# UiT
## THE ARCTIC UNIVERSITY OF NORWAY

# Open Arctic Research Index (Open ARI): A new horizon in Archiving and collecting Arctic datasets

Abu-Alam, T.S.
Universitetsbiblioteket, UiT Norges arktiske universitet; tamer.abu-alam@uit.no; ORCID: 0000-0001-6020-365X

First stage: filtered out none-relevant journals and source institutes

Second stage: filtered out the broken records (e.g. records missing links, text, irrelevant....).

Third stage: compare the records in the temporary database with those that we have in the final database (from the previous filtration process). Removing the duplications and keep only the most updated version.

**Five stages of filtration**

Fourth stage: approve the records from certain sources (e.g. from institutes concerning with the polar data or published in journals related to polar issues).

Fifth stage: check for certain keywords.

Collecting data from major collaborators and metadata providers (e.g. BASE, Polar Data Catalogue, Norwegian Polar Institute, ..... etc).

This option will be used if the metadata provider allow using OAI-PMH service

Filtering the temporary database in three steps filtration process

**Semi final database**

Automatic harvesting of different databases after a permission from the metadata providers.

This option will be used if the metadata provider does not allow using OAI-PMH service

**temporary database**

Uniform the metadata and checking the quality (Automatic/Manual work to reduce the error)

**Final database**

Getting feedback and suggestions/ modifications for new records or existing records from the end-users.

These suggestions will not be added to the database before approval from the admin of Open ARI.

Harvesting all the databases and the existing records on the Arctic and the Antarctic (as possible as we can)

## / Abstract

Open ARI is a planned service at UiT - The Arctic University of Norway in order to collect, sort and archive all the openly available publications and datasets that were published on the Arctic region. This new service will be available as an open access database to users through an interactive searchable front-end.

The pilot project will investigate how such a service can support researchers by making results from Arctic research more visible and better retrievable through a common search index based on a standardized, interdisciplinary metadata set. Moreover and for a better overview for the polar sciences, the new Arctic database will include, as well, examples from the Antarctic region.

As a pilot project, we are developing an existing service at the UiT (i.e. High North Research Documents) by clarifying the need of new technical solution by which the Open ARI will be able to collect all the published material using algorithms that allow the best filtering processes. Also, we are now in a stage to define all the possible national and international collaborators who can support and feed the Open ARI with content. A group of scientist and researchers will be formed as a reference group to define the needs of the scientific community to make sure our final product will meet the interests of the users. By the end of the pilot project, the team will analyse the success opportunities and the challenges in order of planning a full-scale management model.

- Use different social media channels.
- Increase the audience by distributing announcements in journals, newsletters and websites of different organizations and associations deal with polar science.
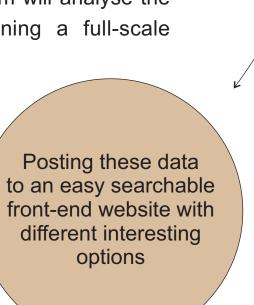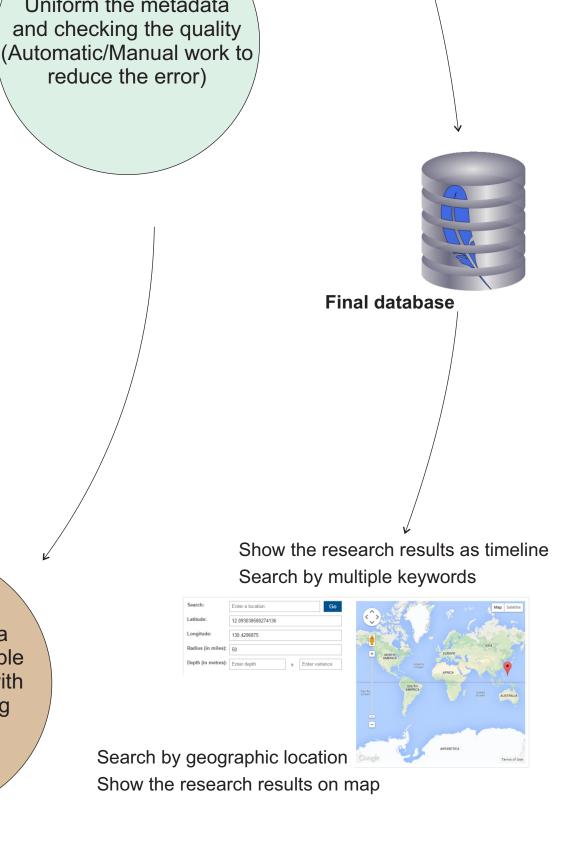- Distribute flyers in well attended conferences and meetings.

