



JOINT EUROPEAN RESEARCH INFRASTRUCTURE NETWORK FOR COASTAL OBSERVATORIES

# Jerico data flow and services: Introducing the Data Management and Virtual Access system

**WP5/WP6**

**Kate Collingridge, Veronique Creach, Leonidas Perivolitas**

[www.jerico-ri.eu](http://www.jerico-ri.eu)

*JERICO-NEXT Summer School 2018  
Operational Oceanography for Blue Growth  
Malta*

# Data management

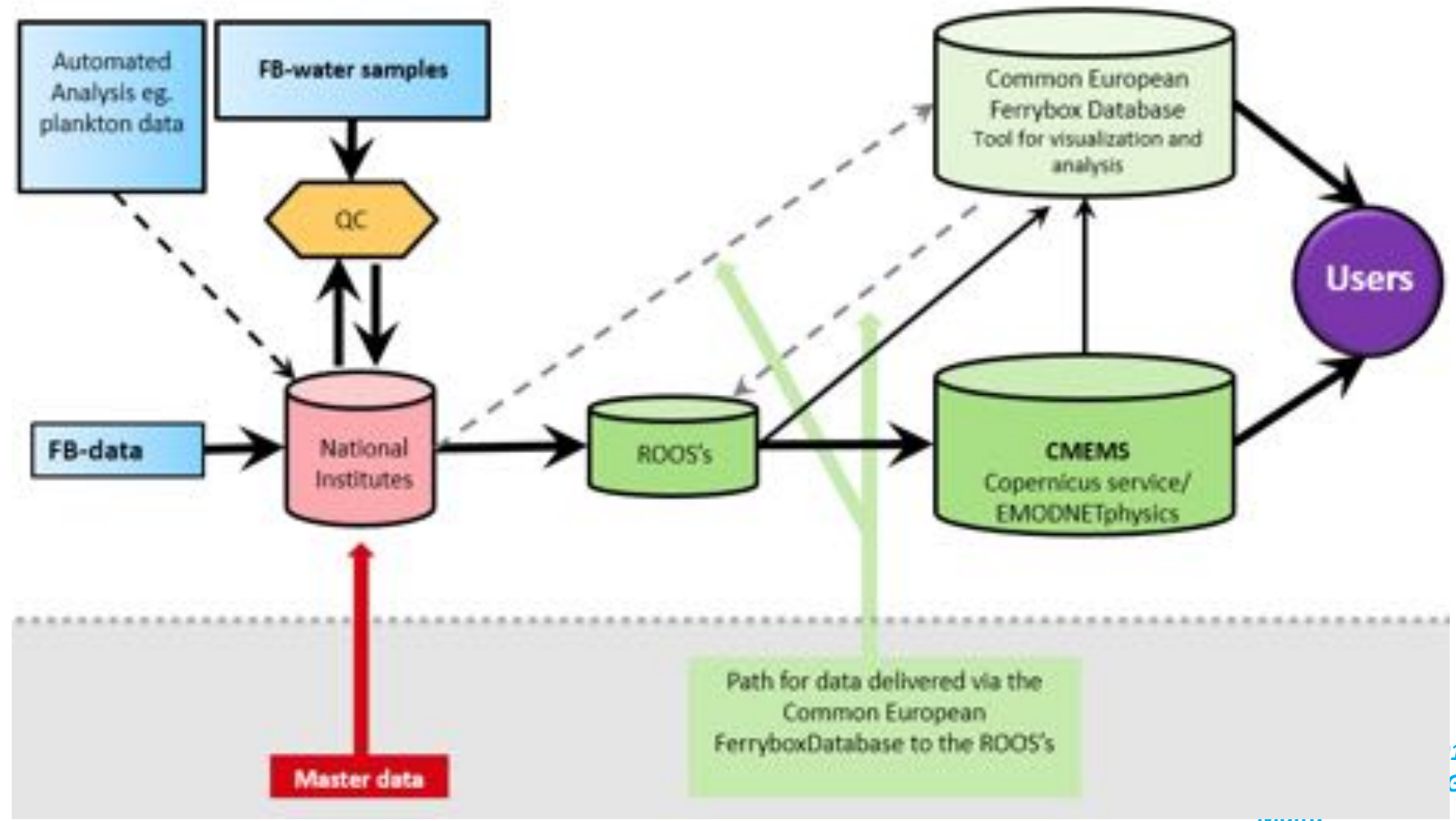


## Objectives

- Integrate biological data in the JERICO-NEXT data portfolio
- Define the project's data policy by enhancing/promoting the open access to the data
- Implement a more efficient platform registration and metadata management system
- Define properly the data flow within JERICO-NEXT
- Improve the quality of measurements derived from platforms that are widely used in coastal monitoring such as the FerryBoxes, HF Radars and Gliders
- Explore the possibility to connect the JERICO-NEXT Data system with a Virtual Access Infrastructure

