MetroSea Milazzo 2022

IEEE Advancing Technology for Humanity

Blue-Cloud: Exploring and demonstrating the potential of Open Science for ocean sustainability

Dick Schaap, MARIS, The Nederlands

Massimiliano Assante PhD, CNR, Italy

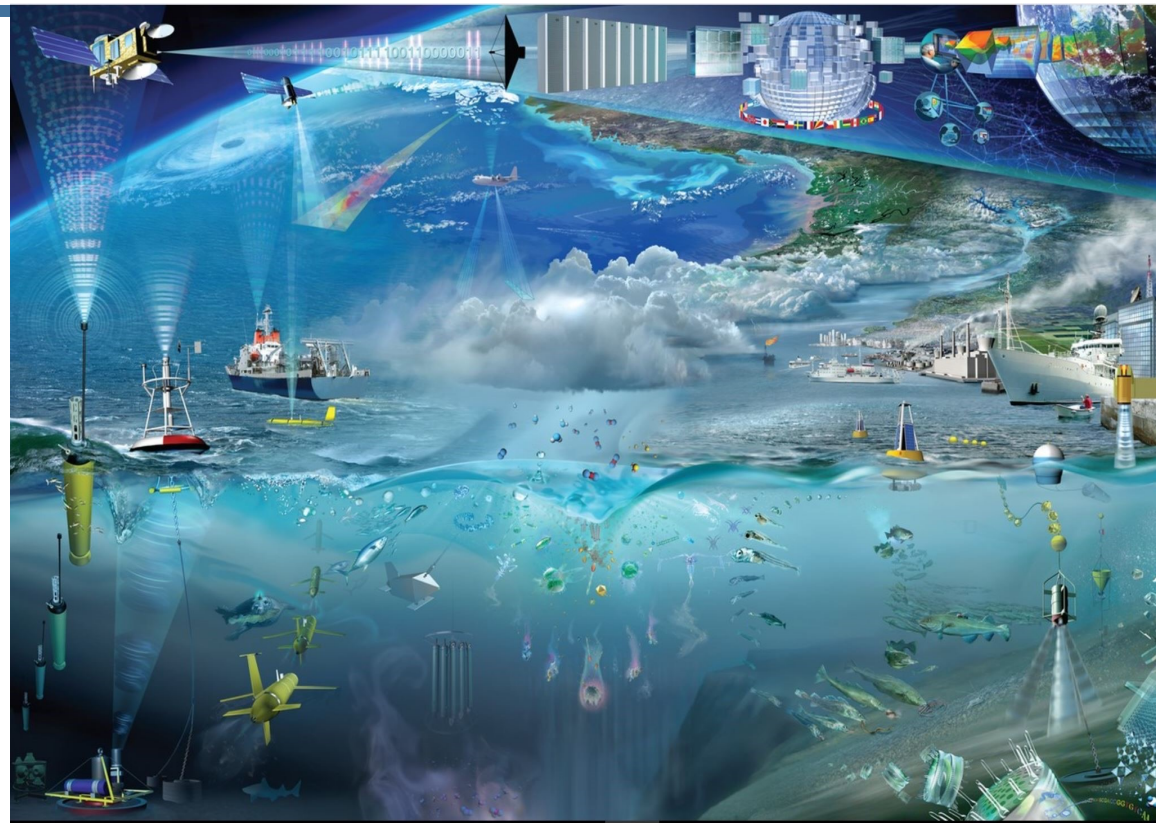
Pasquale Pagano, CNR, Italy

Leonardo Candela, CNR, Italy

Blue-Cloud

Piloting innovative services for Marine Research & the Blue Economy

Acquisition of marine and ocean data

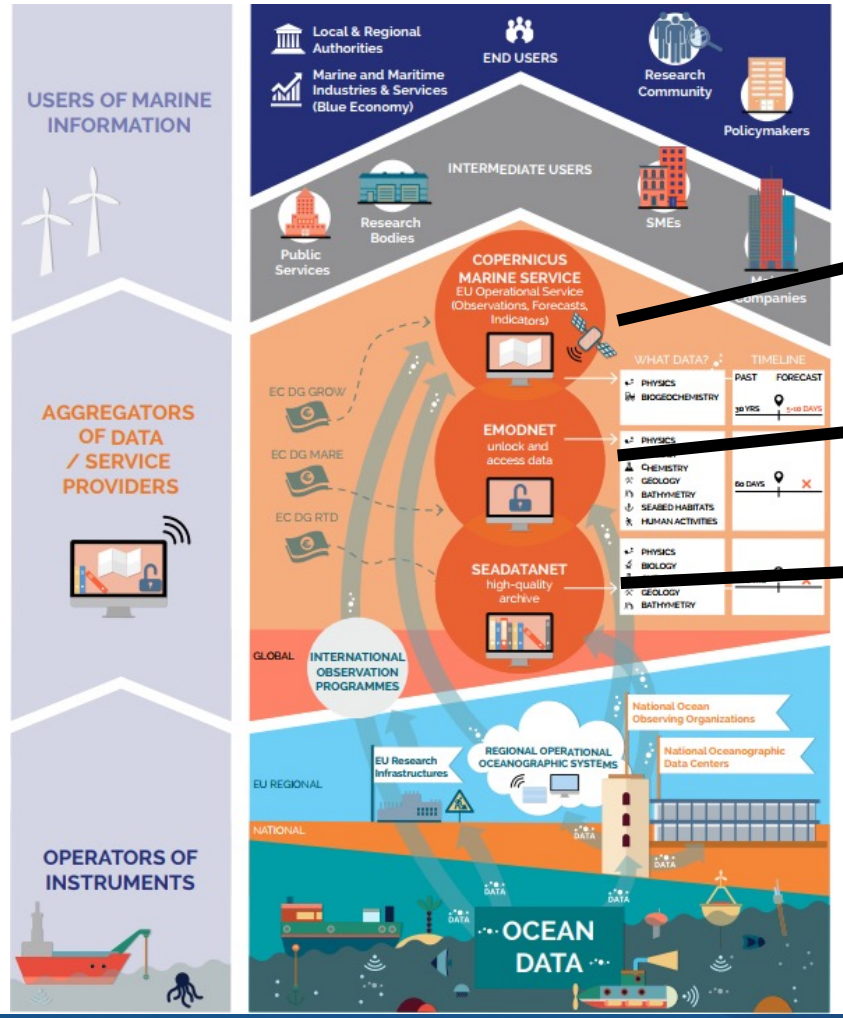


In Europe we spent circa 1.4 Billion Euro a year in marine data acquisition

(According to EMODnet study)

- Scientific Research to gain knowledge and insight
- Modelling (including hindcast, nowcast, forecast)
- Economic activities: shipping, offshore industry, dredging industry, fisheries, tourism, engineering ..
- Environmental Management: monitoring and assessment (water quality, climate status, stock assessment)
- Marine Conventions and Directives, in Europe: Water Framework Directive (WFD), Marine Strategy (MSFD), Marine Spatial Planning (MSP), Coastal Zone Management
- EU Strategies, such as Green Deal, Blue Environment, Blue Economy

European landscape marine data management



EuroGOOS European Global Ocean Observing System

EurOBIS European Ocean Biogeographic Information System

Copernicus Marine Service

ICOS Ocean Thematic Centre

emso ERIC

EMODnet European Marine Observation and Data Network

elixir

ICES CIEM

PANGAEA.

SeaDataNet

Euro Argo

ENA European Nucleotide Archive

EGDI

Data aggregators and providers of data products and services

EU H2020 Blue-Cloud project

Marine thematic contribution to European Open Science Cloud (EOSC)

October 2019 – March 2023; 20 partners

Blue-Cloud Mission

"to promote the sharing of *data, processes and research findings* in the marine domain by delivering a collaborative web-based environment that enables *Open Science;*

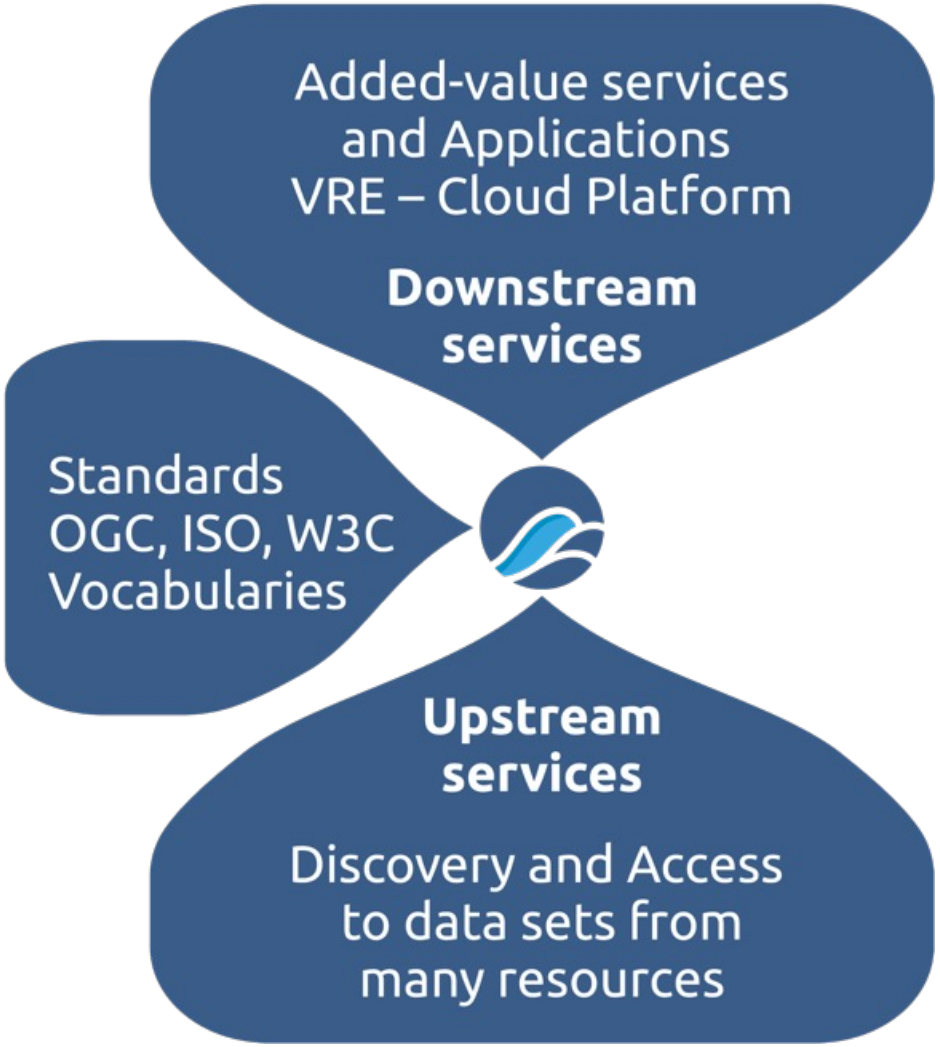
underpinned by simplified access to an unprecedented wealth of marine data resources and interoperable added-value services and products"

