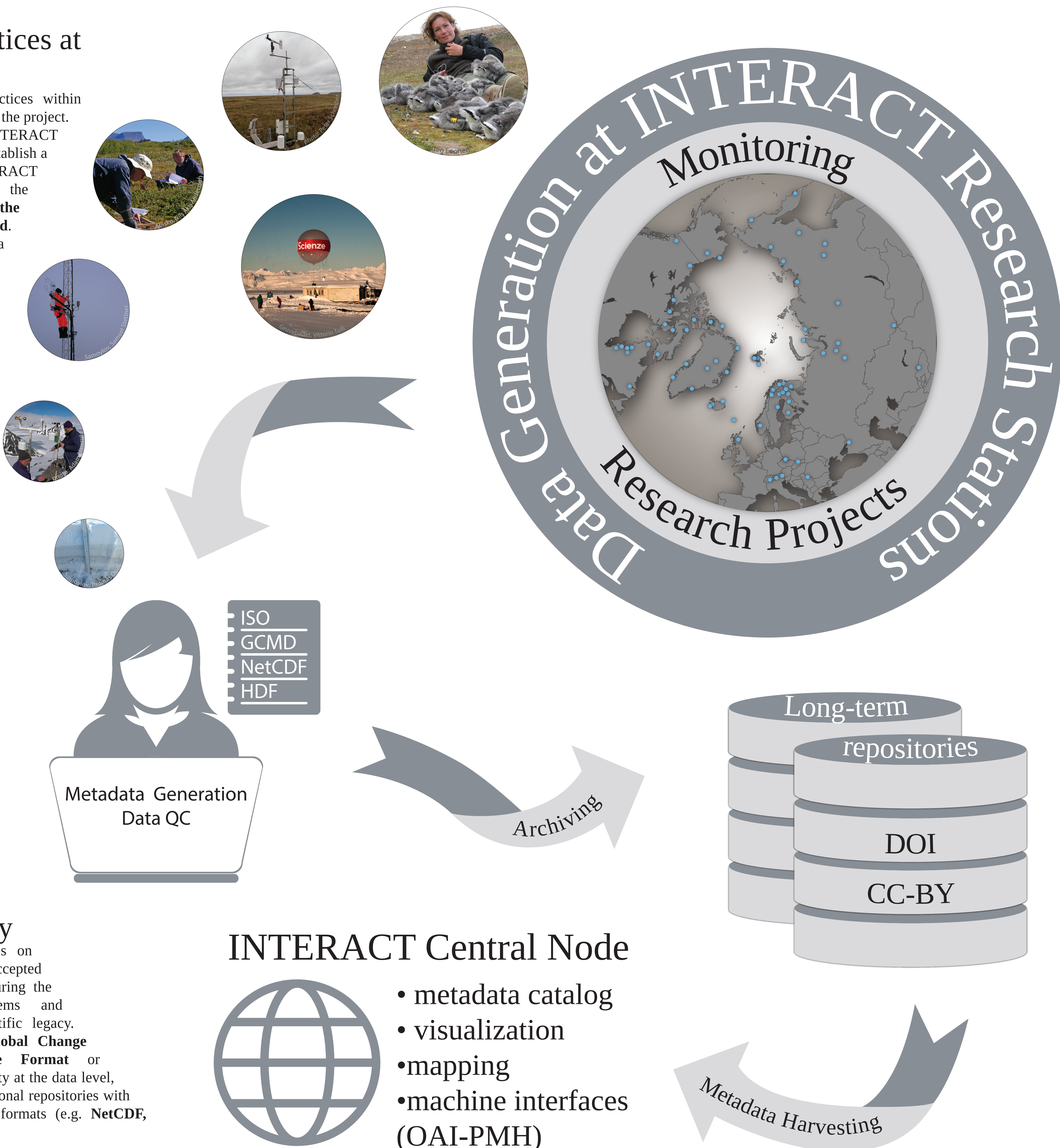
Objective: Establish a unified approach to metadata and data generated by stations in the INTERACT Network

Data Management Practices at INTERACT

A survey of current data management practices within INTERACT was conducted at the beginning of the project. The main finding is that data management at INTERACT stations is highly heterogeneous. In order to establish a unified view on all the data collected by INTERACT stations and demonstrate the benefit of the INTERACT network, **interoperability at the discovery metadata and data levels is required**. The first step towards this is taken through a **Data Management Plan (DMP)** that identifies the general principles regarding common standards and data dissemination principles.

Data Management Plan

INTERACT follows a metadata driven approach. By utilizing **internationally accepted standards and protocols** for documentation and exchange of discovery and use metadata, **interoperability** with international systems and frameworks including WMO systems (e.g. Year of Polar Prediction, Global Cryosphere Watch, Global Atmosphere Watch), and aligned with the activities of the International Arctic Science Committee (IASC) and Sustaining Arctic Observing Network (SAON) Arctic Data Committee (ADC). INTERACT emphasizes **long term data preservation** (as promoted by ICSU-WDS), **community driven best practices** (e.g. RDA), and the principles outlined by the ADC, that promote **free, ethically open, sustained, and timely access** to Arctic data. This approach should provide easy integration with the **H2020 Open Research Data Pilot**, and ensure data access to a variety of stakeholders (e.g. World Meteorological Organization, ESA DUE, GlobPermafrost, researchers, etc.).



Implementation Strategy

The initial data management effort focuses on discovery metadata, utilizing internationally accepted standards, **protocols and vocabularies**, ensuring the interoperability with international systems and frameworks, and the preservation of scientific legacy. Datasets will be documented using the **Global Change Master Directory/Directory Interchange Format** or **ISO19115** standards. To provide interoperability at the data level, long term archival of data across different national repositories with long term mandates in self-explaining file formats (e.g. **NetCDF, HDF/HDF5**) is envisioned eventually.

