INTAROS

# Integrated Arctic Observation System

Research and Innovation Action under EC Horizon2020
Grant Agreement no. 727890

Project coordinator:
Nansen Environmental and Remote Sensing Center, Norway

## Deliverable 1.6

## Data Governance Framework
## (including updated Data Management Plan)
## (revised version after review of period 2)

| | | | |
|---|---|---|---|
| Start date of project: | 01 December 2016 | Duration: | 60 months |
| Due date of deliverable: | 30 November 2019 | Actual submission date: | 19 May 2020 |
| Lead beneficiary for preparing the deliverable: | | NERSC | |
| Person-months used to produce deliverable: | | 0.9 pm | |

Authors: Torill Hamre, Hanne Sagen and Stein Sandven (NERSC), Finn Danielsen (NORDECO), Geir Ottersen (IMR), Agnieszka Beszczynska-Moller (IOPAN), Arnfinn Morvik (IMR), Ingo Schewe (AWI), Martin Enghoff (NORDECO)

| Version | DATE | CHANGE RECORDS | LEAD AUTHOR |
|---|---|---|---|
| 0.1 | 17/10/2019 | Initial version, including D1.2 revised Dec 2018. | T Hamre |
| 0.2 | 22/11/2019 | Updated with partner input and recommended data repositories for INTAROS data. | T Hamre |
| 0.3 | 27/11/2019 | Revisions after internal quality control. | T Hamre |
| 1.0 | 28/11/2019 | Reviewed all content and submitted | K Lygre |
| 1.1 | 15/05/2020 | Revised according to project review | T Hamre |

| Approval | Date: 15 May 2020 | Sign. Coordinator |
|---|---|---|

| USED PERSON-MONTHS FOR THIS DELIVERABLE | | | | | |
|---|---|---|---|---|---|
| ***No*** | ***Beneficiary*** | ***PM*** | ***No*** | ***Beneficiary*** | ***PM*** |
| **1** | NERSC | ***0.3*** | **24** | TDUE | |
| **2** | UiB | | **25** | GINR | |
| **3** | IMR | ***0.3*** | **26** | UNEXE | |
| **4** | MISU | | **27** | NIVA | |
| **5** | AWI | ***0.1*** | **28** | CNRS | |
| **6** | IOPAN | ***0.1*** | **29** | U Helsinki | |
| **7** | DTU | | **30** | GFZ | |
| **8** | AU | | **31** | ARMINE | |
| **9** | GEUS | | **32** | IGPAN | |
| **10** | FMI | | **33** | U SLASKI | |
| **11** | UNIS | | **34** | BSC | |
| **12** | NORDECO | ***0.1*** | **35** | DNV GL | |
| **13** | SMHI | | **36** | RIHMI-WDC | |
| **14** | USFD | | **37** | NIERSC | |
| **15** | NUIM | | **38** | WHOI | |
| **16** | IFREMER | | **39** | SIO | |
| **17** | MPG | | **40** | UAF | |
| **18** | EUROGOOS | | **41** | U Laval | |
| **19** | EUROCEAN | | **42** | ONC | |
| **20** | UPM | | **43** | NMEFC | |
| **21** | UB | | **44** | RADI | |
| **22** | UHAM | | **45** | KOPRI | |
| **23** | NORUT | | **46** | NIPR | |
| | | | **47** | PRIC | |

| DISSEMINATION LEVEL | | |
|---|---|---|
| PU | Public, fully open | **X** |
| CO | Confidential, restricted under conditions set out in Model Grant Agreement | |
| CI | Classified, information as referred to in Commission Decision 2001/844/EC | |

## *EXECUTIVE SUMMARY*

This document contains a description of the Data Governance Framework for the INTAROS project, with an updated version of the Data Management Plan (DMP). The Data Governance Framework defines the procedures for how data management is carried out in the project, including the planning, conducting and monitoring the preparation and distribution of data collections. The DMP describes how new datasets collected or generated by partners in the project, will be managed according to guidelines for FAIR data management in Horizon 2020.

Data governance in INTAROS is pragmatic and geared towards supporting partners in preparing and publishing their data collections. The planning and monitoring activities are carried out by the Data Management Theme Leader and the leaders of the four data generating work-packages in the project. Partners generating data are responsible for making their collections available in line with the recommendations of the DMP. The Data Management Theme Leader, data centre partners (AWI, CNRS, FMI, IMR, IFREMER, ONC, RADI, RIHMI-WDC) and the leader of WP5 ("Data integration and management") (Terradue) are responsible for providing support with technical aspects of data publication and distribution.

INTAROS is pan-Arctic in scope and collect *in situ* observations, extract parameters from satellite data and model projections in several regions and across multiple spheres (themes). The focus areas of INTAROS include Coastal Greenland, North of Svalbard, Fram Strait, the Eurasian Basin, and (5) selected sites in Siberia, Finland, Canada and Alaska. Within these areas, INTAROS partners are collecting new observations and generating high-level data products from different spheres: (1) Atmosphere, (2) Ocean, (3) Sea ice, (4) Marine ecosystems, (5) Terrestrial, (6) Glaciology, (7) Natural hazards, (8) Community-based monitoring. This makes datasets collected or generated within INTAROS relevant for a number of research projects as well as for infrastructures such as EMODNET and GEOSS.

Datasets collected or generated within these spheres by the time of writing are summarised in this document, based on the deliverables from WP 2 ("Exploitation of existing observing systems"), new datasets collected in WP 3 ("Enhancement of multidisciplinary in situ observing systems") and WP 4 ("Enhance community-based observing programs for participatory research and capacity-building"), as well as upcoming model products and derived datasets from WP6 ("Applications of iAOS towards Stakeholders"). Datasets prepared for distribution in WP 2, 3 and 4 have also been registered in the INTAROS Data Catalogue, available at https://catalog-intaros.nersc.no/. This data catalogue will be updated with new datasets collected or generated during the remainder of the INTAROS project.

The DMP recommends standards for metadata and data standards that INTAROS partners should prepare their datasets in, to make it easier for other scientists and stakeholders to reuse the data. Open source tools can help scientists generate metadata and data in standard formats, such as Rosetta, GDAL (Geospatial Data Abstraction Library), NetCDF utilities, and widely used programming languages, such as Python, MATLAB and R, offer libraries that can be used to write customised format converter tools. A dataset prepared in NetCDF format can be made publicly available using data publishing tools like the Thredds Data Server (TDS). INTAROS, together with the Useful Arctic Knowledge (UAK) project has organised several user meetings and one research schools, to build competence in data management within the INTAROS consortium. Additional competence building activities are planned in INTAROS; the training material developed will be made publicly available.

# Table of Contents

# 1. Introduction

The Data Governance Framework (DGF) for INTAROS defines the procedures for how data management is carried out in the project, including the planning, conducting and monitoring the preparation and distribution of data collections. The planning and monitoring activities are carried out by the Data Management Theme Leader and the leaders of the four data generating work-packages in the project (WP2, WP3, WP4, WP6). Partners generating data are responsible for making their collections available in line with the recommendations of the The Data Management Plan. The Data Management Theme Leader, data centre partners (AWI, CNRS, FMI, IMR, IFREMER, ONC, RADI, RIHMI-WDC) and the leader of WP5 ("Data integration and management") (Terradue) are responsible for providing support with technical aspects of data publication and distribution.