Blue-Cloud overarching concept

- Developing and deploying Virtual Research Environment (VRE) with an array of services for configuring and running Virtual Labs for specific analytical workflows, use cases, and demonstrators

- Applying common standards and interoperability solutions for providing harmonized metadata and data

- Developing and deploying harmonized discovery and access to established European marine data management and processing infrastructures



Blue-Cloud

Blue-Cloud federation of major infrastructures

EurOBIS
European Ocean Biogeographic Information System

WEkEO by COPERNICUS

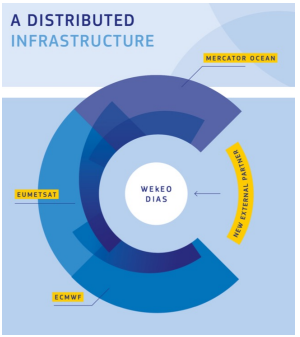
SeaDataNet

D4SCIENCE

EMODnet
European Marine Observation and Data Network

elixir

Euro Argo



EUDAT

ENA
European Nucleotide Archive

EURO-BIOIMAGING

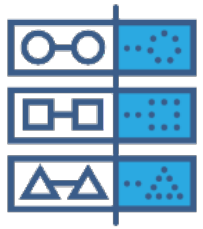
EcoTaxa ßETA

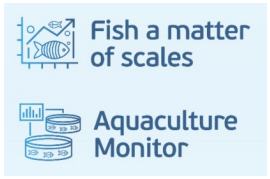
ICOS Ocean Thematic Centre

Blue Data infrastructures

E-infrastructures

Key products and services





- **Blue-Cloud Data Discovery & Access service**, federating key European data management infrastructures, to facilitate users in finding and retrieving multi-disciplinary datasets from multiple repositories
- **Blue-Cloud Virtual Research Environment infrastructure** to provide a range of services and to facilitate orchestration of computing and analytical services for constructing, hosting and operating Virtual Labs for specific applications
- **Blue-Cloud Virtual Labs**, configured with specific analytical workflows to serve as **Demonstrators,** which can be adopted and adapted for other inputs and analyses

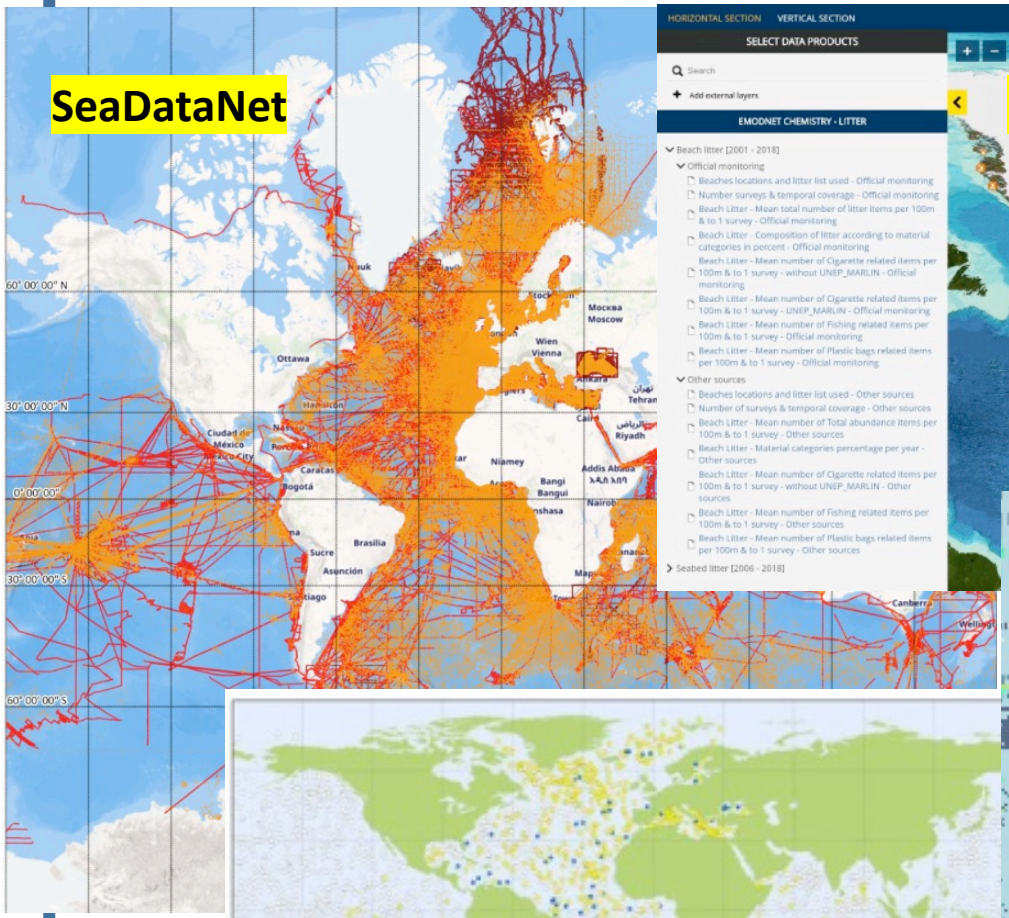
Blue Cloud Discovery and Access service

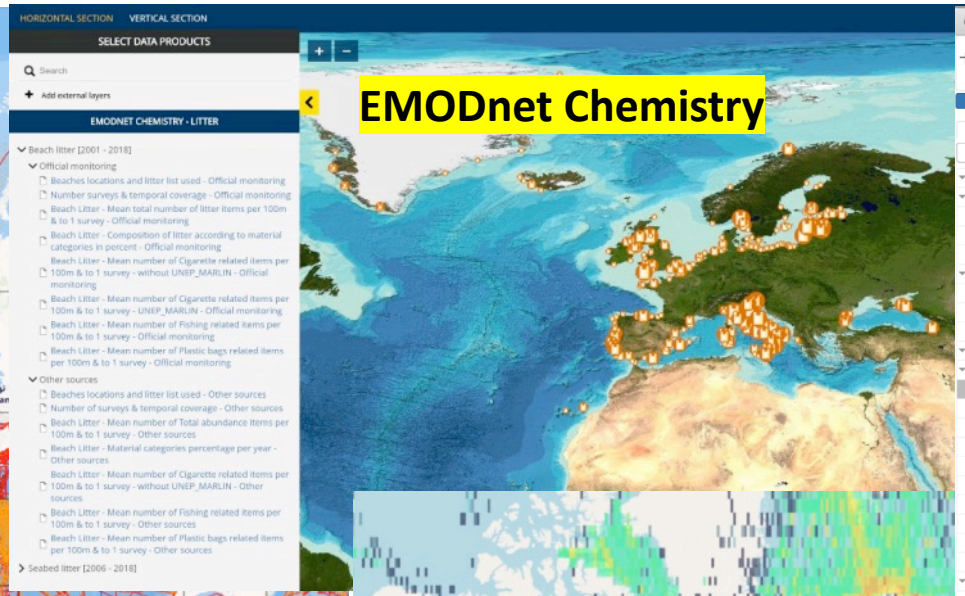
Built and operated by MARIS, CNR-IIA, and CINECA (EUDAT)