# Main significant results and progress on tasks

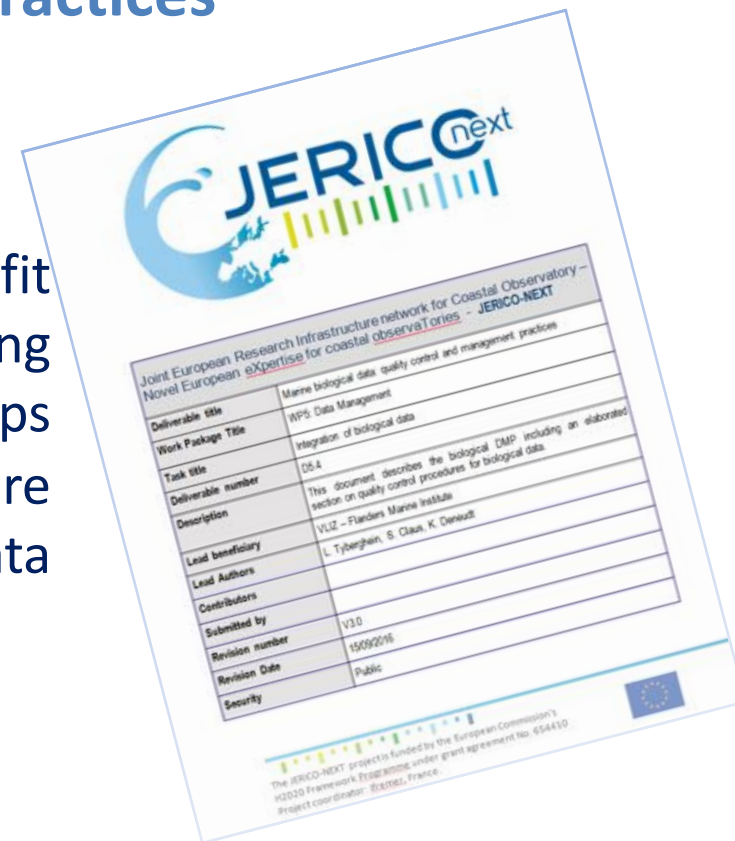
## Task 5.1: Data policy (EUROGOOS) Defining specifications for a European FerryBox data management system