The Data Management Plan (DMP) for INTAROS describes how new datasets collected or generated by partners in the course of the project, will be managed according to guidelines for FAIR data management in Horizon 2020 (European Commission, 2016). This includes a plan for preparation of datasets collected or generated in the project in standard formats documented according to best practices, ingestion of the new data in an established data repository, as well as the long-term storage and curation of data.

This document contains an updated version of the DMP. It provides a summary of datasets generated in the three first years of the INTAROS project, recommendations for data and metadata standards that partners should use to make it easier for other scientists and stakeholders to reuse their datasets, and an overview of capacity building material for data management that is prepared in the INTAROS and Useful Arctic Knowledge (UAK) projects. Scientific datasets collected or machine-generated during project are registered in the INTAROS Data Catalogue, which holds key metadata elements for each dataset (Appendix A). For community-based data special conditions may apply; Appendix B contains the information and consent form for all participants in the community-based activities within INTAROS. Within the stakeholder application work-package, WP6, a range of models will be run, and this will generate model-based data collections. Appendix C contains a template with key metadata elements describing the models used.

The DMP is a living document; it will be updated regularly during the project and used actively to ensure that INTAROS datasets are made publicly available. Later versions of the DMP will provide further details on new actual datasets as they are collected or generated, as well as the general framework for sustainable data management in INTAROS. An updated public version of the DMP will be available in May 2021.

### INTAROS project objectives and concepts

INTAROS (Integrated Arctic observation system) is a research and innovation project under the BG-09-2016 call and will run from December 1, 2016 to November 30, 2021. INTAROS is developing an integrated Arctic Observation System (iAOS) by extending, improving and unifying existing systems in the different regions of the Arctic. The iAOS will integrate distributed repositories holding ocean, atmosphere, cryosphere, terrestrial, and community-based data, and provide a common entry point to data originating from a wide range of observation networks, scientific campaigns and satellite missions, as well as new data generated within the project.

Existing observing systems, data repositories and infrastructure available from partners and collaborators will be the building blocks of iAOS. These observing systems and data repositories are assessed in WP 2 ("Exploitation of existing observing systems"). New data and generated products from INTAROS will be stored in an established data repository based on the outcome of this assessment. Thus, INTAROS will not build up a new e-infrastructure for data storage and preservation, but instead capitalize on the many existing research data infrastructures in Europe, US, Canada and Asia, that hold environmental data for the Arctic.

### Themes addressed by INTAROS

INTAROS has a strong multidisciplinary focus, with tools for integration of data from atmosphere, ocean, cryosphere and terrestrial sciences, provided by institutions in Europe, North America and Asia. In addition to data collected and generated by scientists, the iAOS will also integrate data from local communities. Thus, the themes addressed by INTAROS include:

1. Atmosphere
2. Ocean
3. Sea ice
4. Marine ecosystems
5. Terrestrial
6. Glaciology
7. Natural hazards
8. Community-based monitoring

## 2. General data management principles

Data management in INTAROS is carried out in accordance with guidelines for FAIR data management in Horizon 2020 (European Commission, 2016). This means data collected or generated in the project must be:

- F (Findable) – "making data findable, including provisions for metadata"
- A (Accessible) – "making data openly accessible"
- I (Interoperable) – "making data interoperable"
- R (Reusable) – "increase data re-use (through clarifying licenses)"

A key element of making data *findable* is ensuring that all datasets are accompanied with rich metadata describing the contents and how data has been processed and quality controlled, as well providing a persistent identifier that uniquely identifies every dataset and supports versioning of datasets.

All data generated by the project will be made *available* according to an Open Data policy in line with the recommendations from Horizon 2020. This will enable external parties the right to access and use the digital datasets created by INTAROS, while respecting general terms and conditions as defined in the Grant Agreement and acknowledging that the originating partner retains the ownership of their datasets. The INTAROS consortium will define an open data policy, as part of its data governance framework. The data governance framework will define a set of guidelines for governance of the datasets generated by INTAROS, taking into account, among others, the needs for documenting data quality, procedures for secure long-term storage and curation, as well as a mechanism for data search, retrieval, and use.

INTAROS will make the datasets collected or generated in the project available in standard data formats, with discovery metadata encoded using standard vocabularies. This enables INTAROS datasets to be easily integrated in the iAOS and contributes to making them *interoperable* with in other Arctic data portals and services.

Data collected or generated by INTAROS during field campaigns and extensions to ongoing observations will be publicly available without undue delay. As a general rule, the data provider (i.e. partner) shall publish their data no later than 2 years after the collection period ends. In cases where PhD students or PostDoc researchers rely on data for publications, the embargo period can be extended with up to 2 additional years. This is to support young scientists in advancing their scientific careers. Data from community-based monitoring programmes and collaboration with local communities in the project, will be made available as agreed with the respective communities. Access restrictions may apply to community data. This is done to protect the rights and interests of the local communities and the individuals living there. Both scientific and community-based datasets will be accompanied by a data license clearly stating how the data can be *used* by the scientific and wider user community.

### Data governance

INTAROS will use a simplified version of the DAMA-DMBOK2 data governance framework (Cupola et al., 2014). Focus will be on establishing formal procedures for planning, conducting and monitoring data management activities in the project, given its scope and international character. The planning and monitoring activities are carried out by the Data Management Theme Leader and the leaders of the four data generating work-packages in the project. Partners generating data are responsible for making their collections available in line with the recommendations of the DMP. The Data Management Theme Leader, data centre partners and the leader of WP5 ("Data integration and management") are responsible for providing support with technical aspects of data publication and distribution.

INTAROS relies on existing data infrastructures holding Arctic data. Thus, the organisation and storage of data will be handled by the respective infrastructures. INTAROS role is to recommend a set of such data infrastructures that are mature and sustained through national or international long-term commitments. These data infrastructures may include data warehouses compiling data from different sources optimized for business intelligence applications supporting decision-making for the organisation(s) owning the infrastructure or a given customer segment (e.g. environmental assessment, offshore operations). Alternatively, these aspects of the general DAMA-DMBOK2 data governance framework may be addressed by some of the stakeholder applications within the project.

Data security is key aspect of INTAROS, especially for community-based monitoring (CBM) data. These data are typically managed by the communities themselves, or information and communications technology (ICT) specialists of collaborating research organisations or community networks like ELOKA[1]. INTAROS has taken measures for protecting CBM data from the beginning. For instance, the original CBM data are password protected, and non-sensitive data intended for wider use can only be browsed by guests if the guests first agree to adhere to a set of ethical and appropriate use guidelines, and to cite the data if it is used in publications. Within the project, the data collected as part of CBM activities have been managed

[1] Exchange for Local Observations and Knowledge of the Arctic

by the corresponding local community or community network. The CBM data ownership and data use rights follow principles of "Free, Prior and Informed Consent" (Appendix B). Some of the field experiments in WP3 have been conducted by PhD students or young scientists. In these cases, INTAROS has taken measures to protect their data until they have published a peer reviewed paper or thesis based on these data. In addition to traditional scientific papers, PhD students are encouraged to publish data papers describing their data and to reuse data in other scientific studies or (if permitted) in commercial settings.