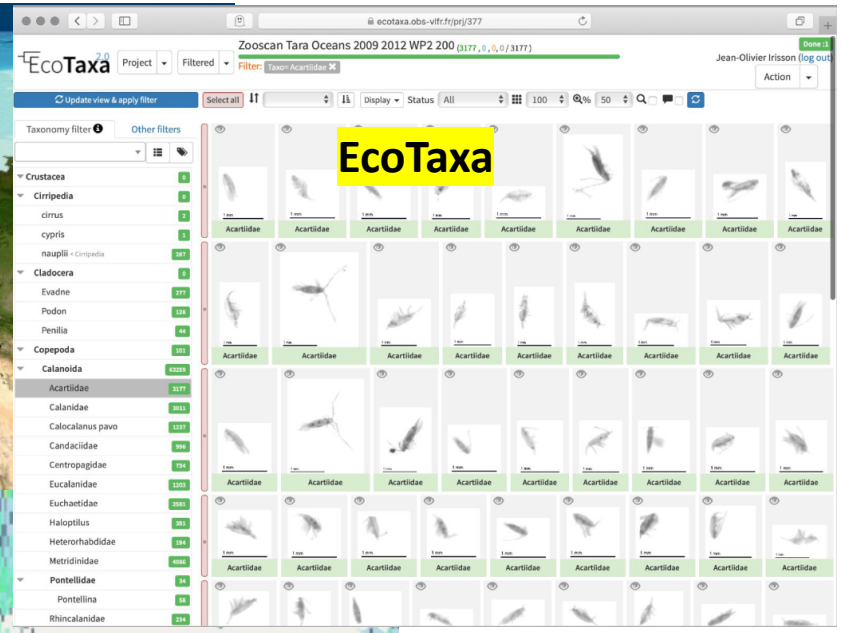
Blue Cloud Discovery and Access service

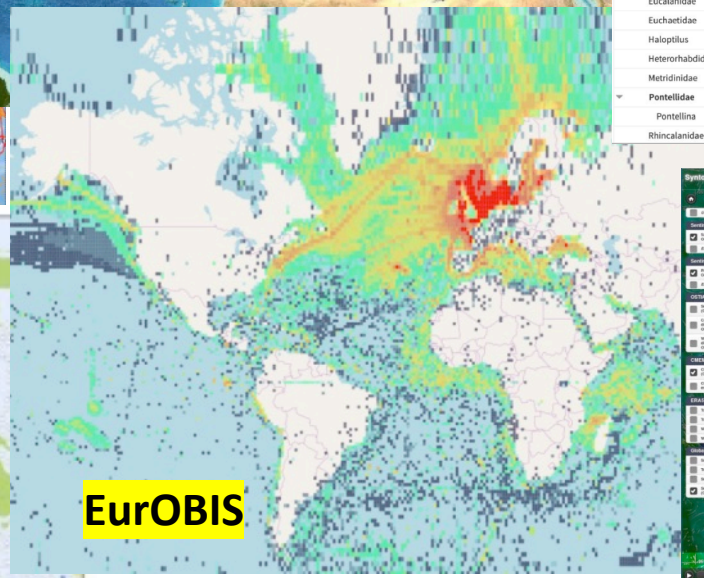
- **Facilitates users**:
  - Federated search for discovering interesting data sets (currently more than 10 million) in a common way
  - Federated retrieval of identified data sets using a shopping basket mechanism
  - Download of data sets or push to Blue-Cloud VRE
- **Facilitates managers of Blue Data Infrastructures**:
  - Wider outreach to potential users
  - Stay informed about data requests and users for their repository
  - Periodic reporting of downloads from their repository
- Built and operated by MARIS, CNR-IIA, and CINECA (EUDAT)

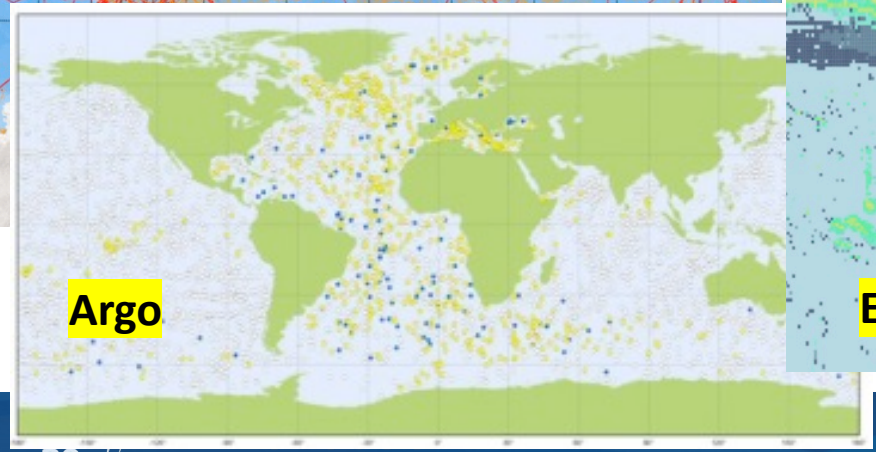
Illustrations of data coverage

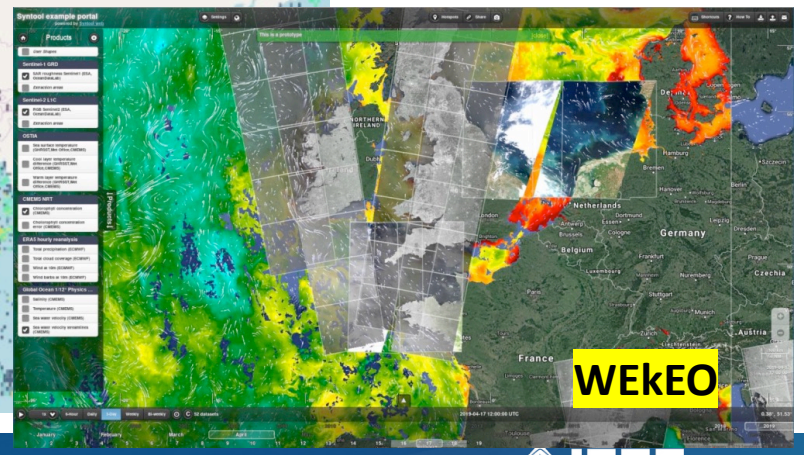












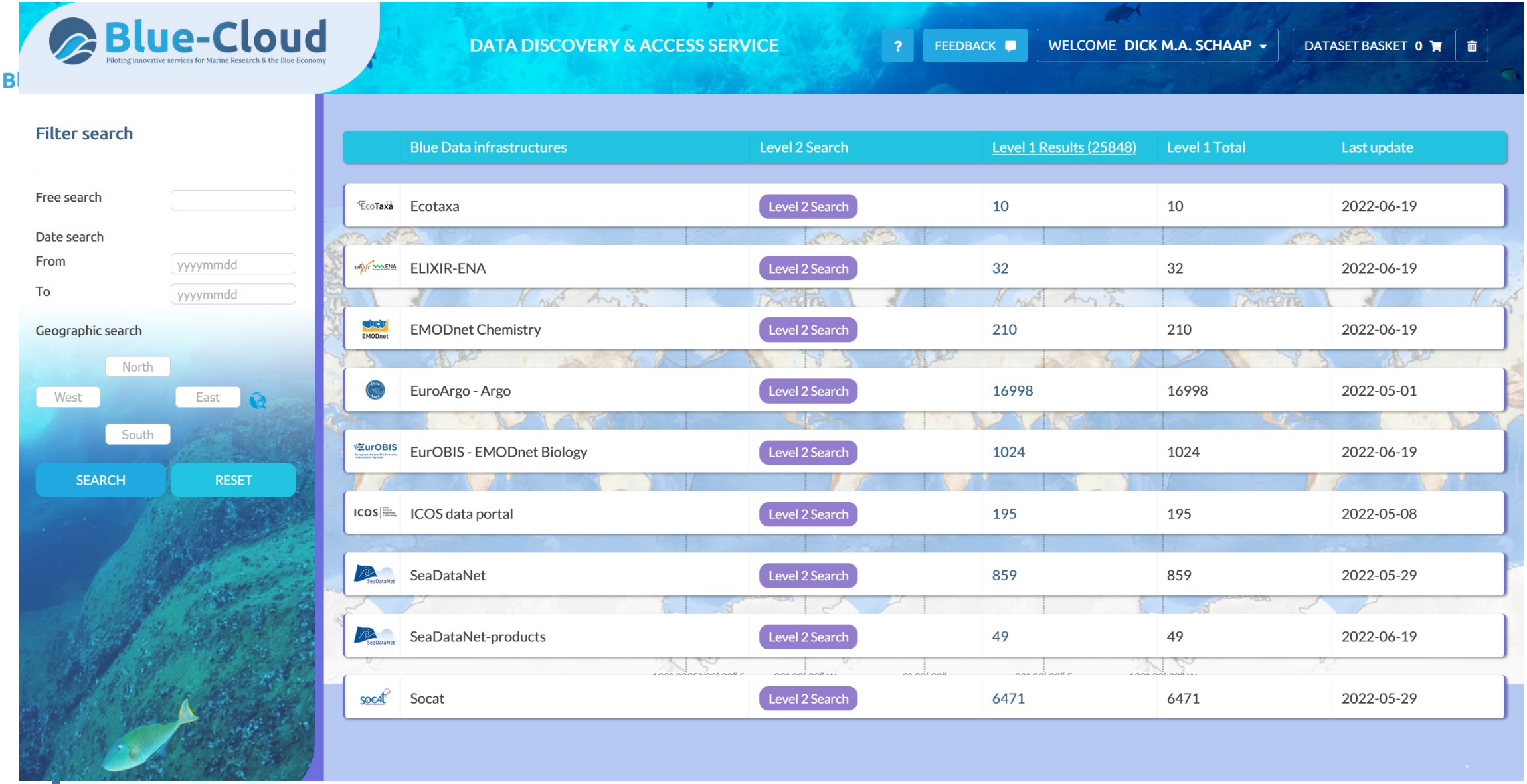
Blue-Cloud

Approach

- **Federated discovery and retrieval** of data sets and data products from the Blue Data Infrastructures

