



# The Ocean Data Dojo project and workshop objectives

Torill Hamre

Nansen Environmental and Remote Sensing Center

Ocean Data Dojo Workshop 1 - Data delivery chains in ice-ocean observing,  
Bergen, 1 November 2022





# Outline

- Svalbard Science Forum & Svalbard Scientific Grants
- Project objectives
- A generic data delivery chain
- Main elements of the workshops
- Example of data delivery chains in this workshop



# Svalbard Science Forum



- The overall objective of Svalbard Science Forum (SSF) is to **contribute to increased scientific quality in research in Svalbard**. The Forum shall contribute to good coordination of research activities in Svalbard. SSF shall **offer integrated and accessible information about all research in Svalbard**. The Forum shall further develop and strengthen collaboration between both individual researchers and research institutions in Svalbard. SSF's work shall **make it easier for researchers and institutions to obtain necessary support and guidance about infrastructure and research services** available in Svalbard. SSF shall also contribute to **place Svalbard research in a larger pan-Arctic and global perspective**.
- Members:
  - The Research Council of Norway (Chair)
  - The University Centre in Svalbard (UNIS)
  - The Norwegian Polar Institute (NPI)
  - Ny-Ålesund Science Managers Committee (NySMAC)
  - Kings Bay AS
  - Svalbard Integrated Arctic Earth Observing System (SIOS)
  - Polish Academy of Sciences (PAS), Hornsund
  - Arctic and Antarctic Research Institute of Roshydromet (AARI), Barentsburg
  - The Norwegian Meteorological Institute (MET Norway)
- Observer:
  - The Governor of Svalbard





# Svalbard Science Forum



- Instruments for implementation include among others
  - Arctic Field Grants
  - Svalbard Strategic Grant
  - Research in Svalbard (RiS)
  - Svalbard Science Conference
  - Research communities (Ny-Ålesund, Longyearbyen, Hornsund and Barentsburg)
- **Svalbard Strategic Grant (SSG)** provides **seed money to advance coordination, collaboration and data sharing** between researchers with a relevance to Svalbard.
- Focus in in Ocean Data Dojo: How can we strengthen and further develop data delivery chains in support of FAIR data management for the Svalbard region. Encourage development of data delivery chains that can be used in other regions and projects.
- URL: <https://www.forskningsradet.no/en/svalbard-science-forum/>





# Objectives

- The Ocean Data Dojo project aims to **identify and propose solutions for closing gaps in FAIR data management practices for ice-ocean observing** around Svalbard across disciplines and projects.

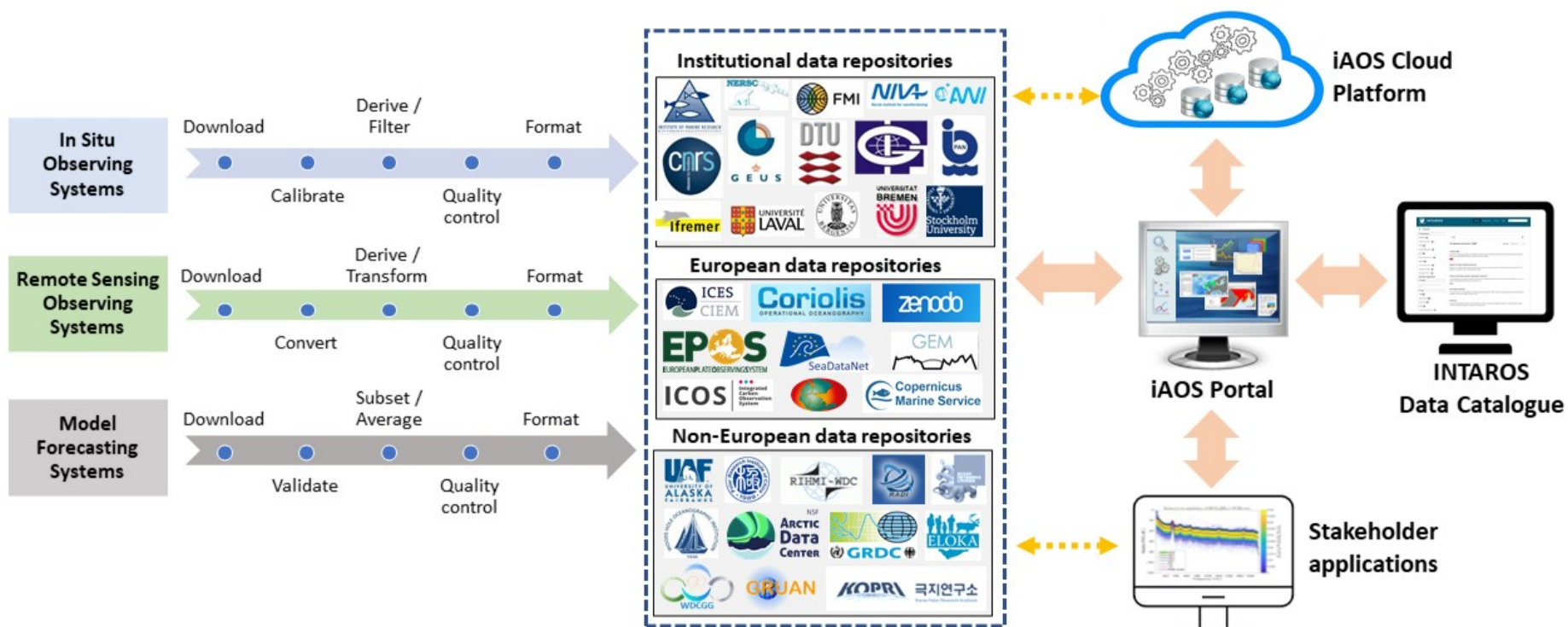
Specific objectives:

- Organise **two workshops to gather documentation about data delivery chains, standards and workflows** used in ongoing ice-ocean research projects. Furthermore, **discuss mechanisms to stimulate collaboration on FAIR data management among Norwegian initiatives.**
- **Promote the workshops and the results** through the CAPARDUS project's web site, and in relevant national and international networks for wider competence building in FAIR data management.
- **Recommend activities to stimulate collaboration** on FAIR data management around Svalbard.



# A generic data delivery chain

- Focus in Ocean Data Dojo: Selected sea ice - ocean in situ observations



INTAROS Data Delivery Chain (Hamre et al., 2021)

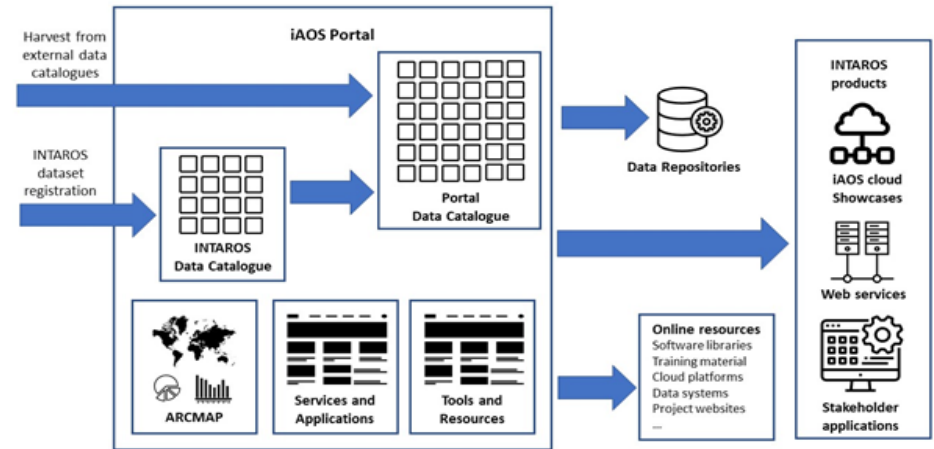




# iAOS Portal



- Data catalogue and portal platform established
- Generic portal design
  - INTAROS data catalogue
  - Portal data catalogue
  - ARCMAP
  - Services & Applications
  - Tools and resources
- Data stored in sustained repositories
- Open standard interfaces
- Portal data catalogue
  - >500 datasets harvested
  - Leveraging open APIs
  - Reusing community plugins



Major components of the iAOS Portal and their interconnections.



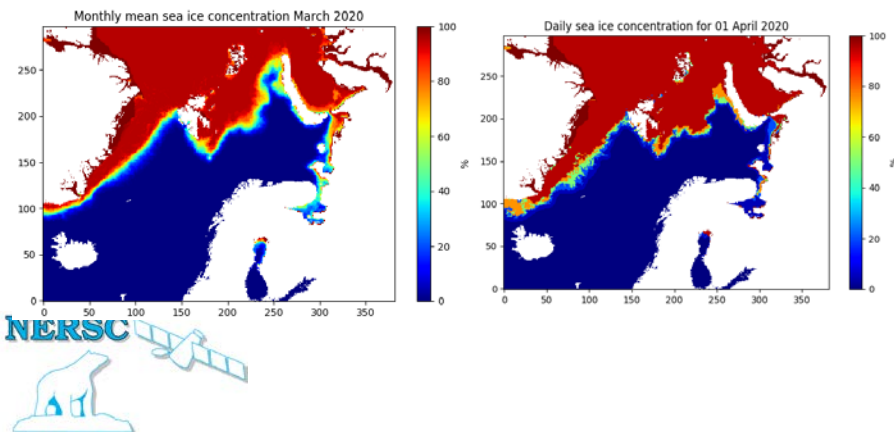
<https://portal-intaros.nersc.no/>





# iAOS Portal

- Promotion spaces for
  - Showcases/Applications
  - Cloud services
  - WPS services
  - Geostatistics libraries
- Open for new entries
- Marketplace for future iAOS developments
- Will be maintained by NERSC



## Services

INTAROS develops services for multiple user segments, including science, environmental and ecosystem management, natural hazards monitoring, risk assessment and support for mitigation planning. Services combine in situ, remote sensing and model data from a multitude of providers to provide a data product that user can apply in their daily work.

Services are developed using the iAOS Cloud Platform and Jupyter Notebook. Some examples of INTAROS services are shown below.

### Geostatistics for gridding in situ oceanographic data

This service was developed to generate ocean temperature and salinity fields for validation of climate model projections. Marine in situ observations are typically scattered in space and time, while models generate gridded data. The service uses geostatistic methods to interpolate a dispersed set of in situ point measurements to a regular grid, allowing comparison with model projections.

The service was applied to a 22 year long time series of CTD data held by the Norwegian Marine Data Centre. In total the input data amounted to 5.5 billion samples measured over 63500 positions (vertical profiles). Figure 1 shows one of the outputs from the service, a gridded field of ocean temperature for the whole time period. Read more about the geostatistics service in INTAROS Deliverable D5.6.

Temperature at 20m depth

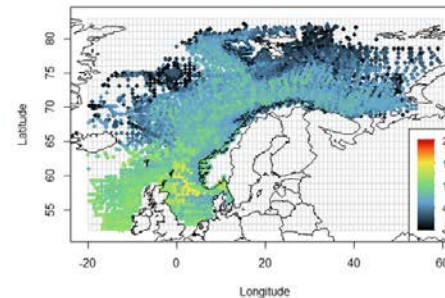


Figure 1. Base map of the whole IMR dataset – Temperature (°C) at 20m depth.

Jupyter Notebook files: RGeostats workshop  
Software packages needed:

- Latest Conda package build for RGeostats (build recipe)
- Latest Conda package build for Rintaros (build recipe)

Developer: ARMINES

### Analysis of passive acoustic data

This service processes and characterizes passive acoustic data, and produces spectrograms and noise statistics plots that can be used for analysis in combination with time series of satellite remote sensing derived parameters. It is implemented using the R version of the open source PAMGuide software package, and has extended support for new data formats (NetCDF) and data access through the OPeNDAP protocol.

The service has been tested with datasets from several sources (NERSC, CNRS, PANGAEA). Figure 2 shows an example of passive acoustic collected by CNRS in Kongstjorden, Svalbard, as part of the INTAROS field campaigns. The spectrum is dominated by low-frequency noise below 10 Hz. Local peaks around 10 Hz and 60 Hz are also seen, which could be mammal vocalization. Intermittent broad-band signals are also seen in the spectrogram. Read more about the passive acoustic service in INTAROS Deliverable D5.7.

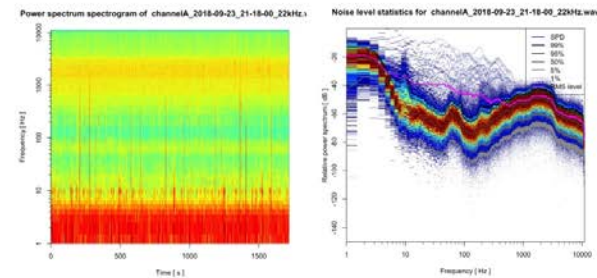


Figure 2. Examples of power spectrum spectrogram (left) and noise statistics plot (right) generated by the passive acoustic service when analysing acoustic data collected in Kongstjorden, Svalbard, during the INTAROS project.