Selected aspects of the DAMA-DMBOK2 data management knowledge area "Data integration & interoperability" is addressed by the development of the iAOS in WP5. The iAOS will provide a set of services to illustrate the integration of multi-source data, offer federated search, various geo-statistical operations, and other services relevant to the stakeholder applications in WP6.

The INTAROS Data Catalogue holds metadata describing the data collections generated by the project. All partners generating data collections in the project, are obliged to register their data in this catalogue. Additional metadata for data collections that can be used in the stakeholder applications in WP6 can be registered as well or harvested from other catalogues offering a standard interface like OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) and OGC CSW (Open Geospatial Consortium Catalogue Service for the Web).

Data collected or otherwise generated within INTAROS originate from many different spheres. Consequently, the processing chains and standards used for calibration and quality checking differ accordingly. Thus, each partner in WP2, 3, 4 and 6, will follow the best practices and use the community standards for data collection within their scientific domain or community-based monitoring. Resulting data collections will be stored in one of the recommended data infrastructures in this report (Section 8), with metadata describing the data quality.

### Metadata and data standards

To facilitate uniform discovery of multi-source heterogeneous data, iAOS will build on established standards for metadata search and retrieval. iAOS will integrate data from a number of spatial data infrastructures and use best practices for search services using OpenSearch with Geo, Time and EO extensions that allow standardized and harmonized access to metadata and data of data providers. This solution will facilitate the aggregation of results between disparate data providers via OpenSearch common standards, allow search engine discovery (using OpenSearch Description Documents) and, mostly importantly, facilitates smooth integration between related server's OpenSearch implementations.

For the INTAROS datasets, partners will use standards such as GCMD DIF and ISO 19115 to hold general descriptive metadata, i.e. discovery metadata, to support flexible search in iAOS. Standard vocabularies such as GCMD Science Keywords and SeaDataNet-2 vocabularies (e.g. for parameter and sensor names, units, keywords) will be used to mark up the metadata in a manner that facilitates machine readable search and retrieval. Metadata following GCMD DIF must be encoded using DIF-9 or DIF-10 XML schemas; metadata following ISO 19115 must be encoded in XML using the ISO 19139 standard.

To facilitate reuse of a dataset, a more detailed description of its content, including among others, parameters and units used, is needed. Such usage metadata are well defined for some types of data, e.g. physical oceanography data from CTDs and EO data from satellites.

However, for other types of data, there is a lack of standards. For instance, while GCMD and ISO19115 can represent general descriptive metadata for acoustic data from scientific experiments, there is currently no standard that can fully represent the usage metadata. Thus, standardization of usage metadata is needed to integrate passive acoustic data from distributed ocean acoustics observation networks.

In the NorDataNet[2] project funded by the Research Council of Norway, NERSC has developed a new data format for the acoustically sensed ocean temperatures from the Fram Strait Multipurpose Acoustic system (Yamakawa et al., 2019). This format is based on the NetCDF Climate and Forecast (CF) Metadata Conventions and the metadata structure developed by the OceanSITES program (OCEANSites, 2010). NERSC has developed a second format, also based on NetCDF/CF and OceanSITES, for the ambient noise data from the Fram Strait Multipurpose Acoustic system. The datasets from EC projects DAMOCLES and ACOBAR have been converted to this format and stored at NERSC. During INTAROS, these datasets will be made available in iAOS as well as NMDC.

To describe usage metadata in a consistent and machine-actionable manner, it is imperative to use standard vocabularies. These offer unambiguous definitions of key concepts and terminology within a scientific domain, with each term having its unique identifier (URIs). This enables computer programs to distinguish between syntactical similar terms and to deduce the semantics (i.e. meaning) of the metadata. As an example, the GCMD keyword vocabulary[3] describes Science and Services Keywords (Olsen et al., 2013), Data Centers, Projects, Instruments, Platforms, Locations, Horizontal Data Resolution, Vertical Data Resolution, Temporal Data Resolution and URL Content Types. The NERC Vocabulary Server (NVS) Common Vocabulary [4] holds an extensive set of vocabularies for parameters, sensors, platforms, ships, organisations, projects, to name a few. These vocabularies have been developed by the British Oceanographic Data Centre in collaboration with the scientific community in a series of projects for the past decades. Currently, the development of the NVS Common Vocabulary is carried out within the frame of SeaDataNet[5], a distributed Marine Data Infrastructure with more than 50 partners and sub-contractors from Europe, Russia, Africa and Australia. A third highly relevant vocabulary is the list of standard parameters names and units in the NetCDF Climate Forecast (CF) conventions (Eaton et al., 2011). NetCDF/CF defines a set of required elements, including descriptive and usage metadata, that should be stored in a NetCDF file to allow users to both discover datasets and decide whether they are fit for a particular purpose. With NetCDF/CF each variable can be described in detail using standard names and units, allowing computer applications to extract, process (e.g. re-grid) and display user selected values. In addition to parameter names and units, the CF also enables definition of, amongst others, time (point or interval), coordinate axis, depth axis and map projection. To describe the data set as a whole, CF includes metadata elements for data set title, name of institution producing the data originally, data source, history (of processing), references (e.g. scientific or technical literature) and comments.

For the datasets themselves, the following standard data formats are recommended (in alphabetical order):

- CSV (comma-separated values)
- Darwin Core (Taxonomic Databases Working Group Standard)*

[2] https://www.nordatanet.no/
[3] http://gcmd.nasa.gov/learn/keywords.html
[4] https://www.seadatanet.org/Standards/Common-Vocabularies
[5] https://www.seadatanet.org/

- NetCDF/CF1.6 (NetCDF/Climate and Forecast metadata convention)*
- GeoJSON
- GeoTIFF*
- JSON (JavaScript Object Notation)
- HDF-EOS (Hierarchical Data Format - Earth Observing System)*
- OGC GeoPackage (Open Geospatial Consortium GeoPackage)*
- Shapefile*

Several of these formats allow for metadata to be encoded in a machine readable manner as part of the data file(s), which is highly recommended. Such formats are marked with '*'in the list above. For the other formats, metadata should be placed in an accompanying file, according to the agreed metadata structure using standards such as the Directory Interchange Format (DIF) Standard from NASA or the ISO 19115:2013 "Geographic Information — Metadata" standard from ISO/TC 211.