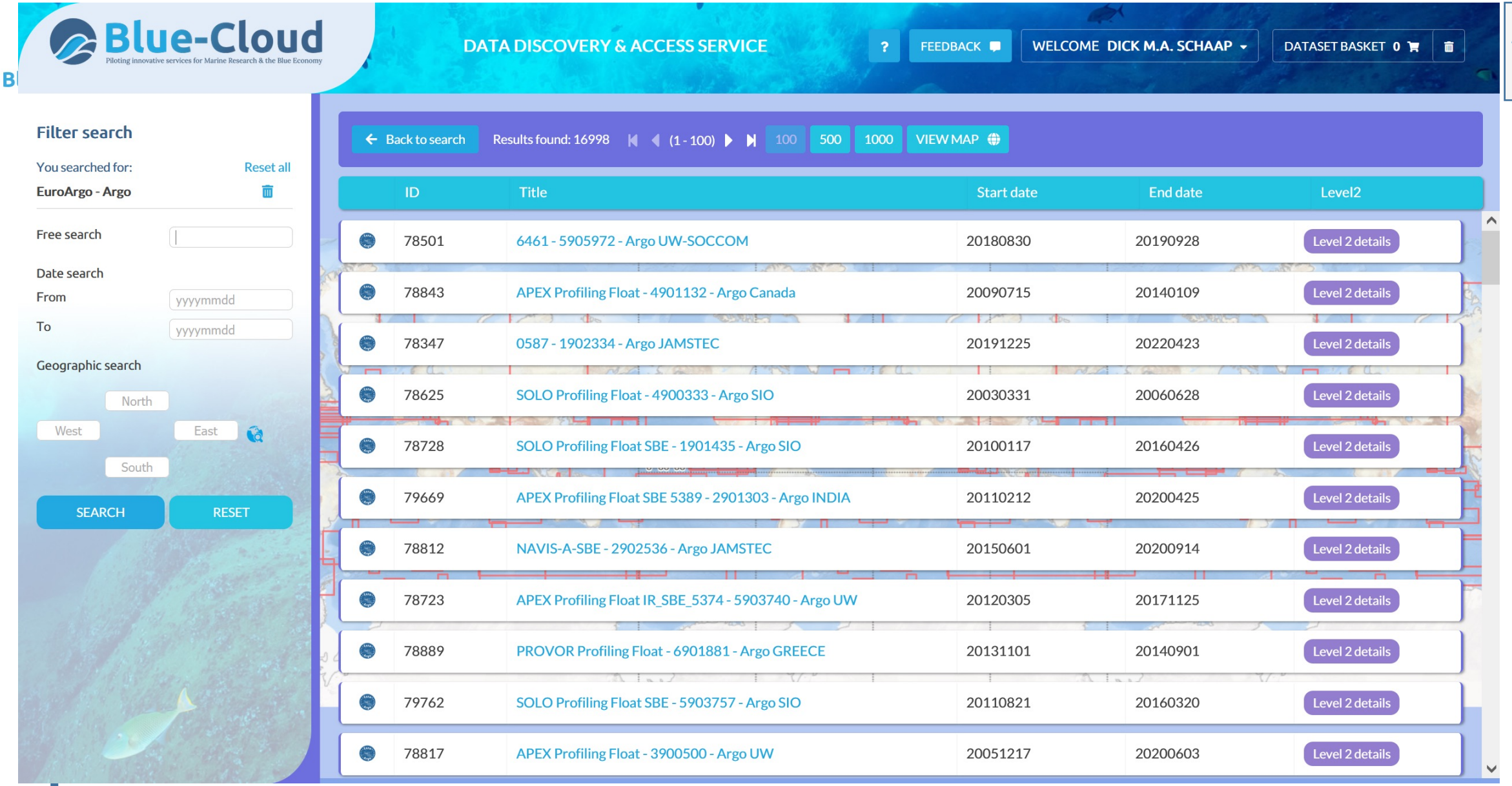
- **Concept of two-step search approach**:
  - First step: identifying interesting data collections and products with few criteria
  - Second step: drilling down with more criteria to select specific data at granule level, where possible, otherwise at collection/products level

- Metadata and Data Brokerage services interacting **Machine-to-Machine** with web services and APIs as provided and operated by the Blue Data Infrastructures