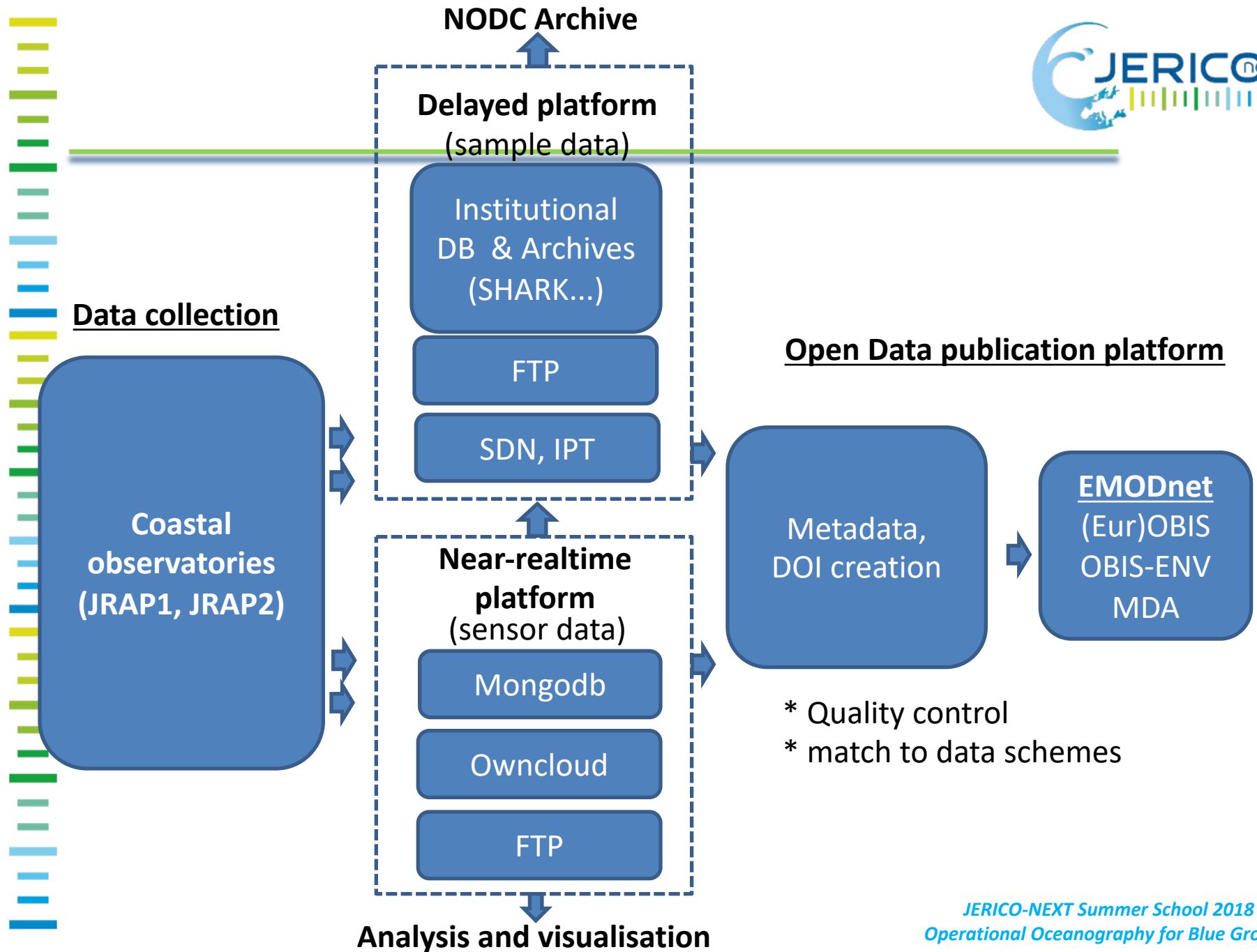


# Biological Data Integration

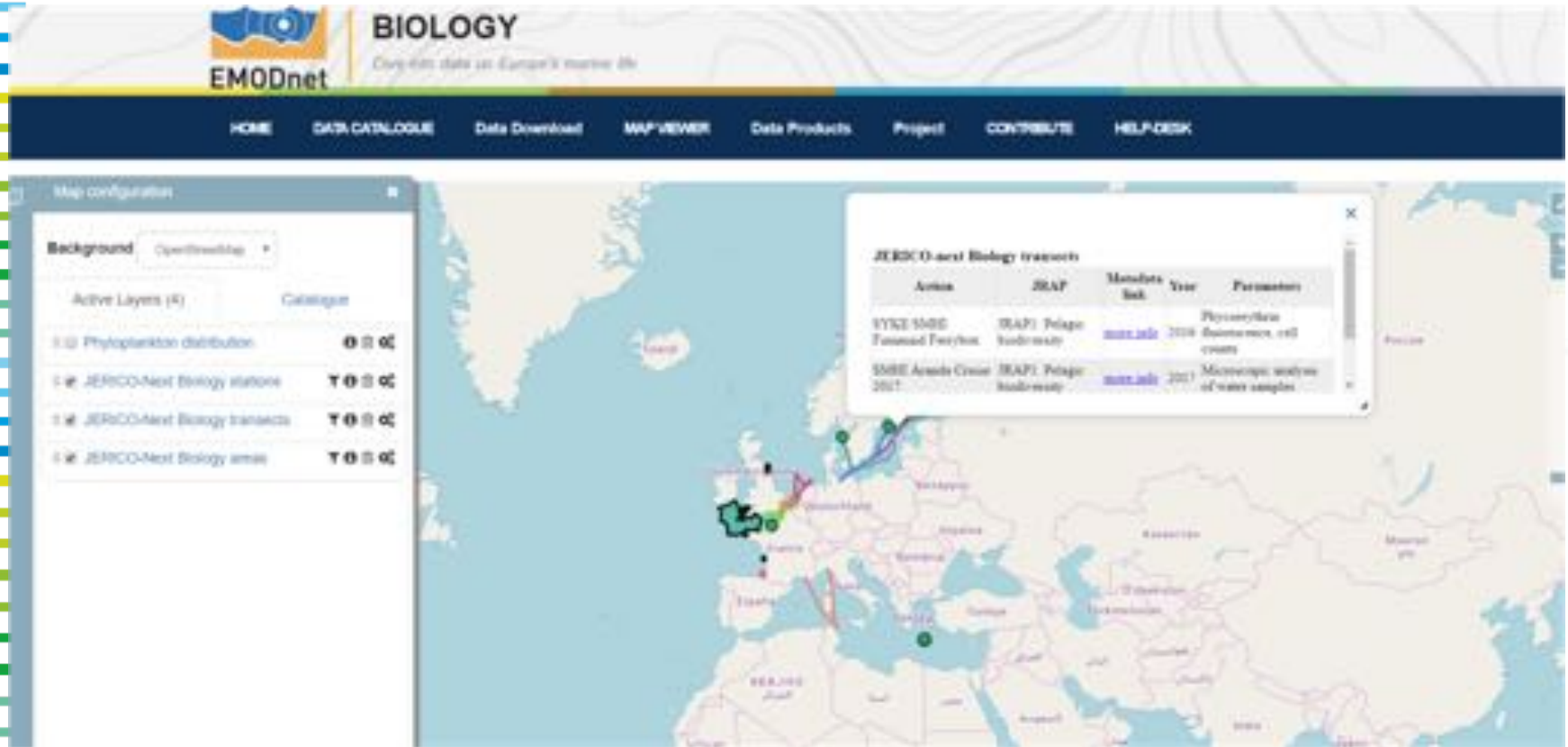
## Marine biological data: quality control and management practices

A number of data types are currently not fit for the integration in global existing databases. Therefore, specific user groups (per sensor type, e.g. flow cytometry) are formed to discuss on data types, data standards and integration.



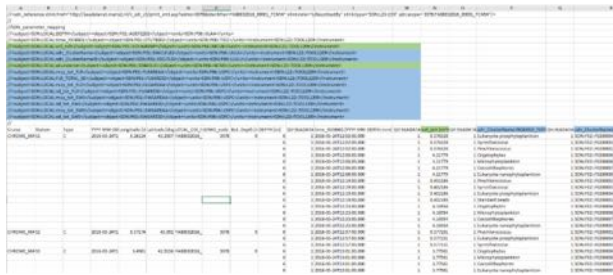


# Data inventory, metadata and data access

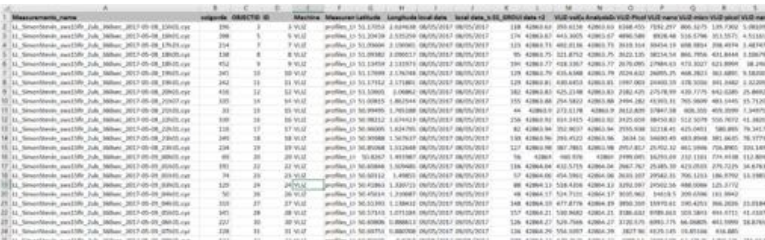


# Data processing: transformation to OBIS-Event Core format

SDN (ODV biology variant):  
CHROME\_MARS2016\_FCMW



VLIZ data from JERICO-NEXT campaign: metadata, example



- **Event Core table**
  - Event (sample) identifier
  - Latitude / longitude
  - Event date...
- **Occurrence extension table**
  - Event & occurrence identifiers
  - Flow cytometer cluster name
  - Species names...
- **Measurements or Facts extension table**
  - Event & occurrence identifiers
  - Flow cytometer cluster identifiers (BODC)
  - Abundances
  - Sampling instrument
  - Volume sampled
  - Optical measurements...

# Enhancement of Quality Control procedures for sensor based biochemical data (SMHI)



## Best practice for quality control of sensor based biochemical data

- The open source software for QC-control, the FerryBox Toolbox used on SMHI, has been used to make manual quality control of the 2016 FerryBox data from the ship Transpaper. During the work some needed features and required changes have been identified e.g. it will also to be configured to be able to deliver an output format of automatic quality controlled data in near real time via CMEMS.
- The new and updated version is ready to be disseminated to the FerryBox users of Jerico Next for evaluation. Ongoing work for making the FerryBox Toolbox freely available.



# Definition of Quality Control procedures for HF Radar data (AZTI)



## Recommendation Report 1 for HFR data

Recommendation at European level for the harmonization of HFR data management, including the following points:

- data format
- metadata structure
- QC flagging scheme
- QC tests.