### Data repositories

INTAROS will not establish new data repositories, but instead utilize existing data repositories and data infrastructures holding Arctic data in Europe, US, Canada and Asia. With data being collected in or estimate for many different regions of the Arctic and across multiple spheres, there is no single data repository that is an optimal choice for storing INTAROS data. The partners have therefore formulated general requirements for the data repositories to be recommended for long-term storage and curation of INTAROS datasets. These data repositories must, among others

- Support metadata standards DIF-9, DIF-10 or ISO 19115
- Prescribe use of standard vocabularies for metadata (such as those listed above)
- Dates shall always be represented in ISO 8601 format
- Prescribe use of standard data formats (such as those listed above)
- Offer a unique identifier for each dataset, such as a Digital Object Identifier (DOI)
- Provide an open data policy and a data license for all datasets
- Have an established and well documented data governance framework
- Offer secure storage with appropriate access control for scientific and sensitive data
- Have sustained funding for regular operation and maintenance the next 10+ years

In addition, the data repositories must provide machine-to-machine interfaces for metadata and data exchange. Specifically, they must support

- Metadata search and access through OAI-PMH and OGC CSW
- Data access through OPeNDAP, OGC WFS or OGC WCS for extraction of user defined subsets of datasets
- Data access through FTP or HTTP for download of full datasets (as files)
- Generation of simple maps (raster images) through OGC WMS of selected parameters in a dataset

### Long-term plans for data management in a Sustainable Arctic Observing System

In 2014, the Sustained Arctic Observing Networks (SAON) established the Arctic Data Committee (ADC) to map Arctic data management projects and services with their locations and relationships, to identify and promote common metadata elements, to provide a guide on data publication and citation, and to conduct interoperability experiments for selected regions

and SBAs (Social Benefit Areas). They also established the Committee on Observations and Networks (CON), to advise the SAON Board on funding, coordinating and extending existing observation systems, and planning for their sustainability. Several Spatial Data Infrastructures (SDIs) and data repositories hold data for either the whole, or part of, the Arctic. These SDIs are operated by different organizations and communities world-wide, making it a challenge to reach agreement on common metadata and data standards, data policies and governance frameworks.

INTAROS has established contact with both SAON committees, ADC and CON, as well as with other initiatives, such as YOPP and GEOCRI, that address data management and interoperability between spatial data infrastructures for Arctic regions. INTAROS is now a member of the SAON Roadmap Task Force, comprising the pan-Arctic forum for developing a Roadmap for a Sustainable Arctic Observing System (SAOS). This is a major deliverable from INTAROS by the end of the project (autumn 2021). Defining a data governance framework that will address different aspects of data interoperability will be a key element of this roadmap.

# 3. In situ datasets collected by INTAROS

INTAROS collects data using platforms such as aircrafts, research vessels and ships of opportunity, stations on land, ice-tethered instruments (SIMBAs and IAOOS platform), gliders, bottom-anchored systems and fixed ocean moorings in WP3 (Enhancement of multidisciplinary in situ observing systems). The data collection in INTAROS is designed to fill selected gaps in the following regions: (1) Coastal Greenland, (2) North of Svalbard, (3) Fram Strait, (4) Eurasian Basin, and (5) sites in Siberia, Finland, Canada and Alaska for terrestrial and atmospheric measurements. Additional activities are focused on integration of seismometers and biogeochemical sensors into the existing monitoring programs by adding an Arctic component to EPOS and ICOS infrastructures.

In situ observing systems in the Arctic are limited due to logistical constraints and the high cost of deploying and maintaining equipment in this region. The sparseness of in situ data is therefore the largest gap in the overall observing system. Therefore, the in situ data collection activities in INTAROS aim at filling some of these gaps, especially in the ocean which is severely undersampled compared to the other spheres (terrestrial, atmosphere, cryosphere).

During the 2017-2018 field season, the pilot phase of INTAROS fieldwork, INTAROS partners have organised or participated in campaigns in Alaska, Eastern Canadian Arctic, Greenland, Fram Strait, North of Svalbard, central Arctic, northern Finland and Siberia. The in situ datasets collected during the above field campaigns will be registered in the INTAROS Data Catalogue once they have undergone quality control and are made available for use by other scientists and stakeholders. As part of the general data management principles INTAROS strongly recommends open publication of collected data no later than 2 years after the collection periods ends and the instruments are retrieved by the PI (ref Section 2)

During the first INTAROS field season, 2018-2019, the observing infrastructure have been implemented by INTAROS partners or already in operation on Greenland, including coastal regions of the Greenland Ice Sheet (GIS) and Greenland coastal waters (in particular the Young Sound), in the Nansen Basin north of Svalbard, in Fram Strait and Svalbard fjords (Kongsfjorden), in the central Arctic Ocean, in Alaska, Eastern Canadian Arctic, northern

Finland and Siberia. Data collected by the INTAROS partners cover many different spheres and include tens of different variables, specific to each sphere. Consequently, the processing chains and standards used for calibration and quality checking differ accordingly as well as data formats and data repositories selected for long-term preserving of collected observations. In general, data from instruments and platforms, operating during the 2018-2019 field season, will be available after the relevant data processing and quality control is completed for each different type of data and the final data products are submitted to open data bases or made available through dedicated ftp servers or websites. Detailed information on data types, data processing and planned storage for each activity in WP3 are provided in WP 3 deliverables (D3.6, D3.7, D3.8, D3.9, D3.14), focused on the results from the first implementation of INTAROS observing platforms and sensors.

New data collected in the coastal Greenland during the first field season include time series of physical ocean variables collected on four moorings in the Young Sound (e.g. temperature, salinity, particles) and measurements on stations during the research cruises, including CTD, carbonate chemistry (pCO2, DIC, and TA) and nutrients profiles. New data on snow water equivalent were collected with SnowFox instruments on the GIS together with standard AWS measurements and new measurements of rainfall rate were also gathered. Acoustic recordings were collected under the sea ice in the Young Sound for bio- and geo-sound studies.

The array of INTAROS moorings north of Svalbard deployed in 2018 was partially recovered in 2019. However due to very heavy ice conditions in the mooring area in summer and autumn 2019 and limited ship time available, one or two moorings will likely have to remain in water until 2020 (at the moment of this report writing, one more late autumn cruise is planned to attempt recovery of these moorings). Instruments deployed on the moorings north of Svalbard provided time series of temperature, salinity, ocean currents (both point measurements and vertical profiles), dissolved oxygen, nitrates, pCO2 and pH, and biological variables (particles, images collected with the underwater vision profiler) and passive contaminants samples. Seismic observations were collected with three ocean bottom seismometers.

In Fram Strait the Autonomous arcFOCE (arctic Free Ocean Carbon Enrichment) system was deployed for the first operational period in 2018-2019 and provided observations of deep-sea bacteria and meiofauna organisms (focus on nematodes) under ocean acidification, i.e. changes in community structure and ecosystem functioning. Time series of temperature, salinity and carbonate chemistry (pCO2, pH, TA) were continuously collected with real-time data availability in Kongsfjorden at the AWIPEV observatory. The long-term acoustic recordings to study the soundscape diversity including benthic fauna sounds, marine mammals' vocalisations, ice sounds, boat noise, wind/wave noise were also collected in Kongsfjorden.

In the central Arctic Ocean, the upper ocean CTD profiles were collected in autumn 2018 with the ice-tethered IAOOS platform together with data from sea ice mass balance instrument, weather station and micro glider. Unfortunately, the IAOOS platform stopped working after 4 weeks, the most likely due to damage by polar bears. In 2019 instead of deploying the second ice-tethered platform, the deep ocean mooring was deployed for one year in the Nansen Basin. Instruments deployed on this mooring will provide measurements of temperature, salinity, dissolved oxygen, ocean currents, carbonate chemistry (pCO2 and pH), sea ice drift and draft and passive acoustics. Measurements of snow-ice-ocean temperature profiles, snow depth and ice thickness were collected with several ice-tethered SIMBA buoys in the central Arctic Ocean. Physical ocean measurements (temperature, salinity, dissolved oxygen and fluorescence) were collected by gliders in the upper 1000 m in the northern Fram Strait (in open

water) along repeated sections. Surface CTD, pCO2, CDOM, chlorophyll a and turbidity measurements were collected by the FerryBox system along the repeated transect between Tromsø and Svalbard. BGC-Argo floats deployed in the Baffin Bay provided standard CTD data and additional biogeochemical measurements (O2, PAR, CDOM, BB, chla, nitrates).