Outreach activities to broaden our audience and getting feedback

Posting these data to an easy searchable front-end website with different interesting options

Show the research results as timeline
Search by multiple keywords

Search by geographic location
Show the research results on map

## / Classification of metadata providers

The Metadata providers can be classified into three categories based on their relation with Bielefeld Academic Search Engine (BASE) which is our main metadata provider.

1- Metadata providers not included in BASE
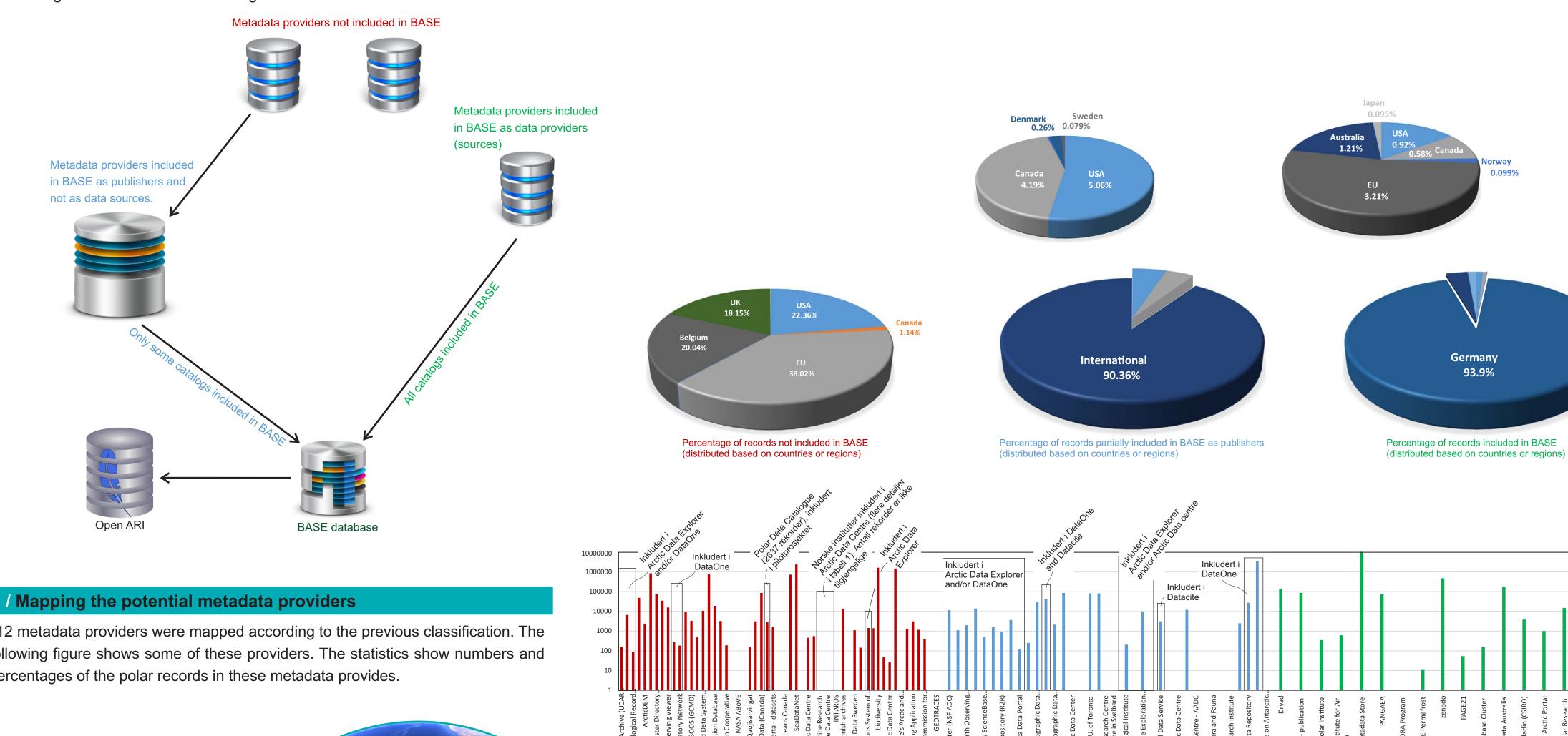2- Metadata providers included in BASE
3- Metadata providers included (partially - as publisher) in BASE. These providers do not give access to all their catalogs.

Metadata providers not included in BASE

Metadata providers included in BASE as data providers (sources)

Metadata providers included in BASE as publishers and not as data sources.

Only some catalogs included in BASE

All catalogs included in BASE

Open ARI

BASE database

To take the OpenARI to the next level - which is not only hosting metadata, but to be a platform to define research gaps and create new forefront knowledge. To do that, we may prepare a web application where scientists can discuss, collaborate and use different databases.

