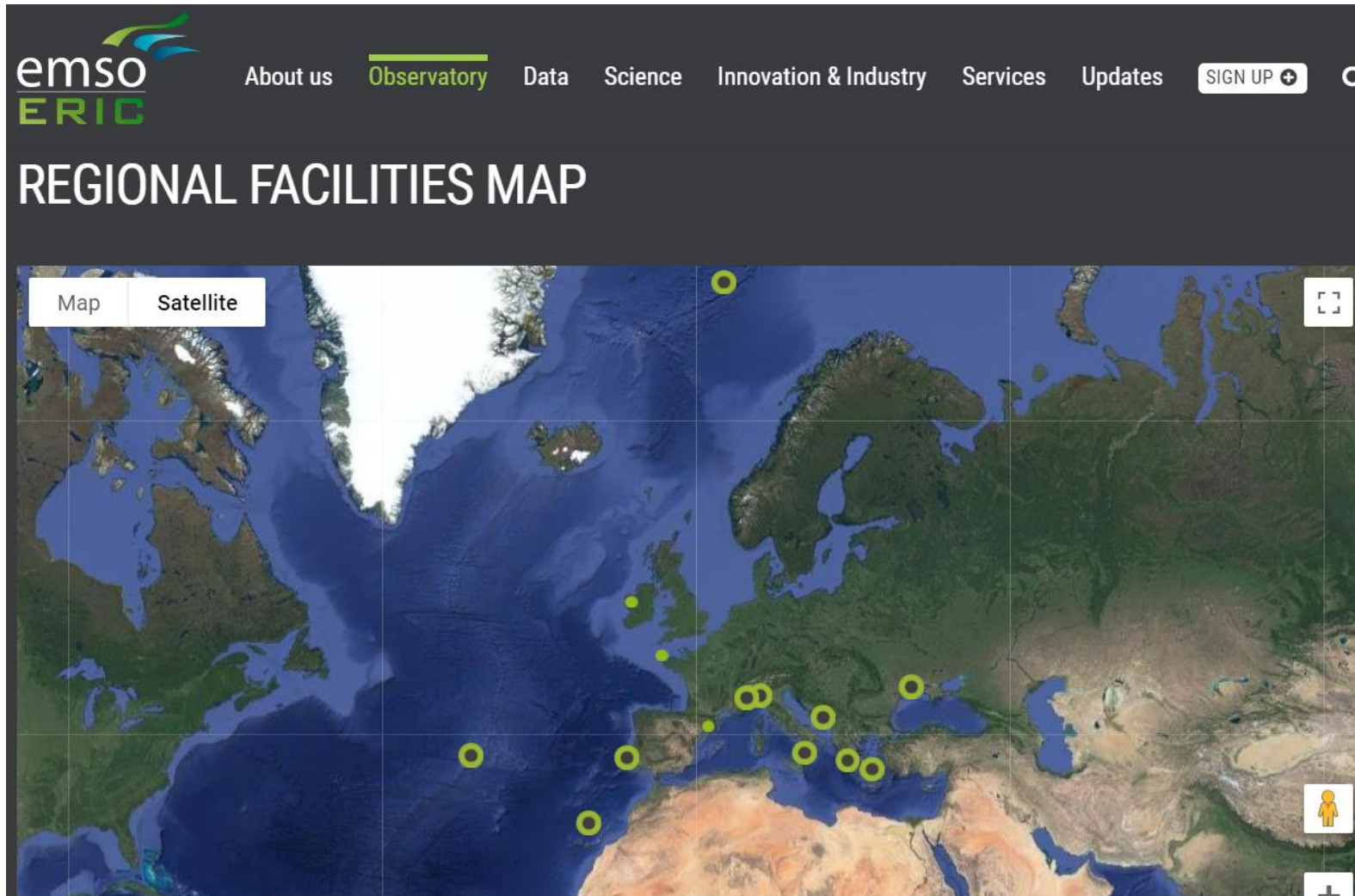


Status della condivisione dei dati in EMSO ERIC

Laura Beranzoli

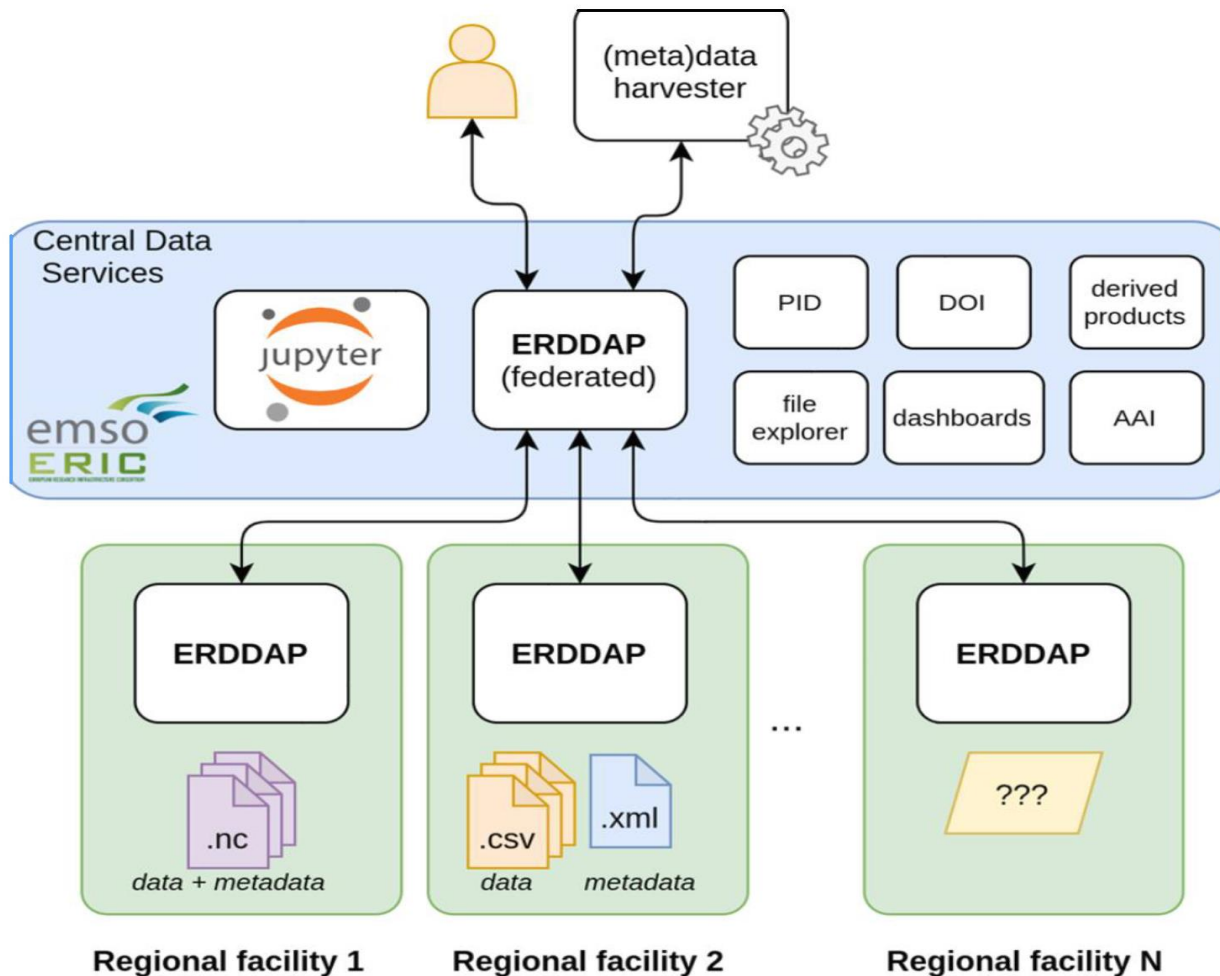
Sessione II - Condivisione di dati nel contesto Europeo e internazionale
Giornata INGV sulla Condivisione dei dati
15 nov 2023