In the Canadian Arctic, measurements of air temperature, wind, radiation, snow temperature, thermal conductivity, snow height, soil temperature, thermal conductivity and water content were collected at 4 monitoring sites. Measurements with the unmanned atmospheric observatory on the icebreaker Oden during the Arctic Ocean 2018 cruise included wind, humidity and air temperature observations, ocean surface temperature, incoming BB LW and SW radiation, precipitation, visibility, cloud bases and fractions, and backscatter, and eddy covariance fluxes of momentum, sensible and latent heat, CO2 and CH4, and aerosols. Soil temperature profiles were measured at different depth at three observations site in the Eastern Arctic. Automated air sampling to collect observations of GHG was implemented in summer 2019 in the North East Siberia.

Data collected by different systems and platforms in 2018-2019 will be stored in relevant data repositories, listed in Section 8. Detailed information on planned data repository for each type of newly collected observations can be found in INTAROS deliverables D3.6, D3.7, D3.8, D3.9, D3.14, and in the INTAROS Data Catalogue for those data that have already been registered there.

# 4. Community-based datasets collected by INTAROS

INTAROS partners mainly work with community-based observing systems in two Arctic local communities, Longyearbyen, Svalbard, and Disko Bay, Greenland. The two selected communities are high-risk regions in terms of climate change impacts as well as loss of biological diversity and can potentially benefit significantly from community-based observing programs. Furthermore, the two communities are characterized by economies which will benefit from efficient and low-cost observing programs at local levels.

During workshops and meetings in the two communities INTAROS partners in WP 4 (Enhance community-based observing programs for participatory research and capacity building) have discussed with the local authorities (Governor of Svalbard; Qeqertalik and Avannaata municipalities) to define a set of topics to be addressed in the project. Based on these topics, relevant existing datasets have been identified, and agreements made to collect new data that can augment the existing records, within the framework of INTAROS and/or existing community-based observation networks. New community-based datasets will also be included in future versions of the DMP and be registered in the INTAROS Data Catalogue.

First, in Qeqertalik and Avanaatta municipalities, contributing to Greenland Government's observing network Piniakkanik Sumiiffinni Nalunaarsuineq (PISUNA), local natural resource experts, typically fishermen and hunters, have formed Natural Resource Councils. The members compile data on species and resource uses. The data are stored at pisuna.org and PISUNA-net (see https://eloka-arctic.org/pisuna-net/). At pisuna.org, the data are stored in their original words, language and format, exactly as they were reported by the fishermen and hunters. In PISUNA-net, English versions of the data are stored in a searchable, web-based database. Second, in Longyearbyen and Qeqertalik municipality, INTAROS has established pilot community based seismic stations for cryoseismological recordings, led by GEUS and

UiB. The community-based seismometers collect data from the local seismic activity. These data are accessible in real-time on https://raspberryshake.net/stationview/. Observations from the community-based seismic stations are included in the GEUS' earthquake bulletin (see https://www.geus.dk/natur-og-klima/jordskaelv-og-seismologi/registrerede-jordskaelv-i-groenland/). They will also become part of the iAOS developed by INTAROS. Third, in Svalbard and Greenland, INTAROS with many partners has begun testing environmental monitoring by tourist cruise expedition operators. This initiative builds on seven citizen science programs: Secchi Disk Study, Cloud Observations, Cultural and Historical Site Photography, Happywhale, eBird, Tidal Glaciers as Hot Spots for Top Predators Feeding, and Plastic Debris on Arctic Shores. The data are stored within databases held by each citizen science programme. Finally, as a spin-off of the INTAROS capacity-development for local communities, community-based monitoring tools are used by Indigenous fishermen, hunters and reindeer herders in Zhigansk and Olenek Districts in Yakutia. The data are stored within databases held by the Republic Indigenous People Organisation of Sakha Republic.

INTAROS will take the gender dimension into account when approaching the targeted stakeholder groups, including local communities (Lygre and Sagen, 2017). An important element in involving local community members in data gathering in INTAROS, regardless of gender, is informing of the project objectives and benefits of participation for their community. The first template for information and consent form for involving local communities in data gathering for INTAROS (see Appendix B), is focused on hunting and fishing which is dominated by males. Later in the project, we will define another template to address topics of relevance for females in these local communities. This will be done in collaboration and dialogue with local communities to engage the women.

# 5. Higher level products based on satellite, in situ and model data

In WP2 ("Exploitation of existing observing systems"), new products derived from existing satellite, in situ and model data are developed and made openly available through existing repositories. Categories of such products include:

- Regionally and seasonally downscaled products based on existing datasets of carbonate system chemistry; nutrients and phytoplankton
- Unified collection of temperature and salinity data from the Arctic Ocean, compiled from a wide range of data sources and thoroughly quality-checked.
- Improved satellite-based ice concentration, ice thickness and ice drift products using new Sentinel data, passive microwave and optical data.
- ESA Sentinel/Copernicus data integrated with *in situ* observations and climate modelling to deduce sub-annual mass loss estimates from the Greenland ice sheet at sub-basin scale.
- The Alaskan transect of 5 eddy covariance towers measuring CO2, H2O, CH4 and energy fluxes will be enhanced to provide continuous, year-round data on GHG concentrations and fluxes, as well as active layer depth, water table depth and snow depth.

Higher lever data products currently prepared as part of Task 2.3 are registered in the INTAROS Data Catalogue. More higher-level products will be added to this catalogue as partners develop new algorithms to integrate and synthesize information from time series or multi-source data. In addition, new integrated products will be developed for selected stakeholders in WP 6

(Applications of iAOS towards Stakeholders) using iAOS (WP 5) and model products. These products will also be registered in the INTAROS Data Catalogue.

# 6. Model results generated for stakeholder applications

Within INTAROS WP6 a range of different models will be applied to provide model output and products targeting diverse stakeholder groups. Because of the diversity in model output and user groups (scientists, fisheries and environmental managers, local communities and more) the model output will be diverse both in size and format. While standard formats exist and will be used, especially for much of the physical model output, other data sets are more challenging. Ways of ensuring that the INTAROS model output is uploaded to established long-term data repositories, like NMDC, are been sought.

A few examples of the many INTAROS model applications are given below:

- The ecosystem models NORWECOM and ATLANTIS will be employed to integrate data from the Barents Sea, originating from a range of in-situ and remote sensing platforms and geographically different locations. This includes in situ hydrographical and biological (NMDC/IMR, ICES) and earth observation (Copernicus CMEMS) data. Output: Integrated physical and biological marine data.
- In situ and remote sensing observations will be combined with ice-ocean models though data assimilation methods. Output: Ice-Ocean reanalysis geernated with data assimilated from in situ and remote sensing observations.
- Integration of satellite data from the Copernicus programme (Sentinel-1), enhanced in situ observations of surface mass balance and mass loss, and innovative modelling to deliver the contribution of mass losses from the Greenland ice sheet and Svalbard glaciers and ice caps to sea-level rise. Output: Calculated solid and fluid ice discharge and total mass loss.