Denmark 0.26%   Sweden 0.079%
Canada 4.19%   USA 5.06%

Japan 0.005%
Australia 1.21%   USA 0.92%   Canada 0.59%   Norway 0.099%
EU 3.21%

UK 18.15%   USA 22.36%
Belgium 20.04%   Canada 3.14%
EU 38.02%

International 90.36%

Germany 93.9%

**Percentage of records not included in BASE** (distributed based on countries or regions)

**Percentage of records partially included in BASE as publishers** (distributed based on countries or regions)

**Percentage of records included in BASE** (distributed based on countries or regions)

## / Mapping the needs, solutions toward a better service

In order to provide the end-users with the best service that will help them in their research and study, we have prioritized twenty requirements that will be included in the OpenARI. These needs can be classified into two categories; a) needs to make the communication between the front-end interface and the back-end database more faster and better performance, and b) needs to make the end-users getting the maximum benefit out of the OpenARI.

Some of these requirements are:

- Add locations value for each record (e.g. Latitudes and longitudes).
- Search using a map.
- Enable showing the results on a map.
- Enable showing the results as a timeline.
- Using an open source user-interface to be sure of keeping the user interface up-to-date.
- Multi-languages user-interface .
- Give the end-users the ability to comment and share different records through different social media channels.
- Give the end-users the ability to upload their original data (i.e. data + metadata). This service can be offered to researchers who do not have the facilities to keep their data online and available to the scientific community. Also, this service can be offered to researchers from UiT and other partners.
- Increase the number of the records in the database by adding more metadata from other metadata providers.
- Better and faster filtration processes of the records.
- Use autocomplete search words algorithm.

## / Mapping the potential metadata providers

112 metadata providers were mapped according to the previous classification. The following figure shows some of these providers. The statistics show numbers and percentages of the polar records in these metadata provides.

Open Data (Canada)
Polar Data Catalogue (PDC)
University of Alberta - datasets
Centre Etude Nordique
Dept. Fisheries and Oceans Canada
ELIDAT
SeaDataNet
ECOMET
Arctic Data Centre
Svalbard Integrated Arctic Earth Observing System
Institute of Marine Research
Norwegian Satellite Earth Observation Database for Marine and Polar Research
Joint WMO/IOC Technical Commission for Oceanography and Marine Meteorology
INTAROS
Norwegian Marine Data Centre (NMDC)
Finnish archives
ARCTIC SPACE CENTRE
Environment Climate Data Sweden
EnviDat
Group on Earth Observations System of Systems
Global Terrestrial Network for Permafrost (GTN-P)
biodiversity
Netherlands Polar Data Center
Italian Arctic Data Center
Archives Hub
Russian Federation Service's Arctic and Antarctic Research Institute
Chinese National Arctic and Antarctic Data Centre
Arctic Research Mapping Application
Global Cryosphere Watch
Ocean Data and Information System
GEOTRACES
Arctic Spatial Data Infrastructure

NSF Arctic Data Center (NSF ADC)
National Snow and Ice Data Center
UCAR NCAR - Earth Observing Laboratory
NASA Earth Observing System
Rolling Deck to Repository (R2R)
Gulf of Alaska Data Portal
NCEI - National Oceanographic Data Center
Environmental Data Initiative
CLIVAR and Carbon Hydrographic Data Office
ICPSR
NCEI National Climate Data Center
NCEI - National Geophysical Data Center
Scholars Portal Dataverse U. of Toronto
Arctic Institute of North America
Yukon Research Centre
Norwegian Meteorological Institute (Met.no)
The University Centre in Svalbard
The Norwegian Meteorological Institute
The Geological Survey of Denmark and Greenland
Swedish National Data Service
UK Polar Data Centre
British Oceanographic Data Centre
Australian Ocean Data Network
Australian Antarctic Data Centre - AADC
Arctic Data archive System
Conservation of Arctic Flora and Fauna
Polar Research Institute
Polar Research Institute of China
OBIS - Ocean Biogeographic Information System
KNB Data Repository
Global Biodiversity Information Facility

**A histogram shows the numbers of the polar records in the different metadata providers.**

Legend:
- Metadata providers not included in BASE
- Metadata providers partially included in inkludert i BASE
- Metadata providers included in BASE
- BASE

## / Progress and time frame

Currently, we are in the midway toward the end of the pilot project. Where we succeed to map the needs, the potential metadata providers, built a new filtration flow in order to increase the efficiency and reduce the time of the filtration process. Also, we have chosen two open source user-interface to be tested. Over the next three months, we will analysis the success opportunities and the challenges in order of planning a full scale management model.

## / Acknowledgments

By scanning me, you will download the poster and will be updated by the progress.

UiT / NORGES ARKTISKE UNIVERSITET
UNIVERSITETET I TROMSØ