Jupyter Notebook files: PAMGuide-R-Tutorial  
Software packages needed:

- R
- PAMGuide
- Jupyter notebook

Developer: NERSC





# INTAROS Zenodo community

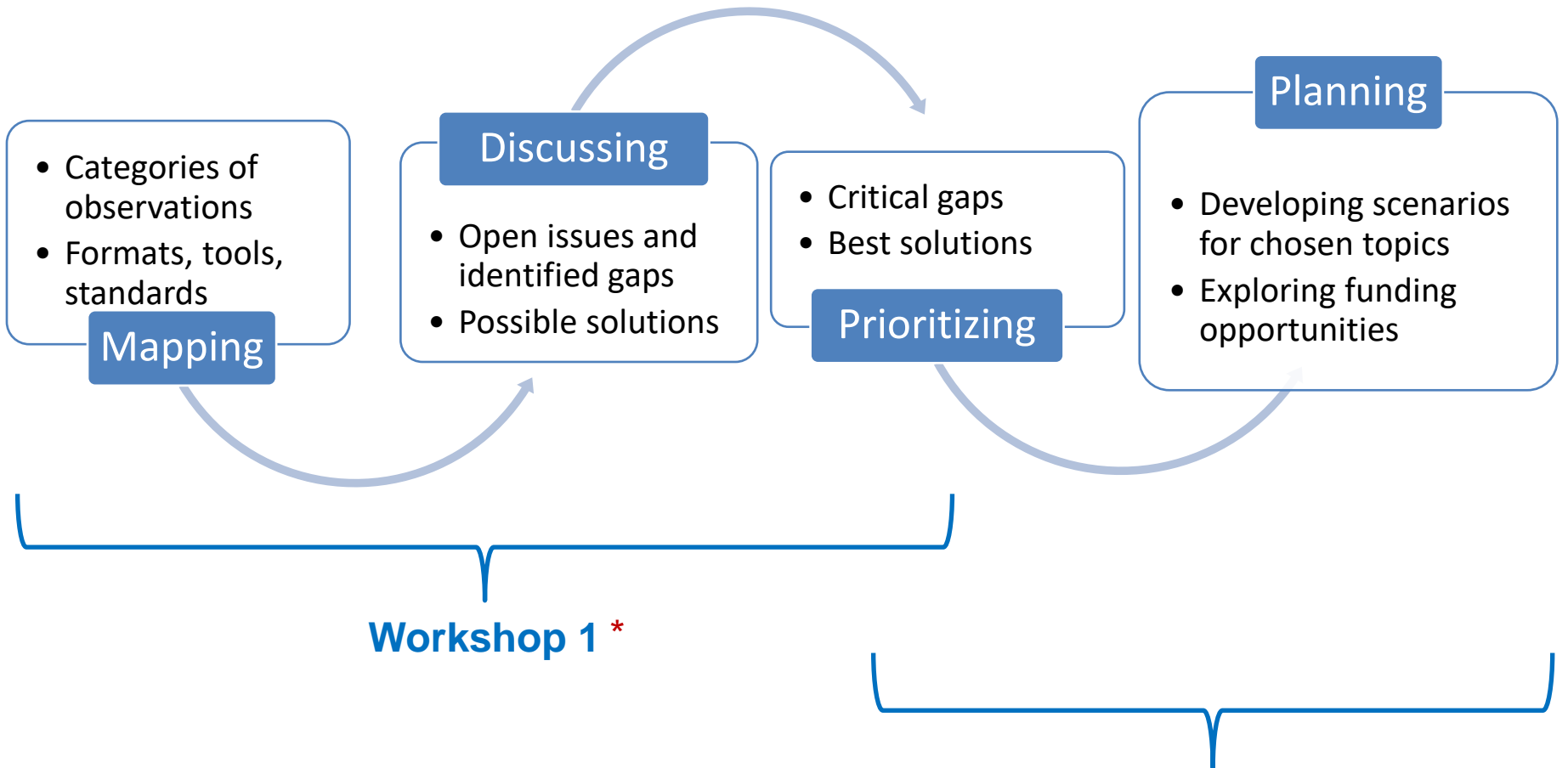
<https://zenodo.org/communities/intaros-h2020>



The Research Council of Norway



# Main element of workshops



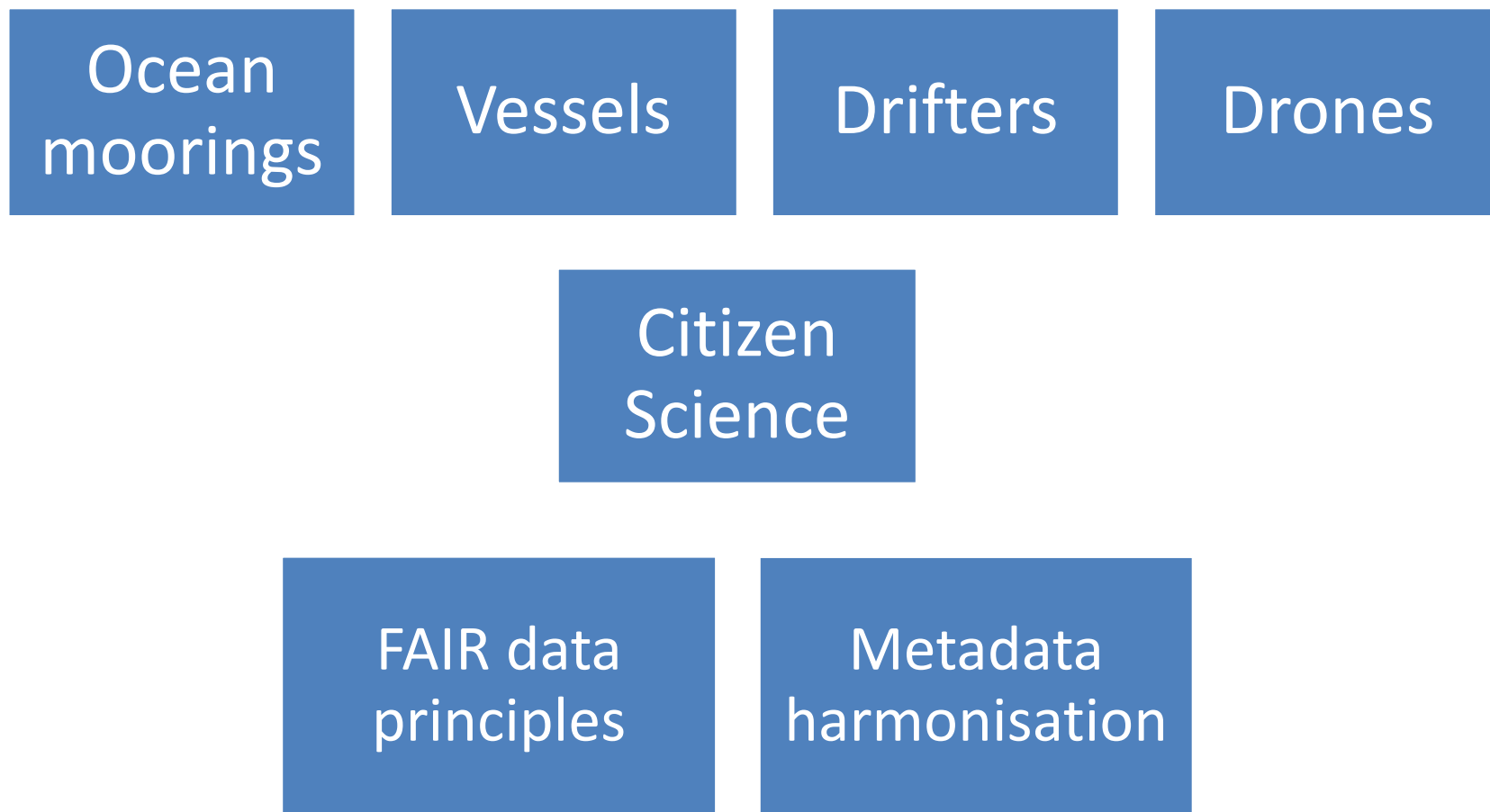
\*Workshop reports and presentations

Workshop 2 \*





# Sample data delivery chains





Questions?

