INTAROS – Integrated Arctic Observing System (iAOS)

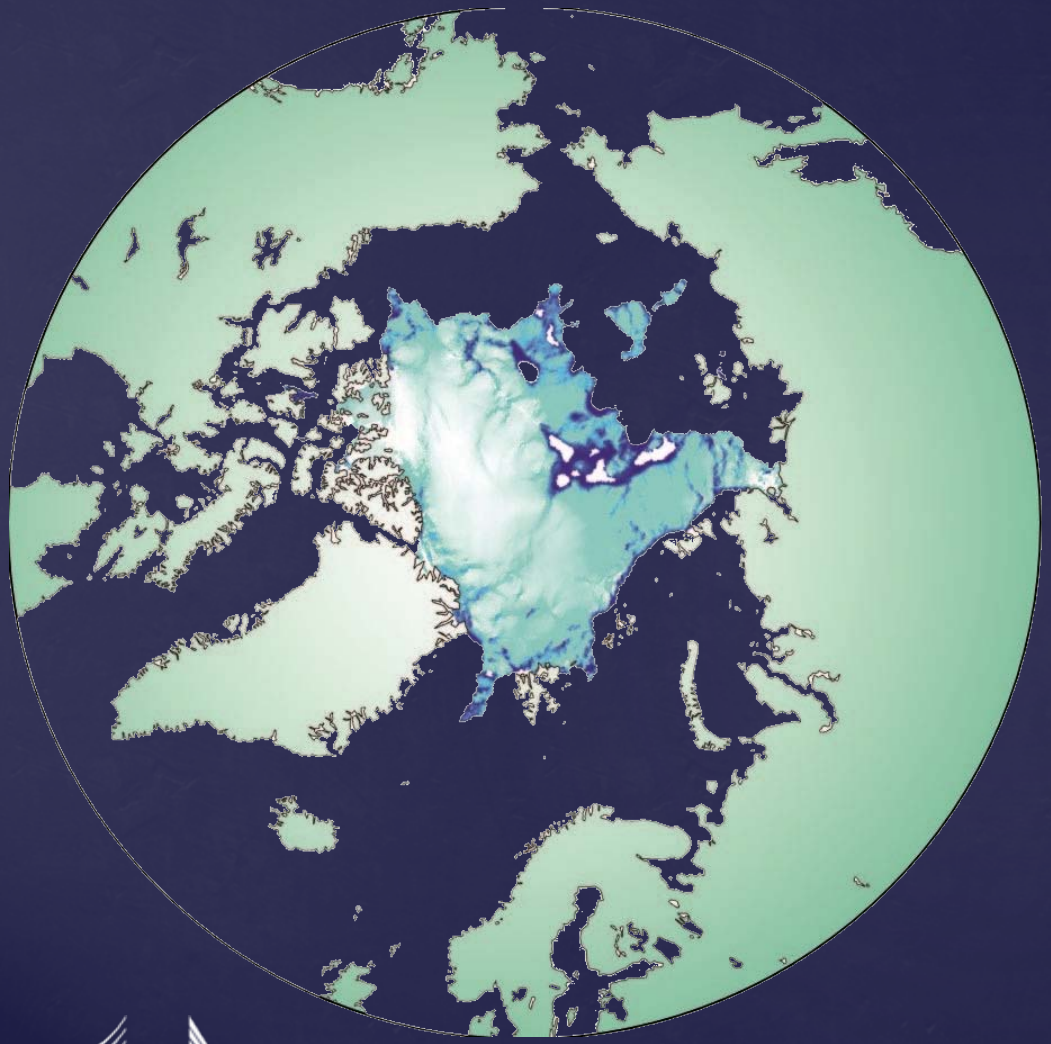
Data Management in iAOS

Torill Hamre
Nansen Environmental and Remote Sensing Center

NERSC

- INTAROS introduction
- Approach to data management
- Data integration and management

INTAROS overall objective



Is to develop an efficient integrated Arctic Observation System by

- extending,
- improving and
- unifying

existing and evolving systems in the different regions of the Arctic

European Commission

Approach to data management

- Data management is intervowen in the INTAROS work plan
- Task 1.4 Data management and data governance framework
  - DMP + FW
- Task 2.2 Exploitation of existing data towards improved data products
- Task 2.3 Compilation of data products from distributed databases and observatories for integration in iAOS
  - DM for data and products from existing databases
- Task 3.1 -3.5 Area specific observations
  - DM for *new* in situ data
- Task 4.4 Make community-based observations accessible for iAOS
- WP5 Data integration and management
- WP6 Applications of iAOS towards Stakeholders

INTAROS WP5
Data integration and management

Pedro Gonçalves, Terradue Srl (lead)

Torill Hamre, NERSC (co-lead)

WP5 main objectives

- To integrate multidisciplinary and distributed data repositories into a scalable and resilient Pan-Arctic observing system, iAOS
- To offer seamless access to observations and derived parameters provide a set of tools for data analysis, transformation and visualization of spatiotemporal datasets

European Commission

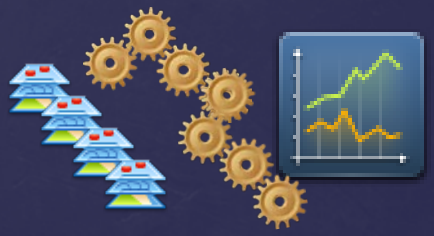
WP5 approach

- build on present observing systems developed over several years and operated with funding from countries and international agencies to ensure that the iAOS is maintained as a sustainable platform
- leverage on existing data repositories for storage and curation of new data collected during INTAROS and provide seamless access to distributed data repositories through a common ICT platform with embedded processing and analysis tools