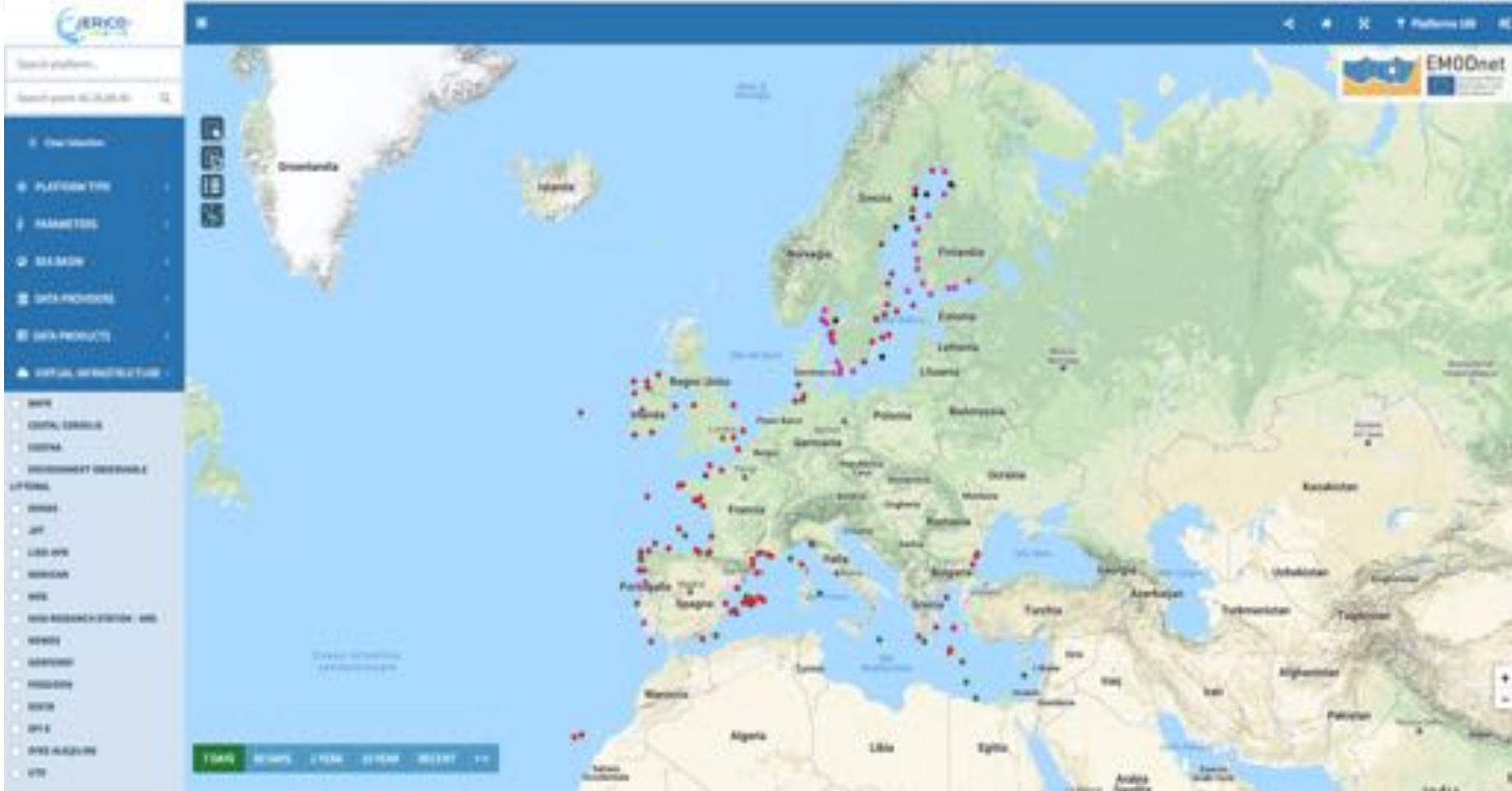
A practical document of recommendations useful for disseminating the results, with clear indications, guidelines and examples

# Scientific calibration procedures on gliders data collection



## Guidelines for the scientific correction of glider data

- Post-mission inter-calibration of glider CTD data
- Archive of delayed-mode corrected glider data.
- SOCIB glider data processing toolbox tested in NRT using SOCIB glider data.
- A first version of delayed mode field correction software is now complete
- **First draft of the deliverable** contains a description of both the method used in the field correction code and the handling of other potential errors in glider data such as spikes, temperature sensor thermal lag and bio-fouling conductivity offsets.



# What is Virtual Access?



**Virtual Access** means “*access to resources needed for research through communication networks ... Examples of virtual access activities are databases available via Internet, or data deposition services*”

It was a mandatory activity for the H2020 proposals such as JERICO-NEXT

## Aims of JERICO-NEXT Virtual Access

---

- Make oceanographic data resources available to a wide range of communities within and outside the EU.
- Provide a free of charge service integrating datasets and derived information from in situ systems (HF radar, FerryBox and fixed platforms).
- Promote the improvement of existing services and potentially the development of new services:
  - *access to in situ data can improve calibration and validation of numerical models and also provide sea truth data that improves the quality of information derived from remote sensing.*

# JERICO-NEXT providers

- 15 partners and 15 infrastructures

Id	VI provider	VI Name	VI Primary Link
1	NIVA	NorFerry	<a href="http://www.niva.no/en/miljoedata-paa-nett/ferrybox-og-satellittdata">http://www.niva.no/en/miljoedata-paa-nett/ferrybox-og-satellittdata</a>
2	HZG	COSYNA	<a href="http://www.cosyna.de">http://www.cosyna.de</a>
3	CEFAS	CEFAS DATA HUB	<a href="https://www.cefas.co.uk/cefas-data-hub/">https://www.cefas.co.uk/cefas-data-hub/</a>
4	FMI	ÜTO	<a href="http://swell.fmi.fi/Uto/latest.html">http://swell.fmi.fi/Uto/latest.html</a>
5	SMHI	SHARK data set	<a href="http://www.sharkdata.se/">http://www.sharkdata.se/</a>
6	SYKE	SYKE-Alg@line	<a href="http://www.syke.fi/en-US/Research_Development/Research_and_development_projects/Projects/Real_time_algal_monitoring_in_the_Baltic_Sea_Alcline">http://www.syke.fi/en-US/Research_Development/Research and development projects/Projects/Real time algal monitoring in the Baltic Sea Alcline</a>
7	CNR-ISMAR	LiSO-HFR	<a href="http://radarhf.ismar.cnr.it">http://radarhf.ismar.cnr.it</a>
8	HCMR	POSEIDON	<a href="http://www.poseidon.hcmr.gr">http://www.poseidon.hcmr.gr</a>
9	SOCIB	SOCIB	<a href="http://www.socib.es">http://www.socib.es</a>
10	CNRS	Environment Observable Littoral	<a href="http://www.obs-vlfr.fr/data/view/eol/surface/">http://www.obs-vlfr.fr/data/view/eol/surface/</a> <a href="http://www.obs-vlfr.fr/data/view/eol/ctd/">http://www.obs-vlfr.fr/data/view/eol/ctd/</a> <a href="http://www.obs-vlfr.fr/data/view/eol/meteo/buoy/">http://www.obs-vlfr.fr/data/view/eol/meteo/buoy/</a>
11	IO-BAS	NOMOS	<a href="http://www.bgodc.io-bas.bg">http://www.bgodc.io-bas.bg</a>
12	AZTI	BHFR	<a href="http://www.euskoos.eus/en/basque-ocean-meteorological-network/high-frequency-coastal-radars/">http://www.euskoos.eus/en/basque-ocean-meteorological-network/high-frequency-coastal-radars/</a>
13	CNRS	SPI-S (Sediment Profile Imagery Software)	<a href="https://spiarcbase.epoc.u-bordeaux1.fr/">https://spiarcbase.epoc.u-bordeaux1.fr/</a>
14	IH	MONICAN	<a href="http://monican.hidrografico.pt/">http://monican.hidrografico.pt/</a>
15	IFREMER	Coastal Coriolis	<a href="http://www.coriolis-cotier.org/">http://www.coriolis-cotier.org/</a>

# JERICO-NEXT Infrastructures

## The platforms

- 6 categories: Ferry/ship; buoy/drifter, HF Radars, station, gliders, cables

## The parameters

- 14 biogeochemical (at least):

most common: Temperature, salinity, chlorophyll/fluorescence, turbidity and (*some are unique:  $p\text{CO}_2$* )

- 4 biological:

plankton, invertebrates, fish, but also mammals (seals) and primary production (FRRf)

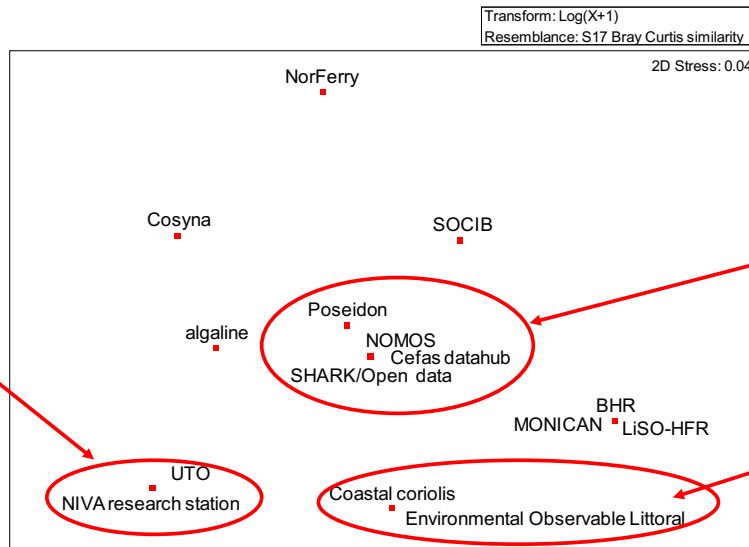
## The products/services

- Real-time (75%) visualisation and archived data sets
- Software

# JERICO-NEXT Infrastructures



platforms

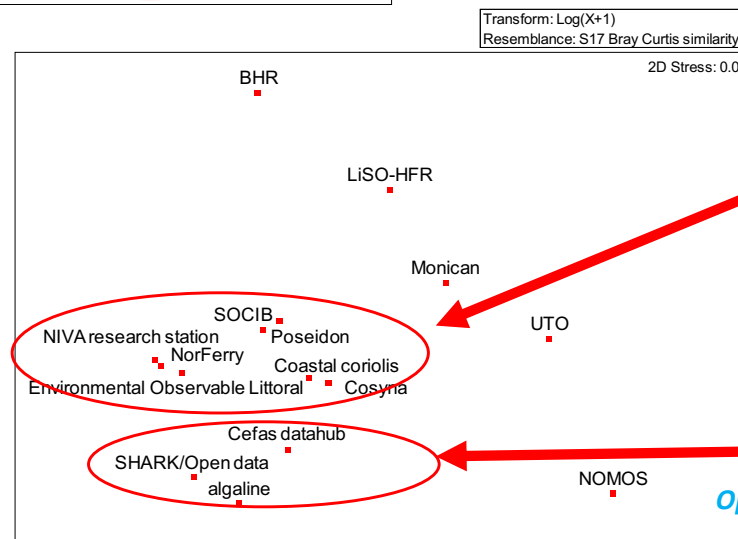


Buoy,  
Ship,  
discrete  
samples

Buoy

Station

parameters



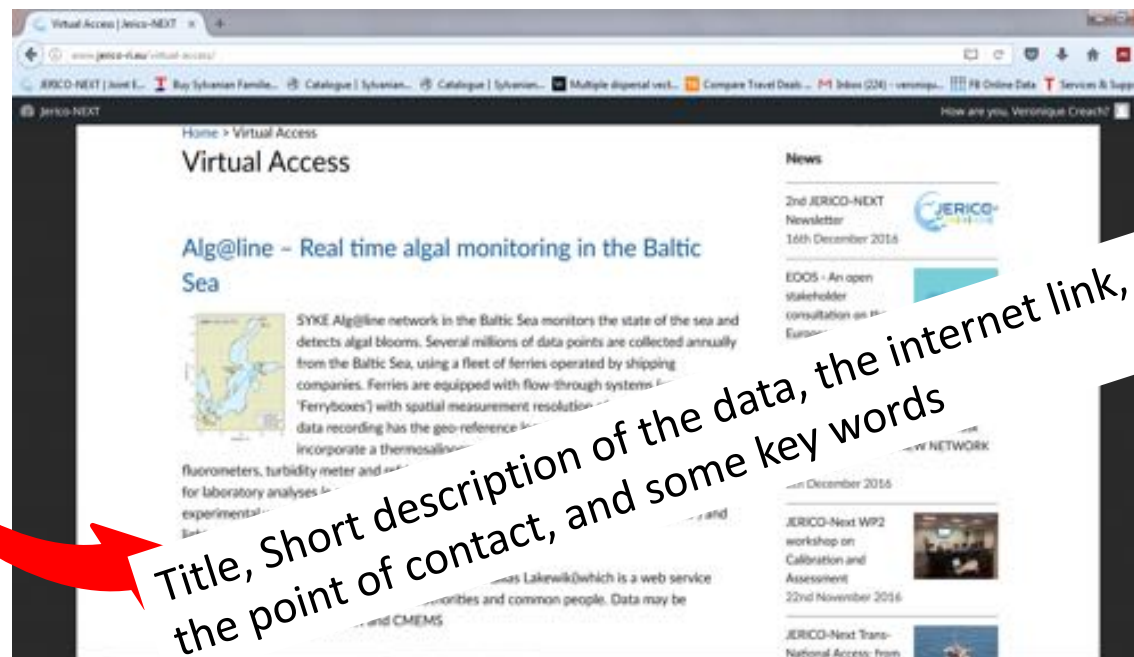
Chlorophyll/fluorecence  
Salinity  
Temperature  
Turbidity

Biological data



# Visualisation on JERICO-NEXT website

<http://www.jerico-ri.eu/virtual-access/>



Provide an interactive map showing where data and data products are available

## JERICO-NEXT Virtual Access

---

- Most/all of these data portals existed prior to JERICO-NEXT
- JERICO-NEXT is a way to advertise these services centrally and make them accessible to more users
- The JERICO-NEXT provides up to 25% of the costs of the infrastructure over the 4 years of the project

**So... What do we expect in return?**

## JERICO-NEXT Virtual Access

---

- Some form of analytics to know that the service is being used
- Development of the service – provide new products, improve services, add data DOIs
- Periodic assessments reported to European Commission

# Virtual Access Assessment

Assessment based on the quantity and diversity of the web based traffic

- Repeat visitors over a period
- Average length of time per visit
- Number of Downloads
- *Locations of enquiries*

PiwiK, AWStats  
Google analytics

*Can show the user coming from JERICO-NEXT website*

# Virtual Access Assessment

Assessment based on the utility of the data and how the data is used across European and international frameworks.

- where the data goes
- who is using it
- what other internal or external functions are improved or elevated due to the information contained on the website.

*Examples will include data being accessed by national and international modelling processes, validation of earth observation data and inputs into European data streams*

When the users are known, get information about the contribution of the data in their product development (only for some VA providers)

# Virtual Access Assessment

---

Assessment based on the integration of data services reported under the VAI

*ie: Follow the publications (DOI reference) referring of the use of the data from the VA providers.*

→ T5.8: Linking JERICO-NEXT activities (WP5/WP6) to a Virtual Access infrastructure

# JERICO-NEXT Virtual Access



## European DATA and metadata portals



## USERS

Scientists

Policy makers

Industry

Public

education

## VA providers website

NorFerry
NIVA Research Station (NRS)
COSYNA
Cefas data Hub
Utö Atmospheric and Marine Research Station
Alga@line
Sharkdatabase and Open data
LISO-HFR
Poseidon
SOCIB Data Centre Multi Platform Observatory
Environmental Observable Littoral
National Operational Marine Observing System
BHFR
SPIArcBase
MONICAN
Coastal Coriolis
NOMOS

Task 5.8

Task 5.8

Via national regulation or directly

Task 8.9

# Release of a JERICO-NEXT catalogue as joined activity between WP5 and WP6

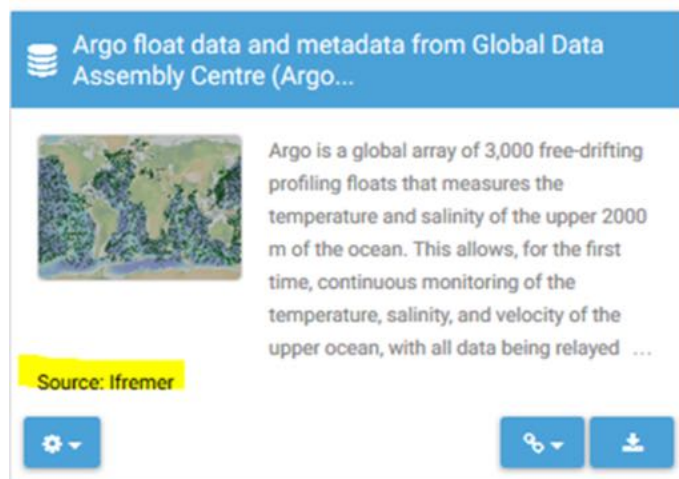
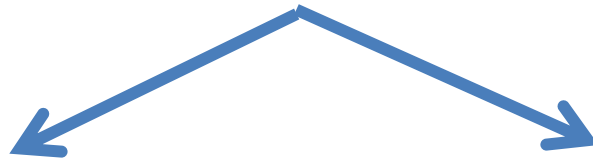
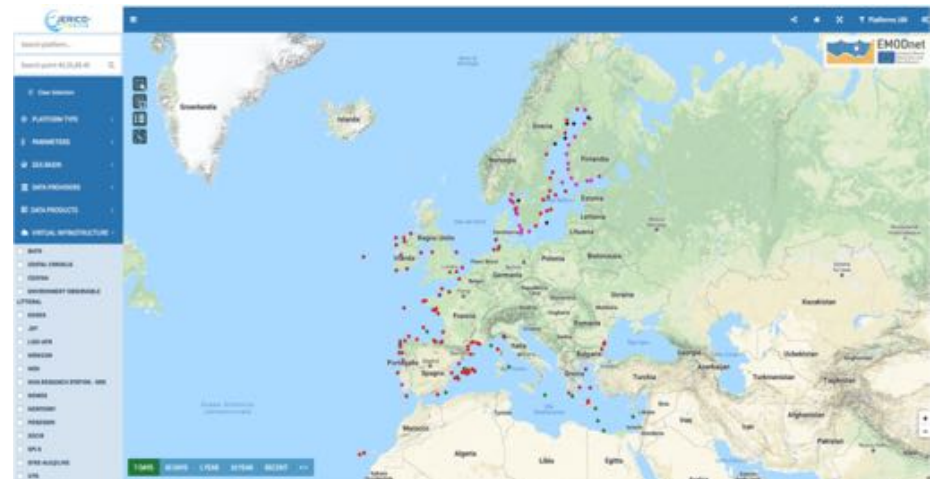


- Catalogue of products and services
- The catalogue will be implemented using Sextant, developed by Ifremer
- Sextant meets the needs of different projects on a regional, European to global scale (EMODnet, CMEMS, SDN etc.) and internationally recognised ISO and OGC standards
- <http://sextant.ifremer.fr>

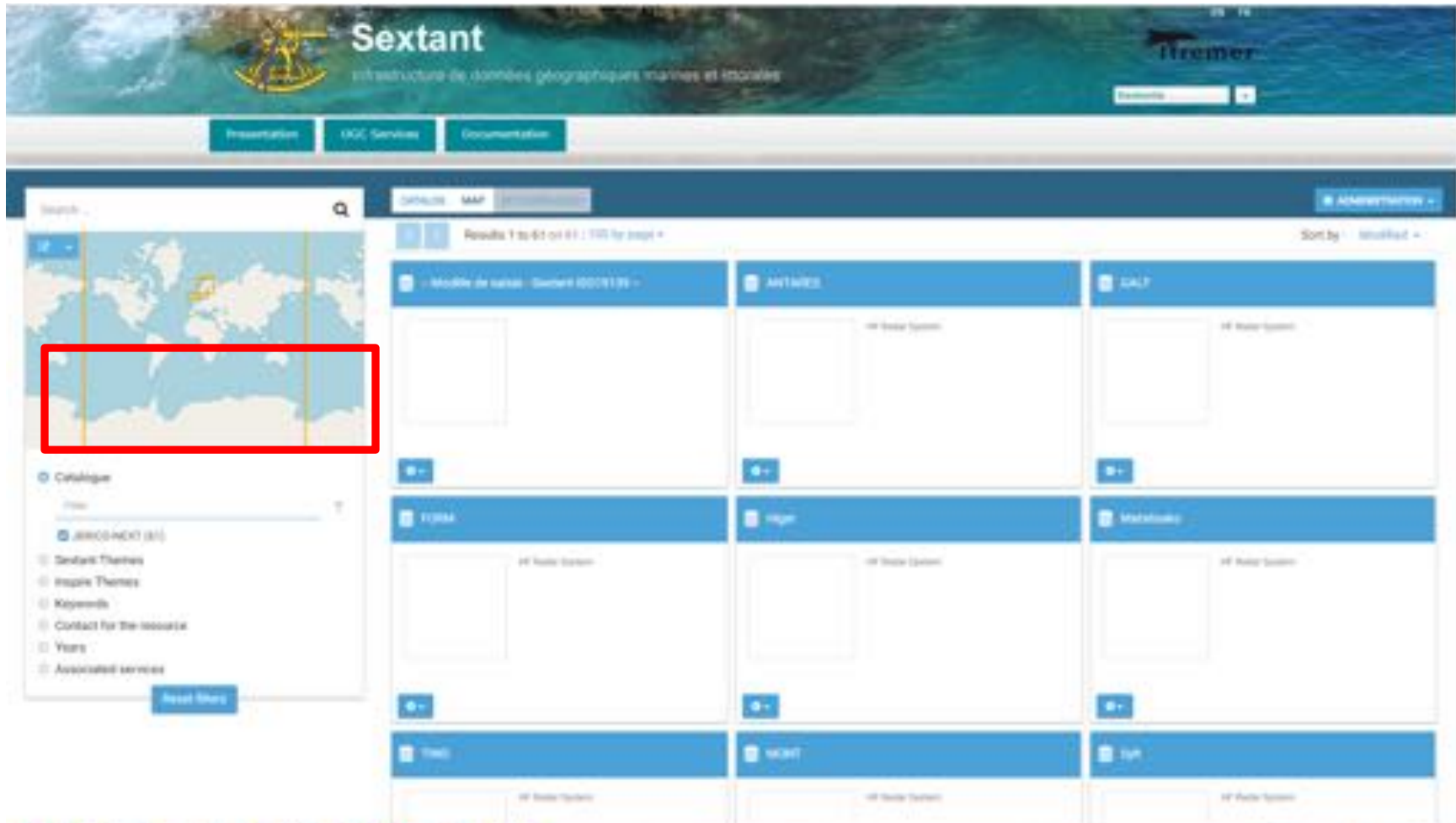


# What should be visible in the catalogue?

- **Level 1:** Virtual access infrastructures
- **Level 2:** Platform list + link to relevant data
- **Level 3:** TNA data
- **JERICO-NEXT products**