# 7. Resources for data preparation and distribution

The costs of making datasets and products obtained within the INTAROS project available in standard formats with ample metadata, and deposit these in an approved data repository, are eligible costs under the Horizon 2020 Grant Agreement. The respective data owners in the INTAROS consortium are responsible for ensuring that their datasets and products are uploaded to one of the existing data repositories recommended by the assessment in WP2. In general, all data and products obtained within INTAROS should go into a secured and long-term storage. In cases where the cost and potential value for long-term storage is questioned (e.g. erratic data due to instrument failure, grave malfunction or loss), the data provider, in dialogue with and with approval of the INTAROS Executive Board, can decide not to save the data. Such incicents must be recorded in the project's progress reports to document why data are not stored.

Open source tools can help scientists generate metadata and data in standard formats, such as Rosetta, GDAL (Geospatial Data Abstraction Library) and NetCDF utilities. Widely used programming languages, such as Python, MATLAB and R, offer libraries that can be used to write customised format converter tools. Some web sites with training material for metadata and data preparation include:

- DataONE Data Management Planning: https://www.dataone.org/data-management-planning
- DMPTool: https://dmptool.org/
- NMDC and NorDataNet use of Rosetta: http://tomcat.nersc.no/rosetta/
- GDAL documentation and tutorials: https://www.gdal.org/
- NetCDF utilities: ncgen and ncdump
- NetCDF/CF validation: http://puma.nerc.ac.uk/cgi-bin/cf-checker.pl
- Global Change Master Directory (GCMD) Keywords
- Links to Training and reference materials compiled by DCC: http://www.dcc.ac.uk/training/training-and-reference-materials

A dataset prepared in NetCDF format can be made publicly available using data publishing tools like the Thredds Data Server (TDS). In November 2018, a webinar was held to exchange experience and build competence in publishing scientific data using TDS. This activity will be followed up by a Data Management Training Workshop at the upcoming INTAROS General Assembly in January 2019 and complemented with training material prepared as part of the Useful Arctic Knowledge (UAK) Winter School held in Longyearbyen, Svalbard, 3-7 December 2018. Training material prepared for these events will be used to build competence in data management within the INTAROS consortium, to facilitate preparation of INAROS datasets in standard data formats with ample metadata compliant with established standards.

An essential part of a Sustainable Arctic Observation System (SAOS) is long-term and secure preservation of observations and derived products. Monitoring of resources used for preservation of data and products within INTAROS will provide input to the estimation of resources needed for SAOS. This is an important part of the roadmap for SAOS to be developed in WP 1 (Requirements and strategy for Pan-Arctic Observing Systems).

# 8. Long-term data preservation and curation

Datasets collected, and data products generated during the INTAROS project will be stored in established data repositories with secured funding for long term preservation and curation. The list of recommended data repositories includes (in alphabetical order):

- **CERSAT** – a data repository of satellite-based products, e.g. Arctic sea ice drift.
- **Coriolis GDAC** (Global Data Assembly Centre) – for glider data (e.g. ocean temperature, biogeochemistry, and acoustic data).
- **European Plate Observing System (EPOS) RI** – a distributed network of seismic stations across Europe, and a data infrastructure for long term storage of these data.
- **Exchange of Local Observations and Knowledge for the Arctic (ELOKA)** – Global repository founded for the purpose of making community-based observing datasets publicly available while respecting sensitivities specific to the management of data related to Indigenous knowledge.
- **Global Runoff Data Centre (GRDC)** – provides access to data from the Arctic-HYCOS observing system (daily and monthly gauged river discharge data from a selection of stations operated by the national hydrological services (NHS) in the Arctic Council member states).
- **Greenland Ecosystem Monitoring Programme (GEM) database** (http://data.g-e-m.dk/) for data from the coastal Greenland observations.

- **Integrated Carbon Observation System (ICOS) RI** - a distributed network of collection of carbon data across Europe, and a data infrastructure for long term storage of these data.
- **NOAA CMDL database and WDCGG (World Data Centre for Greenhouse Gases)** at https://gaw.kishou.go.jp/ for GHG measurements on NE Siberia
- **Nordicana D** collection at www.cen.ulaval.ca/nordicanad/en_index.aspx for thermal and carbon data of permafrost in the Canadian Arctic.
- **Norwegian Marine Data Centre (NMDC)** – a national data infrastructure with a distributed network of providers holding data for ocean areas of Norwegian interest, with a long-term mandate to preserve marine data for these ocean areas.
- **Norwegian node of EIDA (European Integrated Data Archive)** at http://www.orfeus-eu.org/data/eida for seismic data from OBS.
- **NSF Arctic Data Centre** for soil temperatures in the Eastern Canadian Arctic.
- **PANGAEA (Data Publisher for Earth & Environmental Science)** is a data centre holding a wide range of environmental datasets; global coverage.
- **Sea-Ice Portal of University Bremen (UB)** holds long time series of sea ice parameters derived from satellite data (e.g. ice concentration and ice thickness) for polar regions.
- **SEANOE (Sea scientific open data publication),** https://www.seanoe.org/, operated by Ifremer Department of Information Systems and Marine Data for all French data (acoustic recordings, glider data, French mooring data, etc.);
- **SIMBA data repository at FMI** at https://simba.srsl.com/fmi/.
- The **Bolin Centre of Climate Research database** for atmospheric data from the icebreaker Oden).

Most of the data and products will be open, but INTAROS will respect the need to restrict access to sensitive data, such as data collected through community-based observations. The ethical aspects of data management in INTAROS are further described in the next section.

## 9. Ethical aspects

All activities in the proposal meet the national legal and ethical requirements of the 8 Arctic countries. Specifically, the ethics requirements will address: (1) research related to humans, (2) protection of personal data, and (3) third countries.

**1. Humans.** The procedures and criteria that will be used to identify and recruit research participants for the community observing systems of INTAROS are described below.

*a) Procedures:* The recruitment will be undertaken by the project in cooperation with representatives of the local communities. Participation will be entirely voluntary. The community members that the project will invite to participate in project workshops, discussions, sharing of knowledge, observations and experience are seen as 'co-creators' (and not as 'objects' of research).

*b) Criteria:* The project together with representatives of the local communities will identify who among the local communities may be interested in participating. The criteria will be the knowledge and experiences of the community members and their interest in participating. The aim is to obtain the participation of both men and women and different age classes, as they tend to use different natural resources and have knowledge about different topics.

**2. Protection of personal data.** Below we describe the justification for collection and processing of personal data, the procedures involved, and the information sheet.

*a) Justification for collection and processing of personal data:* Information on gender, age classes, and the experiences and interest of community members in community-based observing will be used to contact and recruit participants in relation to the community-based observing systems. Aside from this, personal information is not used in the project. No further personal data are collected.