HOME
Level 1

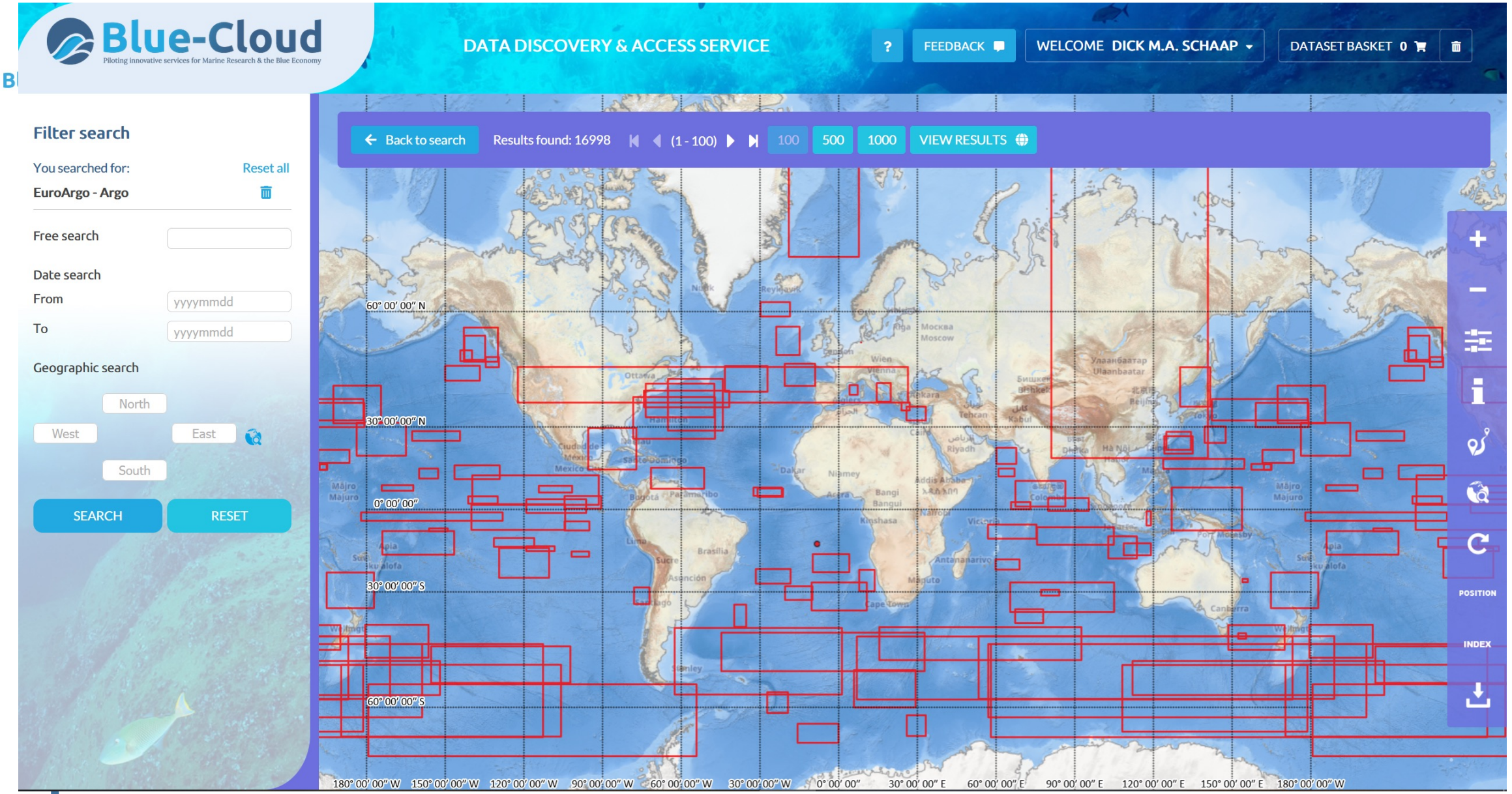
https://data.blue-cloud.org



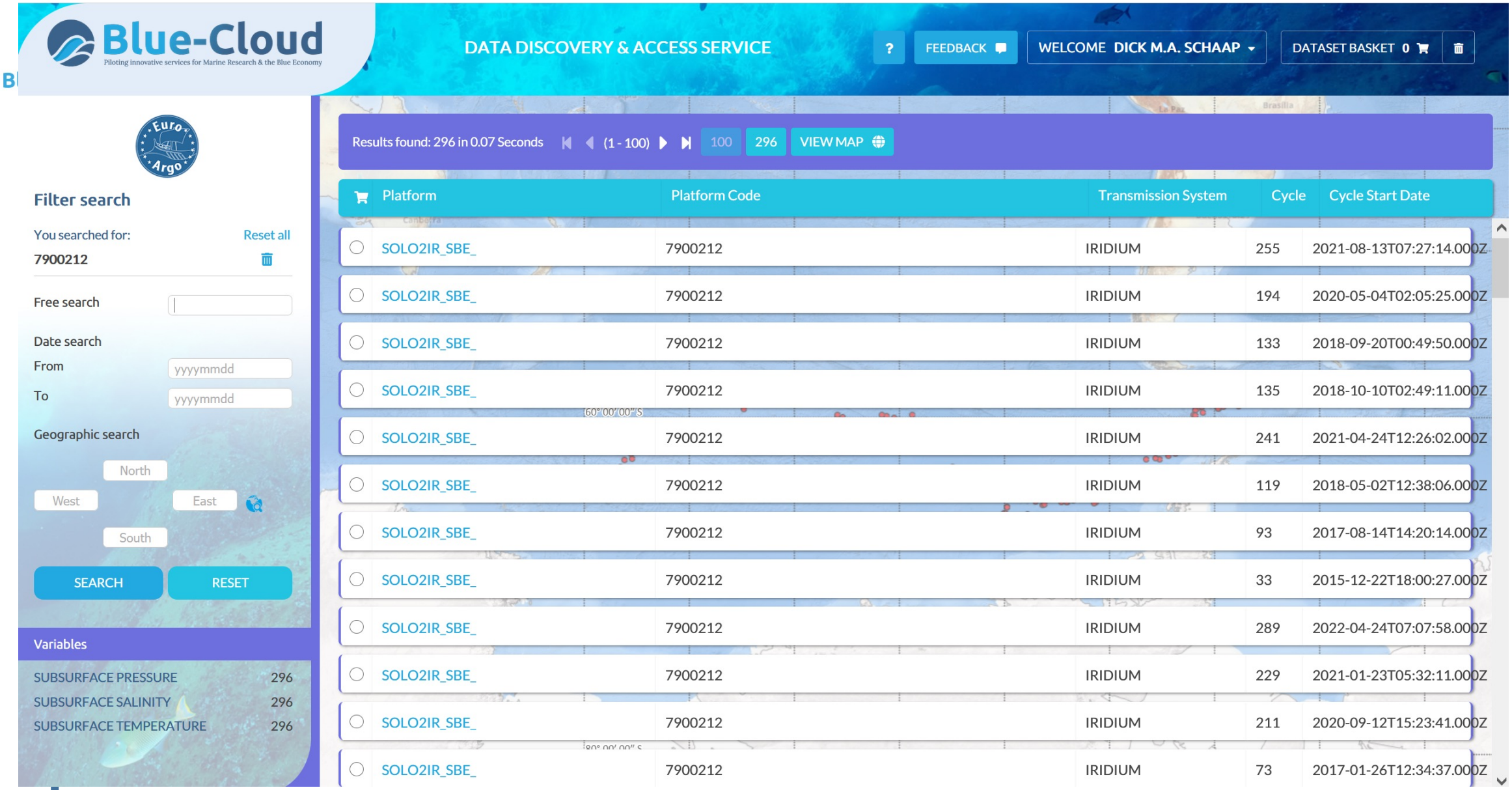
ARGO Level 1

https://data.blue-cloud.org

ARGO Level 1

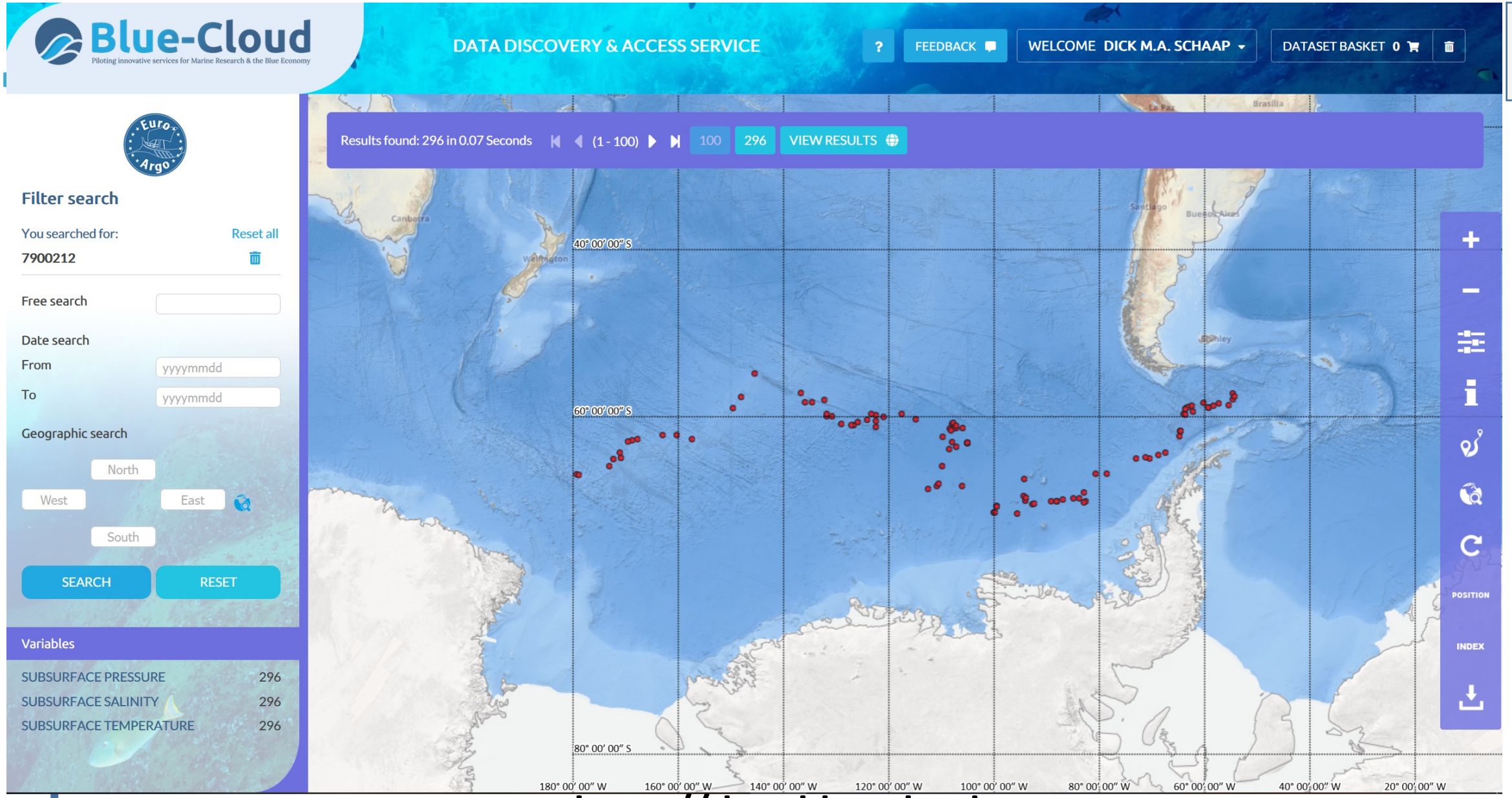


https://data.blue-cloud.org



ARGO Level 2

https://data.blue-cloud.org



ARGO Level 2

https://data.blue-cloud.org

Blue-Cloud

Level 2 Interfacing - Protocols

SeaDataNet	Dedicated API
SeaDataNet Products	OGC CSW service
EMODnet Chemistry	OGC CSW service
EuroArgo - Argo	Dedicated API
EurOBIS – EMODnet Biology	DCAT service
Ecotaxa	Dedicated API
ELIXIR - ENA	Dedicated API
ICOS Marine	SPARQL service
SOCAT	ERDDAP service

MetroSea Milazzo 2022

IEEE Advancing Technology for Humanity

Virtual Research Environment (VRE)

Built and operated by CNR-ISTI (D4SCIENCE) D4SCIENCE

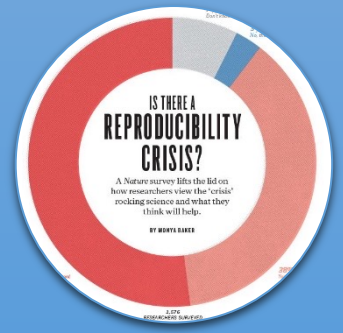
Blue-Cloud

Blue-Cloud VRE

Collaboration

Sharing

Reuse



Reproducibility

Blue Cloud VRE promotes Open Science

Blue-Cloud

Blue-Cloud VRE

System of System

Blue-Cloud VRE is built with dedicated services leveraging on existing e-infrastructures and marine infrastructures, EOSC resources and services

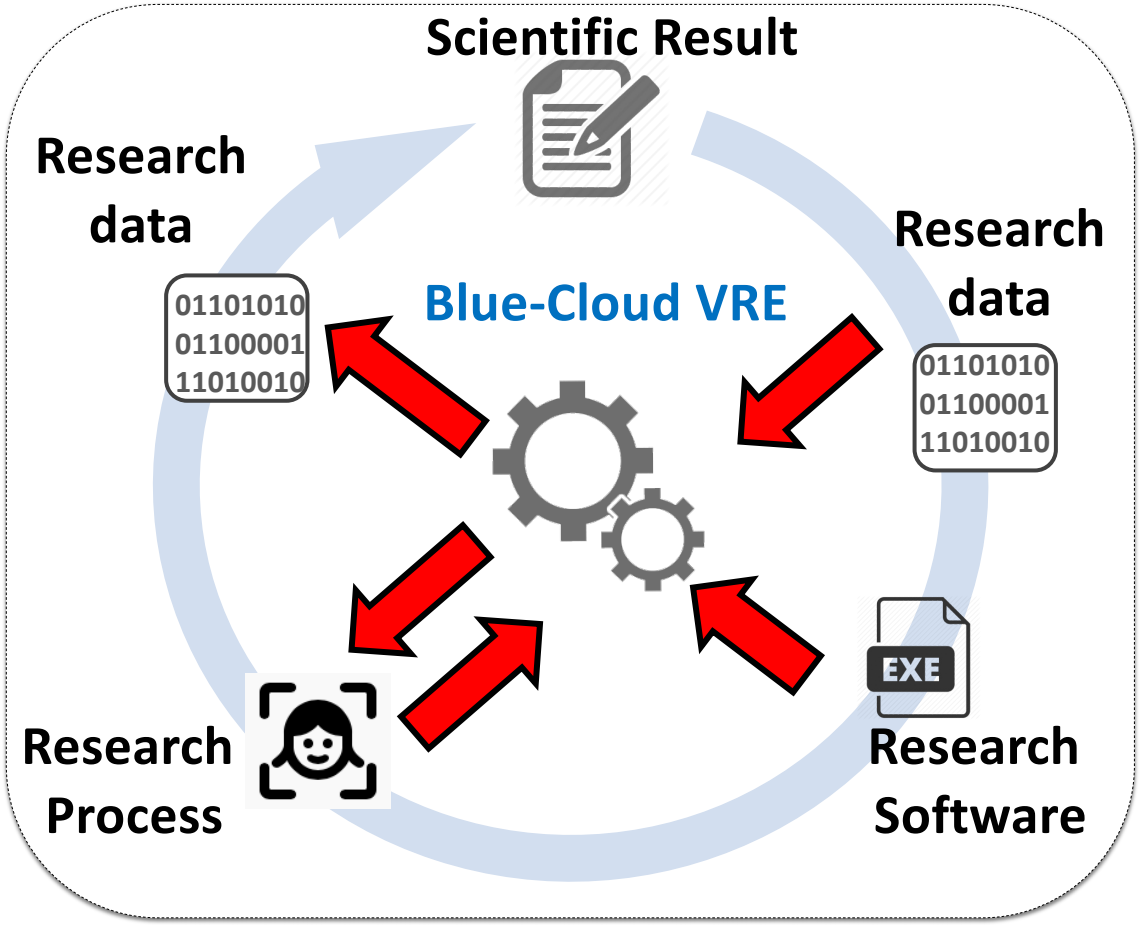
Extensible

Blue-Cloud VRE can integrate services and resources resulting from existing marine initiatives

Open

Blue-Cloud VRE promotes open science and practices

A SoS to support and promote Open Science



Enable

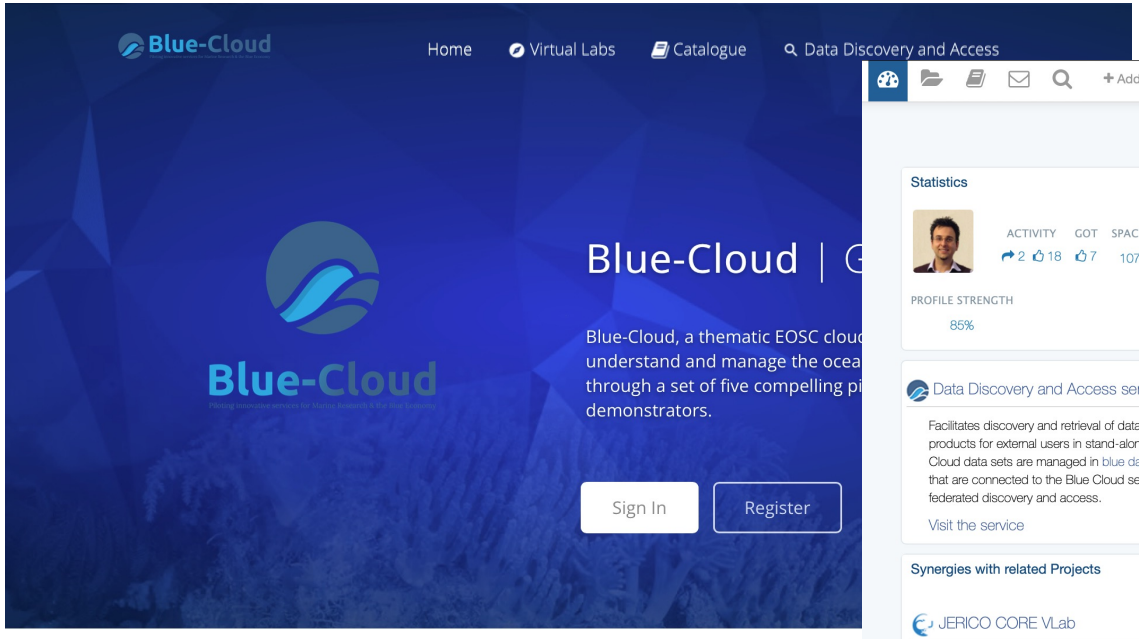
- Repeat, Reproduce, Reuse, Evaluate
- Active collaboration
- Effective sharing
- Provenance and attribution

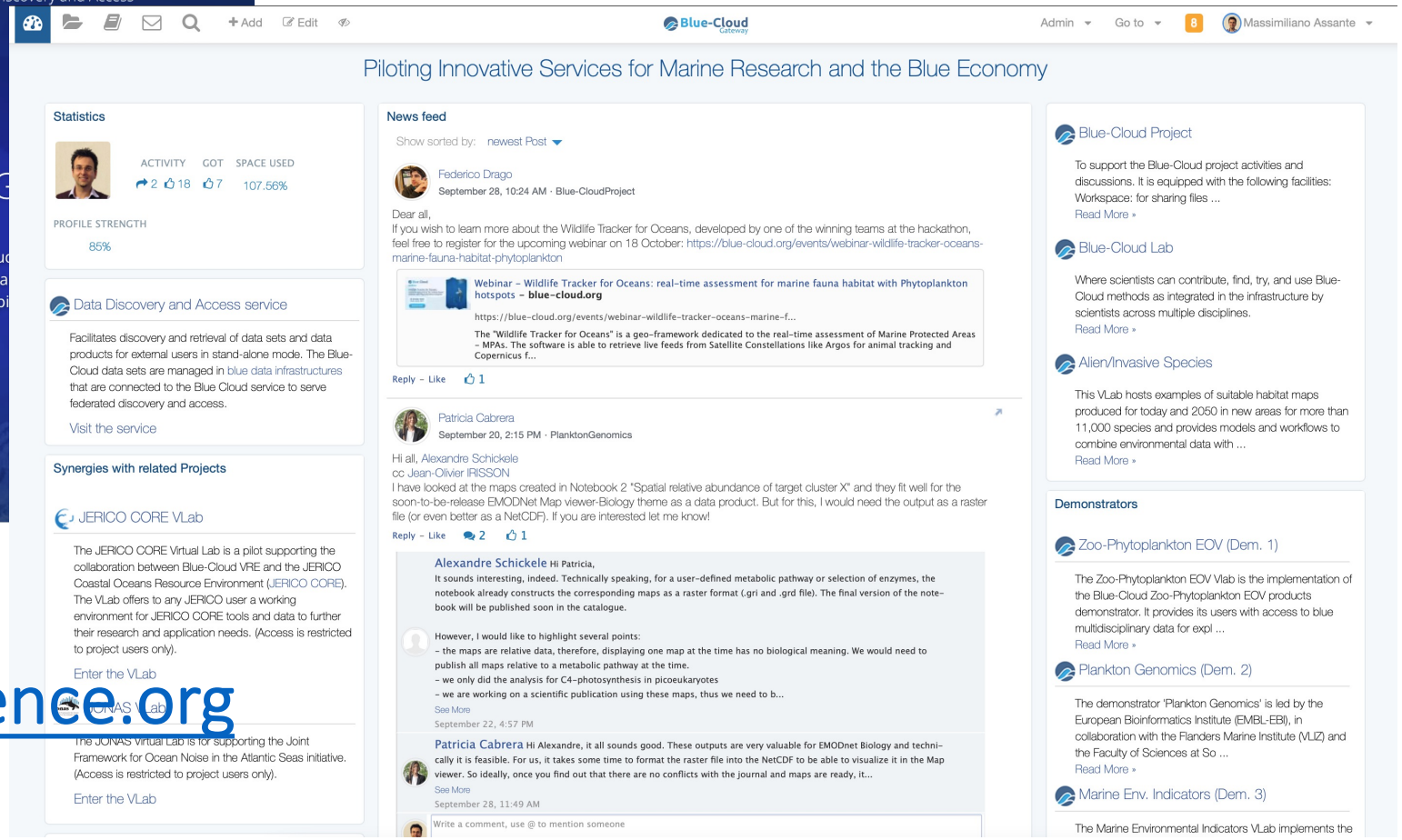
Adopt

- As-a-service approach
- Standards
- Economy-of-scale to reduce operational costs

Blue-Cloud

VRE at-a-glance





https://blue-cloud.d4science.org

Blue-Cloud

One place for store, share, and preserve datasets

- Common **workspace** and **dataspace** to easily exploit technologies and services not designed to work together

File System

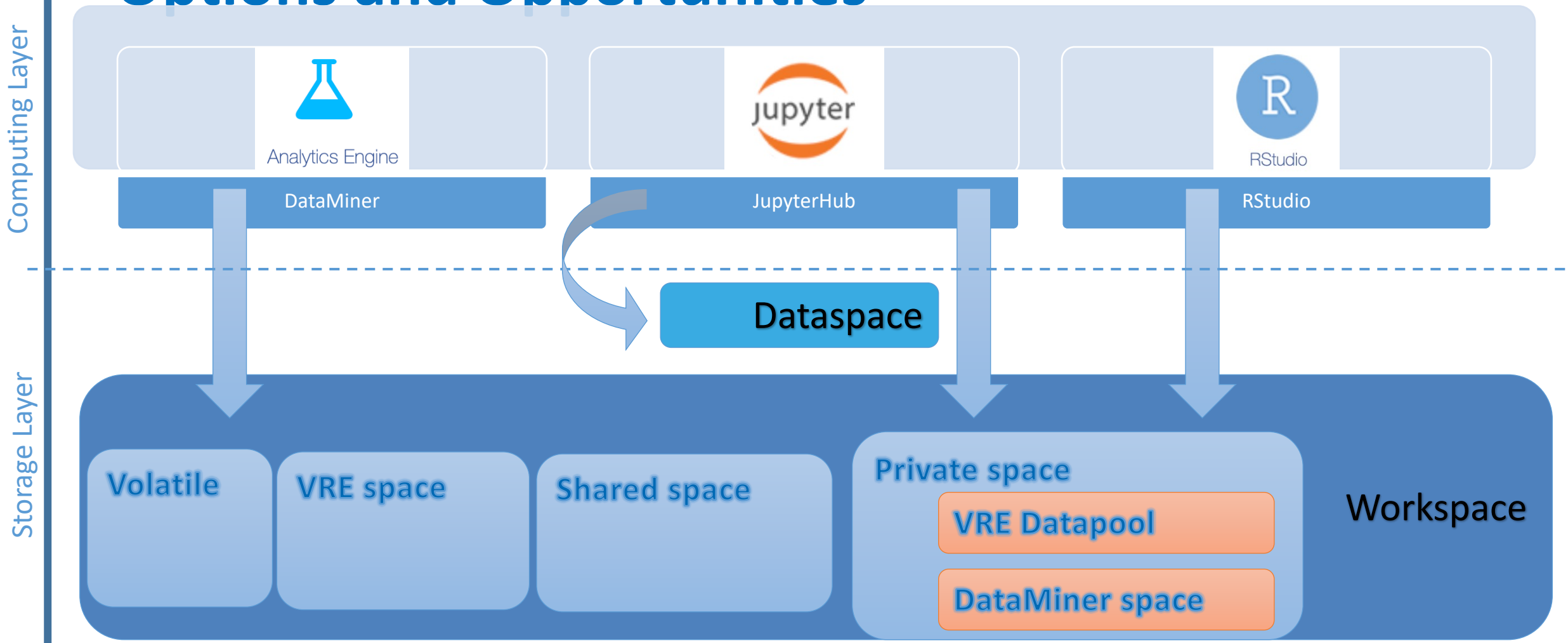
- *Fault-tolerant*
- Fully-replicated
- Encrypted
- Compressed

Versioning

Provenance

Controlled Sharing

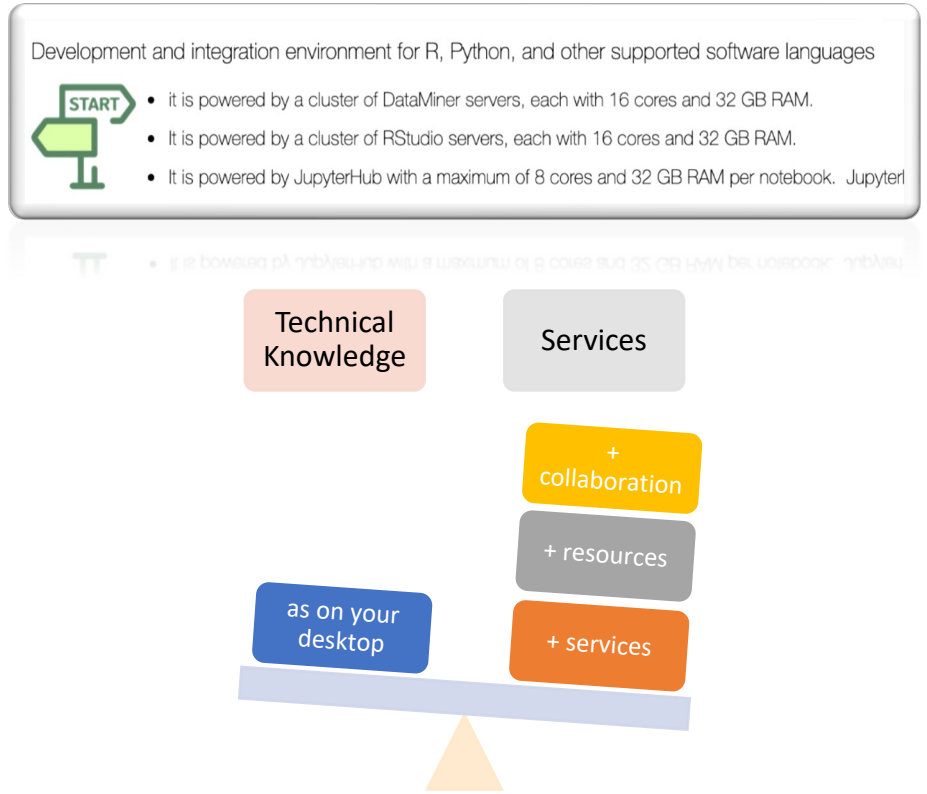
Data Analytics Options and Opportunities



Blue-Cloud

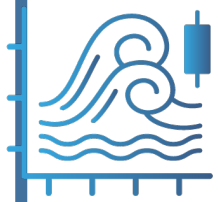
One place for execute analysis and processes

- interactive notebooks via JupyterHub and community-specific applications delivered as a Docker container extend the Analytics framework



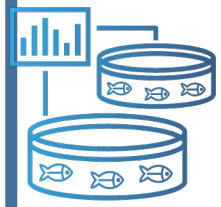
Blue-Cloud

5 Virtual Labs at the VRE

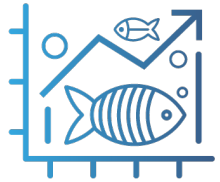


Marine Environmental Indicators

Zoo & Phytoplankton EOV products



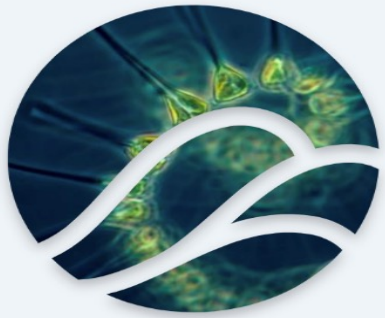
Aquaculture Monitor



Fish a matter of scales

Plankton Genomics

Vlabs for different domains



Biodiversity

Zoo and Phytoplankton EOV products

Genomics

Plankton Genomics

Environment

Marine Environmental Indicators

Fisheries

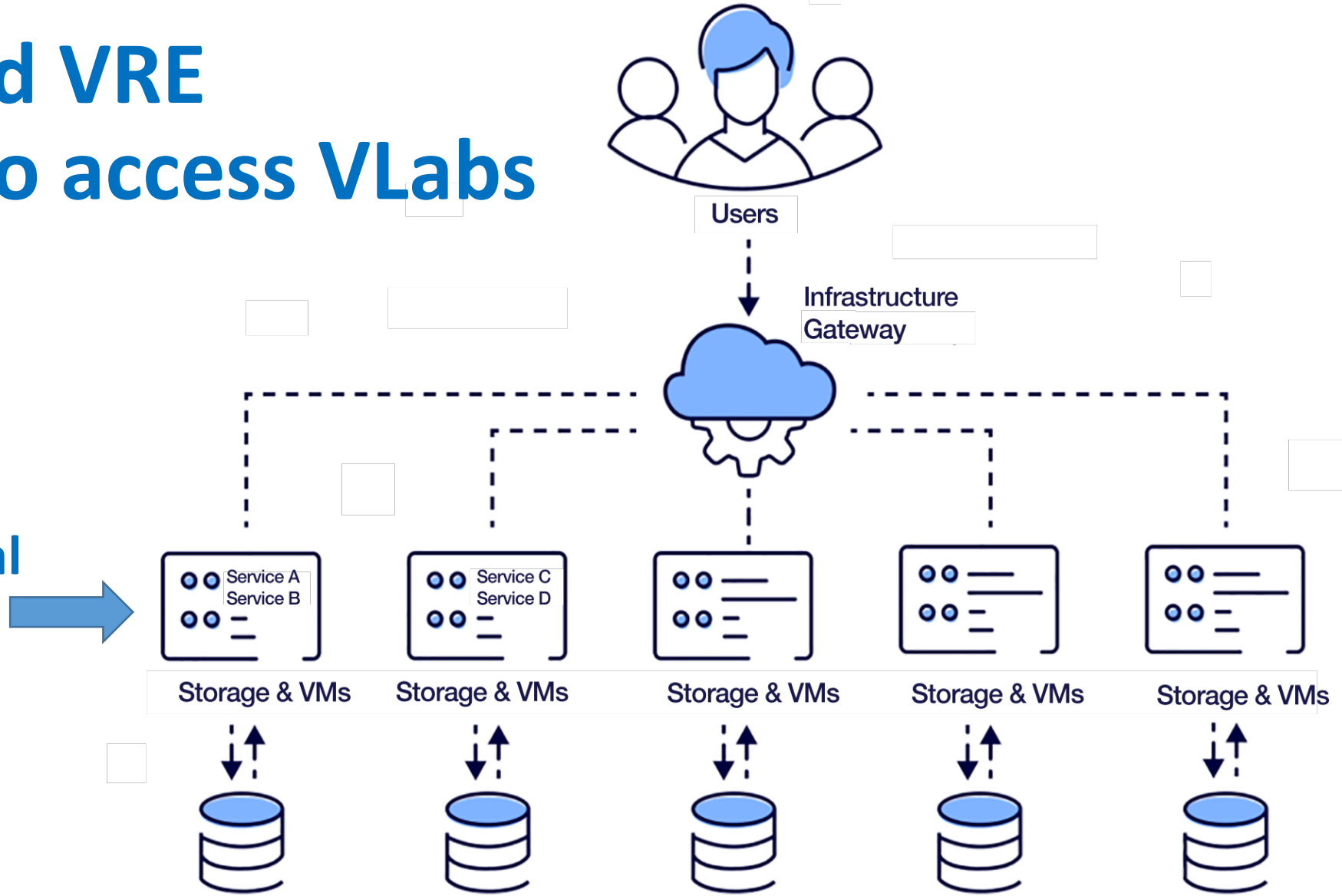
Global Record of Stocks and Fisheries

Aquaculture

Aquaculture Monitor

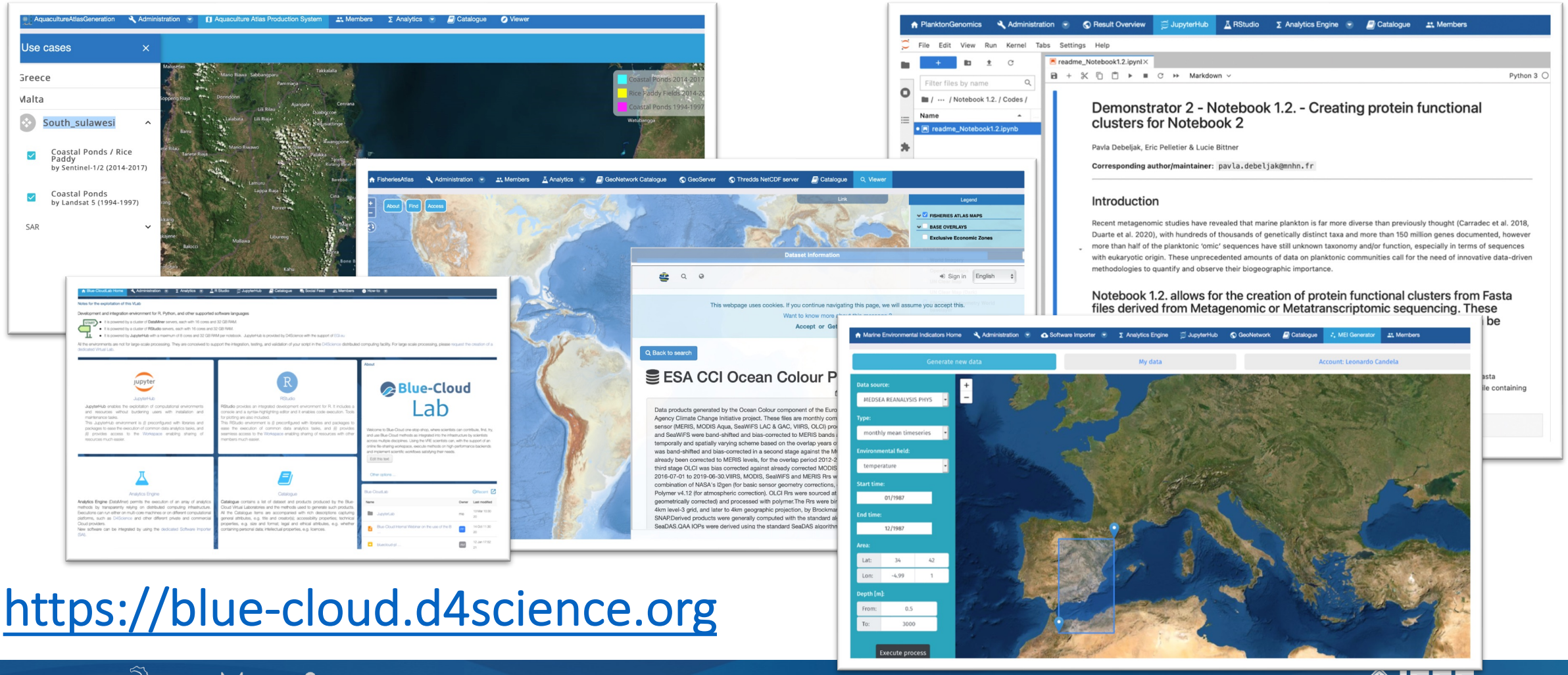
Blue-Cloud VRE gateway to access VLabs

Virtual Labs



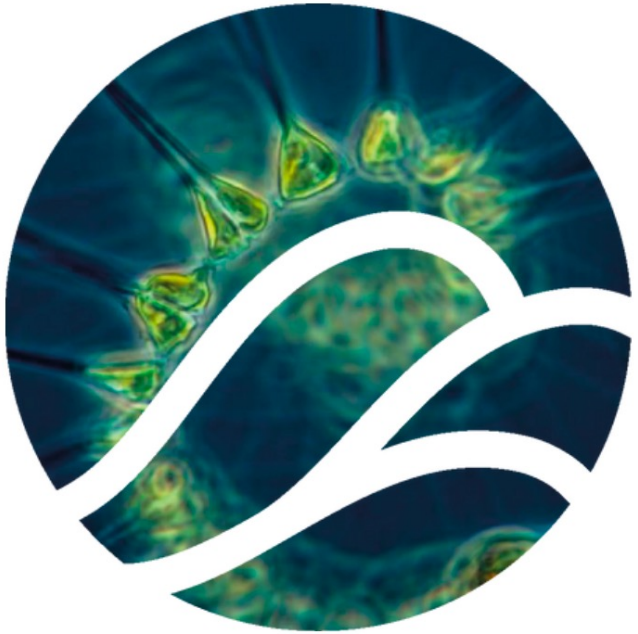
Blue-Cloud

VLabs at-a-glance



https://blue-cloud.d4science.org

Blue-Cloud



Zoo- and Phytoplankton Essential Ocean Variable products

Test the VLab now!

Partners:

VLIZ LOV LIÈGE université

Data sources through Blue-Cloud:
EurOBIS, EMODnet Biology, LifeWatch, GEBCO, SeaDataNet, World Ocean Atlas, NOAA, Copernicus Marine Service, Argo GDAC, GlobColour

Main target users:
Plankton researchers, ocean modellers, data product developers and Blue Data infrastructures, for their data products catalogues and as use cases.

Services introduction:
The Vlab offers three independent services that consist of the combination of different data types (biological, physical and environmental data) to then apply models that generate an output. These are offered in a working space where data and scripts are accessible and reusable.

UN SDGs addressed

13 CLIMATE ACTION

14 LIFE BELOW WATER

SERVICES

Zooplankton Essential Ocean Variable

Zooplankton EOV generates zooplankton gridded maps of six zooplankton species in the North East Atlantic.The workflow uses the DIVAnd software tool (Data Interpolating Variational Analysis in n dimensions) that allows to interpolate

Phytoplankton Essential Ocean Variable

Phytoplankton EOV generates global open ocean 3D gridded products of (1) chlorophyll a concentration (Chla), which is a proxy of the total phytoplankton biomass, and (2) Phytoplankton Functional Types (PFT), as a proxy for phytoplankton diversity, based on

Modelling phyto & zooplankton interactions

Modelling phyto and zooplankton interactions enables users to calculate the relative contribution that limits the growth of phytoplankton by the drivers: nutrients, phosphates, silicates, light and zooplankton grazing.

Milazzo 2022

IEEE Advancing Technology for Humanity

Successor project: Blue-Cloud 2026

Blue-Cloud 2026

A federated European FAIR and Open Research Ecosystem for oceans, seas, coastal and inland waters

MISSION: To develop further the European federation of marine and inland water data management infrastructures & high quality services

A1. DD&AS

A FAIR compliant Data Discovery & Access Services > access to 10+ million open data sets & products by 13 major BDIs

A2. VRE

An Open Science Virtual Research Environment (VRE) federating multiple e-infrastructures > supporting Analytical Big Data Workbenches & VLabs

A3. EOVs

3 EOV Workbenches for highly qualified data collections

3.000 DATA ANALYTICS SESSIONS PER MONTH - 5,000 HTC DATA ANALYTICS JOBS PER MONTH

A4. VLABs - FIVE DOMAIN-BASED VIRTUAL LABS

- Coastal Ocean observations along Europe
- Coastal currents from observations
- Carbon-Plankton Dynamics
- Marine Environmental Indicators
- Global Fisheries Atlas

A7. COMMUNITY

- **All EU countries** engaged
- **3k+** engaged Blue- Cloud community users
- **5k+** followers across all the platforms
- **10+** External Stakeholders

OUTREACH

- **1** Blue-Cloud Hackathon
- **1** Blue-Cloud TV
- **18** Newsletter issues
- **11** Webinars on Blue-Cloud VRE, DDAS & EOV Workbenches
- **3** Blue-Cloud Annual Impact Events
- **3** Ocean Literacy Webinars
- Videos & Interviews

A6. TRAINING ACADEMY & CATALOGUE

- **3** Online training course on Best Practices for FAIR data principles
- **3** Info session & course on the EOV Workbenches
- **2** online webinars dedicated to the BlueCloud VRE
- **2** dedicated to the DDAS and the innovations introduced
- A series of training sessions on how to use the VLabs

POLICY

- Scientific papers & articles
- Restoring healthy oceans, seas, coastal & inland waters in Europe
- Strategic Roadmap 2030 A5. ROADMAP
- Cross-domain expansion factsheets
- Sustainability Business model

DTO Task Force

MetroSea Milazzo 2022

IEEE Advancing Technology for Humanity

Thank you

Massimiliano Assante, CNR, Italy
massimiliano.Assante@cnr.it
Twitter: @maxassante