A data card with a blue header containing the text "Argo float data and metadata from Global Data Assembly Centre (Argo...". Below the header is a small satellite image of the ocean. To the right of the image is a text block: "Argo is a global array of 3,000 free-drifting profiling floats that measures the temperature and salinity of the upper 2000 m of the ocean. This allows, for the first time, continuous monitoring of the temperature, salinity, and velocity of the upper ocean, with all data being relayed ...". At the bottom left, there is a yellow box with the text "Source: Ifremer". At the bottom right, there are two blue buttons: one with a gear icon and another with a download icon.A screenshot of a web interface for JERICO-NEXT. The interface features a dark blue sidebar on the left with a menu of categories: "Classification", "PLATFORM TYPE", "INSTRUMENT", "DATA SOURCE", "DATA PRODUCTS", and "VIRTUAL INFRASTRUCTURE". The main area is a map of Europe and the Mediterranean region, with numerous red and purple dots representing data points. The map is overlaid with a grid and includes a search bar at the top left. The bottom of the interface shows a status bar with "EMOnet" and other technical details.

# Frame of the catalogue

A screenshot of the Sextant web application interface. The header features a banner with a satellite image of the ocean and a yellow sextant, with the text "Sextant infrastructure de données géographiques marines et littorales" and a "Rechercher" search box. Below the banner are three navigation buttons: "Présentation", "OGC Services", and "Documentation". The main content area is a grid of data cards. On the left, there is a search bar and a world map with a red rectangle highlighting the Mediterranean region. Below the map is a "Catalogue" sidebar with a search field and a list of categories: "JERICO-NEXT (3)", "Sextant Themes", "Inspire Themes", "Keywords", "Contact for the resource", "Years", and "Associated services". The main grid contains cards for "Modèle de cartais - Sextant 000/11/39", "ANTARES", "CALT", "TYRMA", "Migra", "MétéoMare", "TMO", "MONT", and "TUN". Each card has a blue header, a search icon, and a "Reset filters" button.

**POSITION VIA**

The data recorded by the oceanographic platform located in the buoy and the lower beam depth stations. Real time position global are received through the RTD2000C port. Platform real time location parameters of the station environment as well as atmospheric parameters in the buoy station.

**LOG GPS**

Surface current data measurements are taken. Global real time position are made available in digital format for the next 48 hours and in real time and delayed mode via a TDR2000 system which provides metadata and data access. The real time global information on real time positioning, atmospheric and

**NOVA Research Station - MRE**

The oceanographic data from the moorings at 1 m and 20 m and in the range of variables at 1 m depth is transmitted via an RTD 2000C. Other parameters that can be available are: Water fluorescence, water clarity and pH. The system will give access to the data used for doing research or for operational use.



**NOVA Research Station - MRE**

Home | Home | Home

**Summary**

**Metadata**

File identifier: 10.1000/0000-0000-0000-0000-0000  
 File character set: UTF-8  
 File name: 11-09-2016\_12:01:00  
 Metadata standard: ISO 19115-2:2009-10-01 - METADATA  
 Metadata standard version: 1.0

**Spatial representation info**

Vector spatial representation

Geometric objects

Geometric object

Geometric object type: Point

**Data identification**

Abstract: 2016 Moored research in the Boller Bay station: the data of the sea and bottom topography. Several stations of data points are collected around the Boller Bay, using a fleet of buoys equipped by drifting velocimeters. Buoys are equipped with flow-through systems (locally called "Pondoms") with depth measurement resolution of about 200 m. The data recording for the georeferenced mapping from GPS, Real-time recording a (Bathymography), underway (hydrography) and CTD (conductivity, salinity, temperature and depth) sensors, water clarity and temperature sampling and providing discrete water samples for laboratory analysis e.g. (chemistry, microbiology, genetic analysis, and experimental work). Additional flow-through sensors e.g. for primary production (PAM) and light attenuation are used periodically. Data is visualized in OGC webpage and in 2D/3D plots. Information is a web service built and maintained in cooperation by scientists and citizens people. They may be downloaded through Knowledge and Science.

Geographic representation: Vector  
 File character set: UTF-8

**Citation**

Title: 2016 Moored  
 Publication: 2016-09-11  
 Creation identifier: 2016 Moored

**Point of contact**

General

Individual name: Roger Kattar  
 Role: Project manager

**Address**

Electronic mail address: Roger.Kattar@met.rug.nl

**Descriptive keywords**

**Parameter (Quantity) Vocabulary (PQV)**

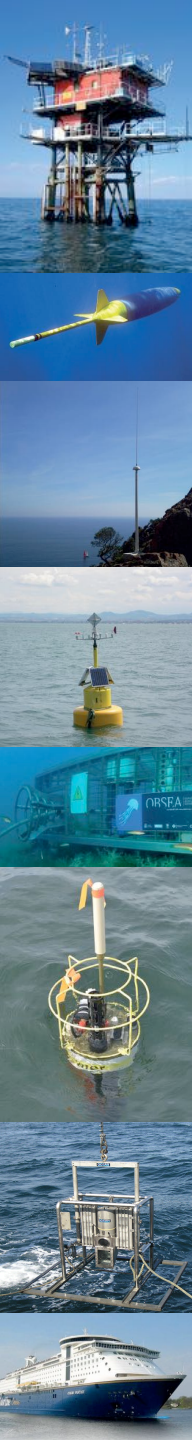
Representation of the water column  
 Atmospheric pressure measurements in water bodies  
 Surface fluorescence parameters  
 Light attenuation in the water column  
 Optical backscatter  
 Water clarity

# Virtual Access Practical



This afternoon (Tues 10<sup>th</sup>)

- Visit VA data portal – Cefas Data Hub
- Download a dataset
- Visualise the dataset using R (and shiny)



Thank you



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654410.