*b) Procedures:* The project together with representatives of the local communities will identify who among the local communities may be interested in participating. The criteria will be the knowledge and experiences of the community members and their interest in participating. The aim is to obtain the participation of both men and women and different age classes (as they tend to use different natural resources and have knowledge about different topics). Information on gender, age, and their experiences and interest in community-based observing is not stored. **The project will comply with the EU directive on data protection** and with any updates it might receive during the lifetime of the project.

*c) Templates of the information sheet/informed consent:* The information sheet/ informed consent form that will be communicated to the participants is provided in Annex 2 in the Description of Action.

d) The participants will be compensated for the time they use on community-based observing. In each area, *every participant will be treated equally favourable and they will obtain the same daily compensation, irrespective of personal characteristics*, e.g. religion, gender or age. The project has set aside funding for compensating community members for lost work-time for the time they spend on community monitoring activities in the two focal communities of community-based observing in INTAROS: Longyearbyen in Svalbard and Disko Bay in Greenland.

**3. Third countries.** The ethical standards and guidelines of Horizon2020 will be rigorously applied, regardless of the country in which the research is carried out. For the activities in Svalbard, the project will need a research permission from the local authorities. The project will obtain the research permission from the local authorities in Svalbard before the start of the research. The project will keep these permissions available at any time if needed for the European Commission representatives or for ethics reviewers.

## 10. References

Cupola, P., Earley, S., & Henderson, D. (2014). DAMA-DMBOK2 Framework. Online at https://www.dama.org/sites/default/files/download/DAMA-DMBOK2-Framework-V2-20140317-FINAL.pdf (retrieved 17 Oct 2019).

Eaton, B., J. Gregory, B. Drach, K. Taylor and S. Hankin, NetCDF Climate and Forecast (CF) Metadata Conventions, Version 1.6, 5 December, 2011.

European Commission, 2016. H2020 Programme Guidelines on FAIR Data Management in Horizon 2010. Version 3.0. 26 July 2016.

Lygre, K. and H. Sagen, 2017. INTAROS Deliverable 8.2 Gender and diversity action plan.

OceanSITES (2010) OceanSITES User's Manual NetCDF Conventions and Reference Tables. Version 1.2. June 29, 2010.

Olsen, L.M., G. Major, K. Shein, J. Scialdone, S. Ritz, T. Stevens, M. Morahan, A. Aleman, R. Vogel, S. Leicester, H. Weir, M. Meaux, S. Grebas, C.Solomon, M. Holland, T. Northcutt, R. A. Restrepo, R. Bilodeau, 2013. NASA/Global Change Master Directory (GCMD) Earth Science Keywords. Version 8.0.0.0.0.

Yamakawa, A., Dushaw, B., Sagen, H. and T. Hamre (2018). Data standardization for long-term underwater acoustic observation - Ocean temperature and ambient noise data in Fram Strait, NERSC Technical Report No. 391, Version 1.0, Revised March 2019.

## 11. Acronyms

| Acronym | Explanation |
|---|---|
| | |
| arcFOCE | arctic Free Ocean Carbon Enrichment |
| ADC | Arctic Data Committee (under SAON) |
| AWIPEV | French - German Arctic Research Base at Ny-Ålesund / Spitsbergen |
| AWS | Automatic Weather Station |
| CBM | Community Based Monitoring |
| CDOM | Colored dissolved organic matter |
| CMEMS | Copernicus Marine Environment Monitoring Service |
| CON | Committee on Observations and Networks (under SAON), |
| CTD | Conductivity Temperature Depth |
| DIF | Directory Interchange Format (metadata standard) |
| DMP | Data Management Plan |
| DOI | Digital Object Identifier |
| ELOKA | Exchange of Local Observations and Knowledge for the Arctic |
| EMODNET | European Marine Observation and Data Network |
| EPOS | European Plate Observing System (infrastructure) |
| FTP | File Transfer Protocol |
| GDAL | Geospatial Data Abstraction Library |
| GEO | Global Earth Observation |
| GEOCRI | GEO Cold Regions Initiative |
| GeoJSON | JSON format for geographic data structures |
| GEOSS | Global Earth Observation System of Systems |
| GeoTIFF | metadata standard allowing georeferencing information within a TIFF |
| GIS | Greenland Ice Sheet |
| HDF-EOS | Hierarchical Data Format - Earth Observing System |
| HTTP | HyperText Transfer Protocol |
| IAOOS | Ice - Atmosphere - Arctic Ocean Observing System |
| iAOS | Integrated Arctic Observation System (platform) |
| INTAROS | Integrated Arctic Observation System (project) |
| ICOS | Integrated Carbon Observation System (infrastructure) |
| ISO 19115:2013 | "Geographic Information — Metadata" (standard) |
| ISO/TC 211 | International Organization for Standardization/Technical Committee 211 |
| JSON | JavaScript Object Notation |
| NASA | National Aeronautics and Space Administration (USA) |
| NetCDF/CF | NetCDF/Climate and Forecast metadata convention |
| NSF | National Science Foundation (USA) |
| OAI-PMH | Open Archives Initiative Protocol for Metadata Harvesting |
| OGC CSW | Open Geospatial Consortium Catalogue Service for the Web |
| OGC WFS | Open Geospatial Consortium Web Feature Service |

| | |
|---|---|
| OGC GeoPackage | Open Geospatial Consortium GeoPackage |
| OGC WCS | Open Geospatial Consortium Web Coverage Service |
| OGC WMS | Open Geospatial Consortium Web Map Service |
| OPeNDAP | Open-source Project for a Network Data Access Protocol |
| PISUNA | Piniakkanik Sumiiffinni Nalunaarsuineq |
| SAON | Sustained Arctic Observing Networks |
| SAOS | Sustainable Arctic Observing System |
| SDI | Spatial Data Infrastructure |
| SIMBA | Sea Ice Mass Balance Buoy |
| TDS | Thredds Data Server |
| UAK | Useful Arctic Knowledge (project) |
| WP | Work-package |
| YOPP | Year of Polar Prediction |

# Appendix A. Template for dataset descriptions.

**Table 1. Template for describing datasets in the INTAROS Data Catalogue.**

| Description | Value |
|---|---|
| Title: | Short text naming the dataset and outlining its content. |
| Abstract: | Short description of contents of dataset, how data was collected or generated, what processing steps data have undergone, quality control procedures applied, estimated uncertainty. |
| Parameter name(s): | List of parameters contained in the dataset. |
| Project/Program name(s): | Name of project(s)/program(s) that supported data acquisition. |
| Observing system: | Name of observing system that collected the data. |
| Tags: | Keywords associated with the dataset. Supports fast search. |
| License: | License under which the dataset is made available. |
| Organisation: | Name of the organisation owning the dataset. |
| Source: | Online access point for dataset, e.g. URL to a Thredds server. |
| Version: | Version of the dataset. |
| Principal Investigator: | Name of principal investigator(s). |
| PI e-mail: | Email address of principal investigator(s). |
| Data Curator: | Name of data curator for the dataset. |
| Data Curator E-mail: | E-mail address of the data curator. |

# Appendix B. Information and consent form for participants in community-based observing

(Text below can be provided in the national language for the INTAROS local communities.)