EMSO – Dimensione dell'infrastruttura



- **11 siti** che ospitano sistemi osservativi di
 - profondità (>300 m)
 - complessità
 - estensione variabile, sia cablati sia autonomi
- **3 siti test**
- Acquisizione di serie temporali di **misure relative a varie discipline**

EMSO data management architecture and distribution as a federated ERDDAP system



Data discovery and access based on ERDDAP

versatile to download data in standard data formats such as .csv, .nc (NetCDF), .json, .mat, .xls ,and others

ERDDAP supports OPeNDAP, to remotely access parts of a dataset without downloading the entire file.

Data Management @ EMSO – organization

- **ERIC Central Management Office:**

- Data Officer
- Information Technology and Data Services Expert
- Data Systems Engineer
- IT Analyst
- Data Systems Technologist (INGV in kind)

(~3 FTE)

- **Data Management Service Group:**

- 1 member per regional facility (in kind from Research Organisations) (~1 FTE)

EMSO Data Policy

- Data Policy under construction according to Open Science
- EMSO utilizza la piattaforma open Zenodo per pubblicare i documenti ed alcuni data set particolarmente importanti; Zenodo prevede l'attribuzione di un DOI ad ogni documento e anche data-set e il suo uso contribuisce alla diffusione dei dati



ENVRI-FAIR, Marine RIs implementation plan

EOSC dashboard

Argo data selection

Argo EOSC discovery

Marine EO
broker VRE
cross domain

ENVRI-hub

etc...

Front office

Stakeholders services:
EOSC Blue cloud,
Copernicus, Argo...

F

search engine

A

OGC
WMS WPS
SensorThings

ERDDAP API

Sparql endpoint

data API
cloud services

I

R

metadata API

vocabulary server

argo ontology

Back office

Machine-to-machine
services Implementation
plan within ENVRI-FAIR
for **F A I R**
enhancements

Original dataset

Collection of
NetCDF files

Repositories

metadata
warehouse

data
warehouse

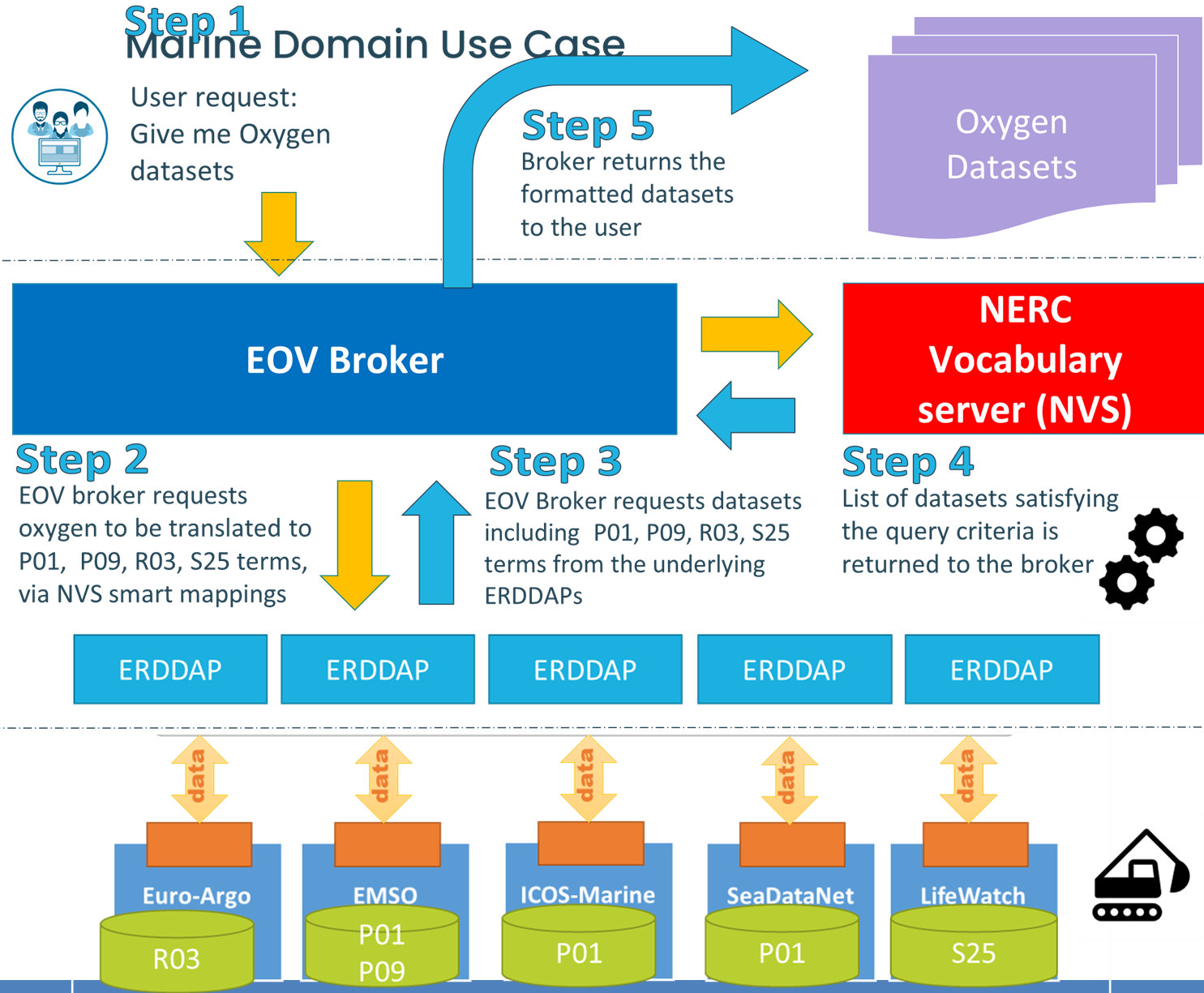
FTP, HTTPS
warehouse

Marine RI infrastructures

Argo, EMSO, Lifewatch,
ICOS, SeaDataNet data
repositories



The Marine USE CASE

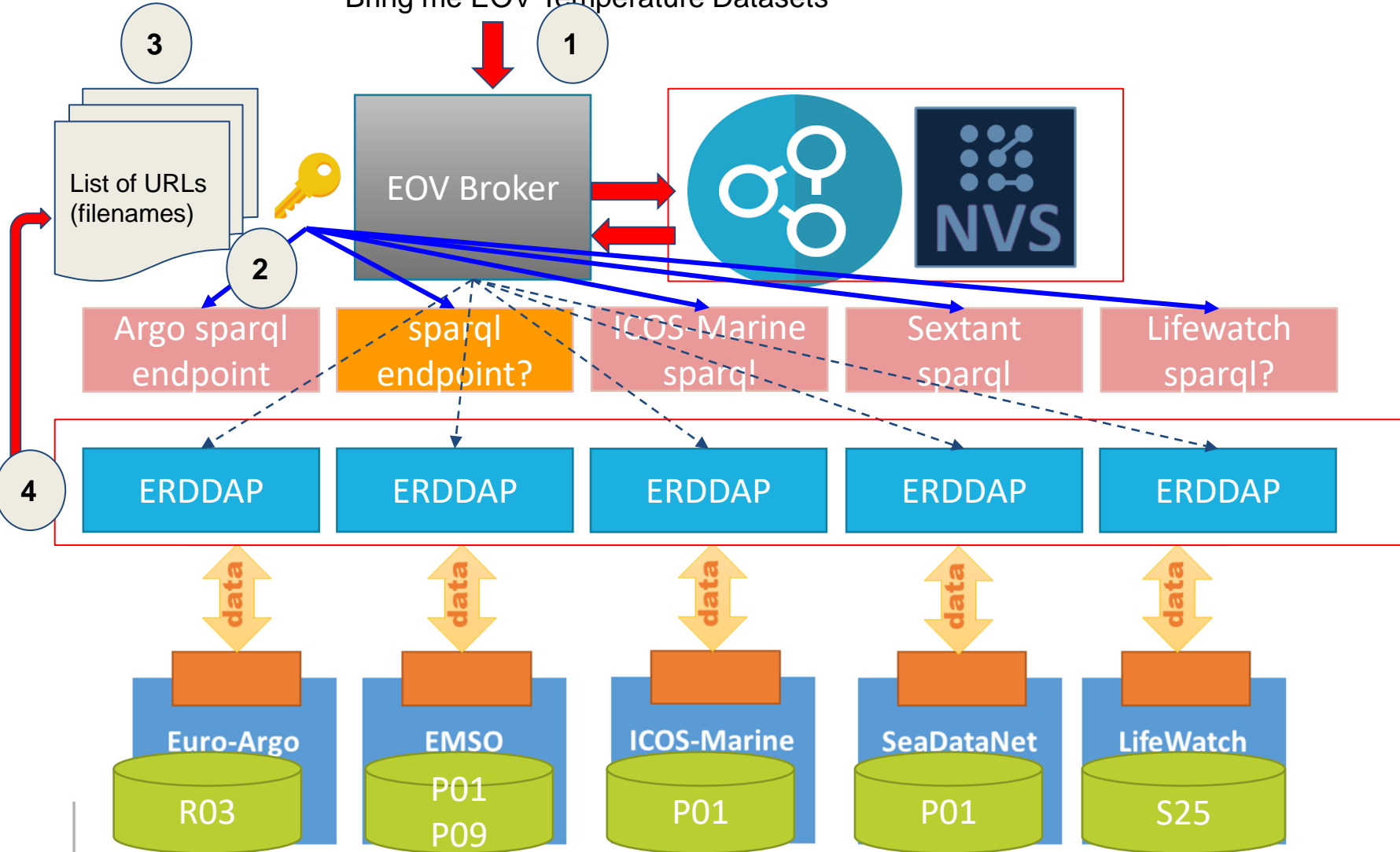




EOV demonstrator V2

Scenario II

Bring me EOV Temperature Datasets



Check what parameters the sparql endpoint include
e.g. **Datafiles** in SDN contain **P01** parameters while metadata contain **P02**

L'impegno attuale dell'INGV nel data management di EMSO

- processo di *harmonisation* di dati e metadati relativi all'oceanografia per la facility *Western Ionian Sea*; implementazione di *best practices* e convenzioni relative a dati e metadati:
 - OceanSITES, Climate and Forecast, SeaDataNet e BODC
 - NERC Vocabulary Service
- Tools for oceanographic data visualisation and analysis: definizione dei requisiti, test e validazione dei tools
- Adozione di standard per dati/metadati di rumore acustico (geofisica e bioacustica)
- Adozione degli standard utilizzati in EPOS per dati/metadati sismologici (nell'ambito di GeoInquire)

NEREIDE

NEw REsearch Infrastructure Data-center for EMSO

- Infrastruttura tecnologica avanzata sviluppata nel PON InSea da INGV presso la facility INFN di Portopalo
- comprende risorse per il calcolo e per la memorizzazione dati per facilitare l'archiviazione, l'elaborazione e la condivisione dei dati acquisiti dalle facilities di EMSO
- si basa su una piattaforma di *cloud computing*, per lo sviluppo e la pubblicazione di servizi digitali (es., per la *Data Science*, archiviazione e gestione)

NEREIDE può offrire servizi alla comunità di EMSO e non solo, se potenziata con adeguate risorse umane



emso

ERIC
EUROPEAN RESEARCH INFRASTRUCTURE CONSORTIUM

Observing the ocean to save the earth