WP5 specific objectives

Provide the underlying infrastructure for the Integrated Arctic Observing System and associated tools for IT management and support

Provide a framework to ease the discovery and retrieval of data from existing spatial data infrastructures, and a set of tools for data analysis, transformation and visualization



Develop new geo-statistical methods for interpolation of spatiotemporal datasets, and provide a set of tools for data analysis, transformation and visualization

Process new observations from WP2-4, and store the generated datasets in an IAOS enabled repository

WP 5 tasks

Task 5.0 Scientific and operational coordination

Task 5.1 System requirements and architecture consolidation

Task 5.2 IAOS platform deployment and operation

Task 5.3 Integrate data from existing repositories into iAOS

Task 5.4 Development of geo-statistical methods for data integration

Task 5.5 Integration of new processing services

Task 5.6 iAOS portal development

Task 5.7 Synthesis of IAOS infrastructure deployment and operation

WP5 Data integration and management

Task 5.1 System requirements and architecture consolidation

- analyse interfaces of the databases selected in WP2
- define the data search and access services in iAOS.
- define the iAOS essential user functionality for
  - Data discovery,
  - View/Browse,
  - Data download,
  - Data transformation, and
  - Workflows
- include a standard format for defining processing chains
- assignment of DOIs for the newly generated datasets
- define the architecture driven by a cloud infrastructure

WP5 Data integration and management

Task 5.2 IAOS platform deployment and operation

- provide the underlying infrastructure for iAOS
- offer a Cloud Developer Sandbox for algorithm and service developers
- include cloud orchestration, storage virtualisation, Virtual Machine (VM) provisioning, scaling & cloud bursting
- operate on a Cloud Platform under a virtualized data center
- define, execute and manage services of interconnected VMs

European Commission

WP5 Data integration and management

Task 5.3 Integrate data from existing repositories into iAOS

- provide a framework to ease the discovery (i.e. search) and retrieval (i.e. access) of data (remote-sensing, in situ)
- use best practices for search services using OpenSearch
- facilitate the aggregation of results between disparate data providers
- selected tools to support data access to the relevant archives will be made available on the Cloud Sandbox environment
- define the roadmap to support a federated cloud solution for existing data infrastructures

European Commission

WP5 Data integration and management

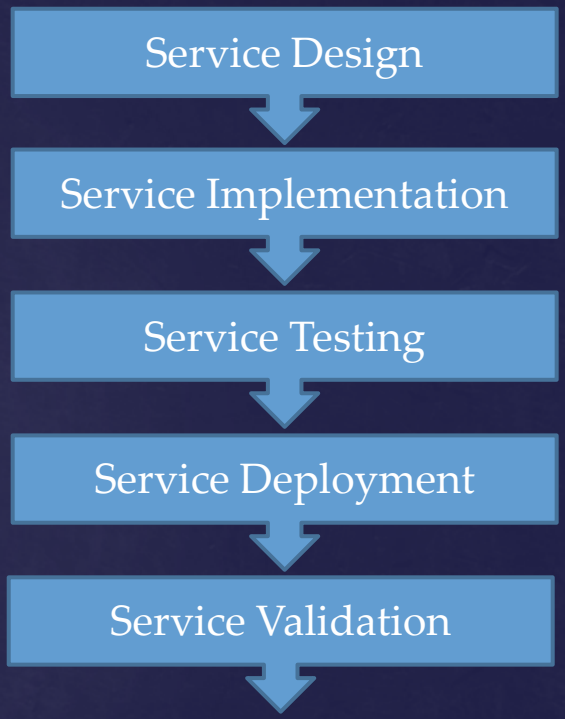
Task 5.4 Development of geo-statistical methods for data integration

- exploits data processing tools in the Developer Cloud Sandboxes from Task 5.2
- develop new geo-statistical algorithms for combining multi-source
- data into new data products
- implement the developed algorithms as part of a geo-statistical library RGeostats
- add functionalities to the iAOS Cloud Developer Sandbox using the GEOSLIB library to increase performance
- Make the developed algorithms available for service development in Task 5.5

WP5 Data integration and management

Task 5.5 Integration of new processing services

- exploit the data processing tools and geo-statistical algorithms
  in the Cloud Developer Sandboxes
- support the full life cycle of the integration of new processing services
- offer an environment where scientists have simultaneous access to data , tools and processing resources
- it will allow implementing services within a shared Platform as a Service (PaaS)
- selected data processing services will be developed and made available to demonstrate the iAOS platform



European Commission

WP5 Data integration and management

Task 5.6 iAOS portal development

- provide an intuitive user interface to the search, access and processing services in iAOS
- provide a joint entry point to the integrated data repositories and the developed services
- visualize retrieved multi-source data in a common map projection with basic GIS operations
- enable execution of the developed processing services

European Commission

WP5 Data integration and management

Task 5.7 Synthesis of IAOS infrastructure deployment and operation

- summarize the experiences with deploying and operating the iAOS infrastructure
- identify technical and non-technical challenges that need to be addressed to establish
  a sustainable pan-Arctic integrated observing system
- make recommendations for the roadmap for a Sustainable Arctic Observing System (SAOS)

European Commission

Summary

- INTAROS will have a strong focus on data integration and management, from initial planning to implementation in iAOS, in line with international standards on quality and metadata
- iAOS will integrate a range of existing data repositories and infrastructures, and build a layer on top for unified search and disovery, browsing and downloading, processing and analysing
- Recommendations from the design and implemenation of iAOS will feed into the Roadmap for SAOS (Sustainable Arctic Observing System)