Monitoring in (..name of area..) of natural resources by local people for improved management

**Background.** The climate is changing. Many people in the Arctic face huge challenges. They rely on natural resources for food, income or both. Maintaining life requires observation of the environment.

Scientific knowledge of the environment is incomplete. Scientific monitoring in the Arctic is difficult. Local herders, fishers, hunters and other environmentally interested people observe the environment all year-round. Their observations and knowledge are, however, not consistently quantified, analyzed, or used for resource management.

INTAROS is a new international project. The project aims at developing an Arctic observation system to improve resource management. The observation system will be based on both community members' and scientists' monitoring. With support from the European Union, the project runs for 5 years from 2016 to 2021. This form is about the community-based monitoring activities of INTAROS. If you are interested in participating in the community-based monitoring activities of INTAROS, it is necessary for us that you fill out the form.

**What.** A simple system for community-based monitoring by herders, fishers and hunters of:

- Animals that you hunt (such as geese, ducks, foxes),
- Attacks by predators,
- Fishing activities and fishing methods,
- Quality of pasture and reindeer conditions in your area,
- Use of resources in your area by people from within and outside community,
- Changes in climate and the environment around you (snow, ice, pollution)

**Why.** Your observations, when regularly collected and shared, can be used to influence the way resources are being used in your area. To improve your livelihoods. To strengthen your rights to the use of the land. Your knowledge is important.
Examples of results may be:

- Better hunting regulations for animals that you hunt (such as geese and other)
- Better management of predators,
- Improved and more sustained access to fish,
- Improved addressing of pollution,
- Better addressing of challenges to management of pasture,
- Better acknowledgement of the rights of your own community to use of your land

**How.** Five steps:

1. The most experienced and interested herders, fishers and hunters establish a community monitoring group.
2. Notebooks. You record observations of natural resources during field trips
3. You summarize your observations in a summary format at meetings in the group every 3-month, you analyse trends, discuss challenges and management initiatives

4. You provide your summarized information to (..organisation..) and authorities
5. You present key observations at community meetings one time each year

**When.** Monitoring should only be done as part of your routine herding, hunting or fishing activities. After every field trip you note your observations in a calendar. Every three month you meet with other members of the community monitoring group. You discuss and agree on trends in natural resources. If you want, you propose management actions to (..institution..) and the authorities.

**Who.** Any local person interested in natural resources of their areas in (..municipality..) can participate. People and communities will decide on their own if they see a benefit in this and if they want to participate in the project. They will decide what they want to monitor. Participation is on a voluntary basis and people are not paid to do the monitoring; they should do it because they think it may help them sustain their resource use.

**Questions:**

1. Do you agree to participate in this activity? _____
2. Do you agree to be contacted later on, in the context of INTAROS? _____

Personal data will be treated with confidentiality. Participation will be entirely voluntary. Participants in community-based monitoring are free to leave the activity at any time. Just inform the project's contact person.

**More information:** (..contact person in the local area, institution, telephone numbers..).

# Appendix C. INTAROS model questionnaire

Template for model descriptions with examples.

**Table 2. Template for describing of models used in the INTAROS project and two examples.**

| Description | Value |
|---|---|
| Model name: | Full name and abbreviation |
| Brief description of model: | Short description of the model. |
| Use/application in INTAROS: | Short description of its use/application in the project. |
| Used in these INTAROS WP(s) and task(s): | Identify which task(s) the model is used in. |
| Theme(s) (ocean, atmosphere, sea ice, glaciology, terrestrial, communities): | List the theme(s) in which the model is used within the project. |
| Geographical area: | Name the regions the model cover in the project. |
| Grid size and time step: | List the spatial and temporal resolution |
| Driving forces (input data and forcing): | Describe the input data and forcing fields used. |
| Model period (years, season…): | Time period(s) and type of simulation (e.g. hindcast). |
| Model used in other models within INTAROS: | Yes/No. If yes, name the other model(s). |
| References (maximum 3): | List papers or reports describing the model or analysis of model results. |
| Additional comments: | Other relevant information. |
| Contact person: | Name and e-mail of contact person for the model description. |
| Contact organisation | Name of the organisation of the contact person. |

| Description | Value |
|---|---|
| Model name: | NORWECOM.E2E |
| Brief description of model: | Module-based End-to-end model where an NPZD model is coupled to IBMs for some selected main species, an ocean acidification module and a contaminant module. |
| Use/application in INTAROS: | Indices and monitoring program evaluation in the Barents Sea |
| Used in these INTAROS WP(s) and task(s): | Tasks 6.2 and 6.8 |
| Theme(s) (ocean, atmosphere, sea ice, glaciology, terrestrial, communities): | Ocean |
| Geographical area: | Barents and Nordic seas |
| Grid size and time step: | 10-20 km, 1 hour |

| | |
|---|---|
| Driving forces (input data and forcing): | Atmosphere (wind and short wave radiation), ocean (currents, water level, salinity, temperature, ice) |
| Model period (years, season…): | 1995-2017 (hindcast), 2006-2070 (downscaled climate simulations) |
| Model used in other models within INTAROS: | Not yet |
| References (maximum 3): | Hjøllo, S. et al. 2012. Modelling secondary production in the Norwegian Sea with a fully coupled physical/primary production/ individual-based Calanus finmarchicus model system. Marine Biology Research, 8: 508–526<br><br>Utne, K. et al. 2012. Estimating consumption of Calanus finmarchicus by planktivorous fish in the Norwegian sea using a fully coupled 3d model system. Marine Biology Research, 8: 527–547 |
| Additional comments: | |
| Contact person: | Morten D Skogen, morten@imr.no |
| Contact organisation | Institute of Marine Research, Norway |

| **Description** | **Value** |
|---|---|
| Model name: | ElmerIce |
| Brief description of model: | Parallel finite-element model based on the open-source multi-physics code Elmer. |
| Use/application in INTAROS: | Model-based demonstration of calculations of ice discharge from selected glaciers to the ocean, aimed to predict the contribution of glaciers to sea level rise. |
| Used in these INTAROS WP(s) and task(s): | Tasks 6.4.3 and 6.4.4 |
| Theme(s) (ocean, atmosphere, sea ice, glaciology, terrestrial, communities): | Glaciology |
| Geographical area: | Northwest of Greenland and South of Svalbard |
| Grid size and time step: | 50-100 m, 1 day-1 week |
| Driving forces (input data and forcing): | Glacier geometry, Surface mass balance, Surface velocity, submarine melt rate |
| Model period (years, season…): | Melt season |
| Model used in other models within INTAROS: | No |
| References (maximum 3): | Gagliardini, O. et al. 201). Capabilities and performance of Elmer/Ice, a new-generation ice sheet model. Geosci. Model Dev. 6, 1299–1318.<br><br>Otero J. et al. 2017. Modeling the Controls on the Front Position of a Tidewater Glacier in Svalbard. Front. Earth Sci. 5:29. |
| Additional comments: | |
| Contact person: | Jaime Otero, jaime.otero@upm.es |
| Contact organisation | Universidad Politécnica de Madrid |

---------- END of DOCUMENT----------

INTAROS

This report is made under the project
**Integrated Arctic Observation System (INTAROS)**
funded by the European Commission Horizon 2020 program
Grant Agreement no. 727890.

Project partners: