



Blue-Cloud

Blue-Cloud VRE

The platform for developing & operating VLabs

Pasquale Pagano

CNR – ISTI

Blue-Cloud Open Conference

Brussels, 8 December 2022





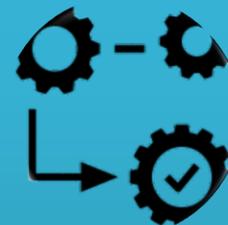
Blue-Cloud VRE



One place to
discover and
access



One place to
store, share, and
preserve



One place to
execute analysis
and processes





Blue-Cloud VRE

System of Systems

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

Extensible

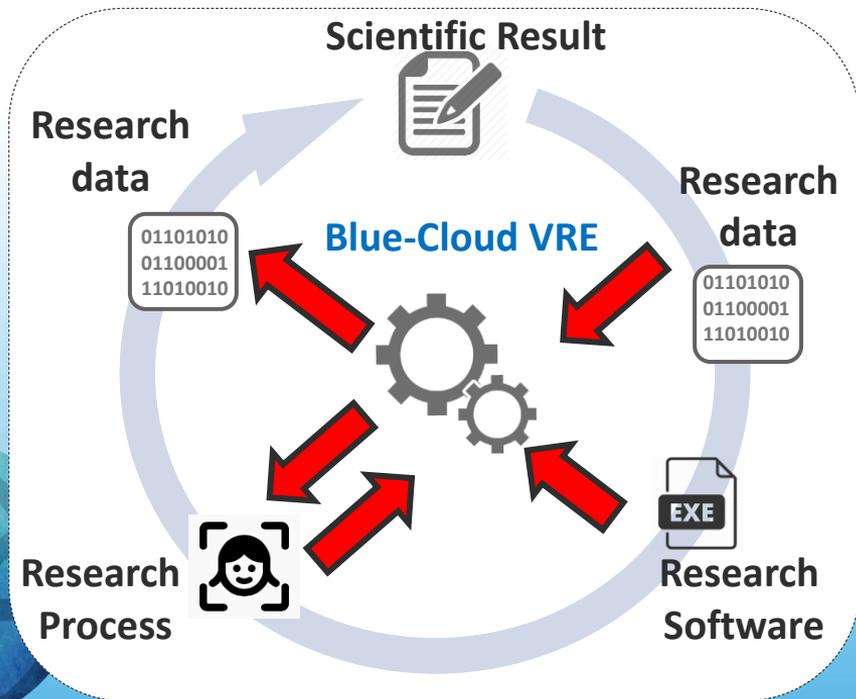
Blue-Cloud VRE integrates services and resources resulting from existing marine initiatives

Open

Blue-Cloud VRE promotes open science and practices

Blue-Cloud VRE

A SoS to support and promote Open Science



Enable

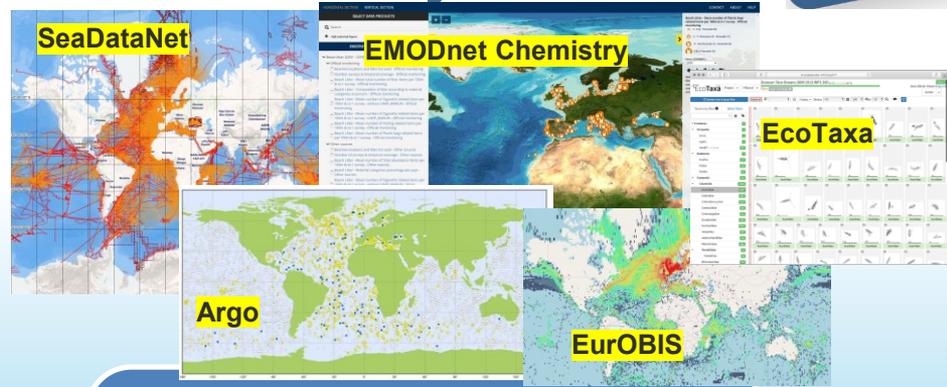
- Repeat, Reproduce, Reuse, Evaluate
- Active collaboration
- Effective sharing
- Provenance and attribution

Adopt

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

One place for discovery

Blue Data Infrastructure	Types of data sets	Logo and link
SeaDataNet CDI service	Marine physics, bathymetry, chemistry, geology, geophysics, and biology observation data sets	
EMODnet Chemistry data products	Marine chemistry data collections and interpolated map products	
EuroBIS - EMODnet Biology	Marine biogeographic data collections with taxonomy and distribution	
Euro-Argo and Argo GDAC	Ocean physics and marine biogeochemistry observation data from Argo floats	
ELIXIR - European Nucleotide Archive (ENA)	Nucleotide sequencing data and information on marine species	
EcoTaxa	Taxonomic annotation data of images on planktonic biodiversity	
SeaDataNet data products	Aggregated marine data collections and climatologies, such as for Temperature & Salinity	
ICOS-Marine	Long-term oceanic observations of carbon uptake and fluxes for understanding the global carbon cycle	
SOCAT - Surface Ocean CO2 Atlas	SOCAT version 2020 with quality-controlled surface ocean fCO2 measurements from 1957 to 2020	
EMODnet Bathymetry	EMODnet Bathymetry World Base Layer is used as base map in the interface	

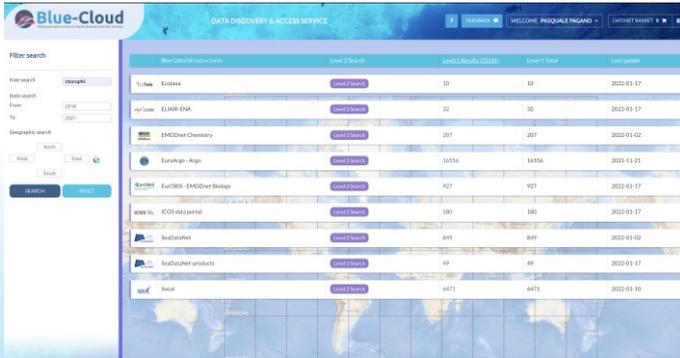


The first step enables users to identify interesting data collections, with free search, geographic and temporal criteria



The second step enables users to drill down per interesting BDI to get more specific data sets at granule level, using again free search, geographic and temporal criteria

One place for access

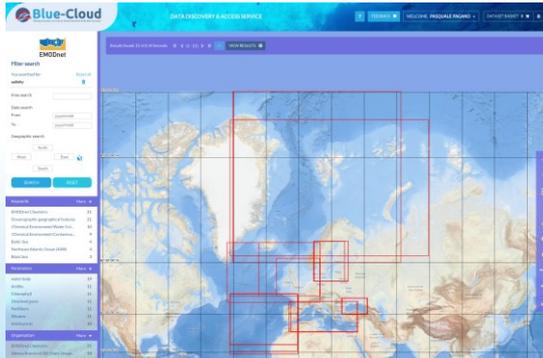


Blue-Cloud DATA DISCOVERY & ACCESS SERVICE

Filter search:

From: To: Geographic search:

Blue Data Infrastructure	Level 2 Search	Level 3 Results (2/28)	Level 3 Total	Last update
ELINE	Estima	<input type="button" value="Level 3 Search"/>	10	2022-01-17
ELINE-ENA	ELINE-ENA	<input type="button" value="Level 3 Search"/>	32	2022-01-17
ENDDnet Chemistry		<input type="button" value="Level 3 Search"/>	207	2022-01-02
EnrApkg - Arpa		<input type="button" value="Level 3 Search"/>	16556	2021-11-31
EnrObs - ENDDnet Biologi		<input type="button" value="Level 3 Search"/>	927	2022-01-17
KOSM	KOSM dataset	<input type="button" value="Level 3 Search"/>	100	2022-01-17
SeaDataNet		<input type="button" value="Level 3 Search"/>	849	2022-01-02
SeaDataNet products		<input type="button" value="Level 3 Search"/>	89	2022-01-17
Sea4	Sea4	<input type="button" value="Level 3 Search"/>	6471	2022-01-10



Blue-Cloud DATA DISCOVERY & ACCESS SERVICE

ENDDnet

Filter search:

Map showing geographical distribution of datasets with red bounding boxes.

Complete your order

Order #	ECI	Dataverse	Source	Description	ID
22987	426	SeaDataCloud Black Sea Ten Salinity Climatology V2	North Sea - Eutrophication and Acidity aggregated datasets 1974/2010-2021	Arctic Ocean - Contaminants aggregated datasets 1974/2015-2018	
22986	426	Black Sea gridded climatology datasets - 1st cycle intermediate content at 1°P	Baltic Sea - Contaminants aggregated datasets 1974/2018-2021		
22985	426	Mediterranean Sea - Temperature and salinity Historical Data Collection SeaDataNet V1		IT	02-11-2021 16:20 Downloaded 13-12-2021 16:05
22984	426	Baltic Sea - Temperature and salinity observation collection V2		IT	02-11-2021 16:20 Downloaded 13-12-2021 16:05
22983	425	AD1185h Sea	ENDDC_Buette_11185_13	IT	02-11-2021 16:20 Downloaded 13-12-2021 16:05
22982	425	ENDDC_Buette_10855	ENDDC_Buette_10855_18	IT	02-11-2021 16:20 Downloaded 13-12-2021 16:05
22981	426	ENDDC_Buette_10855	ENDDC_Buette_10855_1	IT	02-11-2021 16:19 Downloaded 13-12-2021 16:05
17835	417	Asteroid distribution data from Deep-sea beds of European seas - taxonomic species check-list of benthic invertebrates living deeper than 2000 m in the seas bordering Europe		IT	27-09-2021 11:03 Downloaded 02-12-2021 15:27

Compose and submit shopping request at the granule level



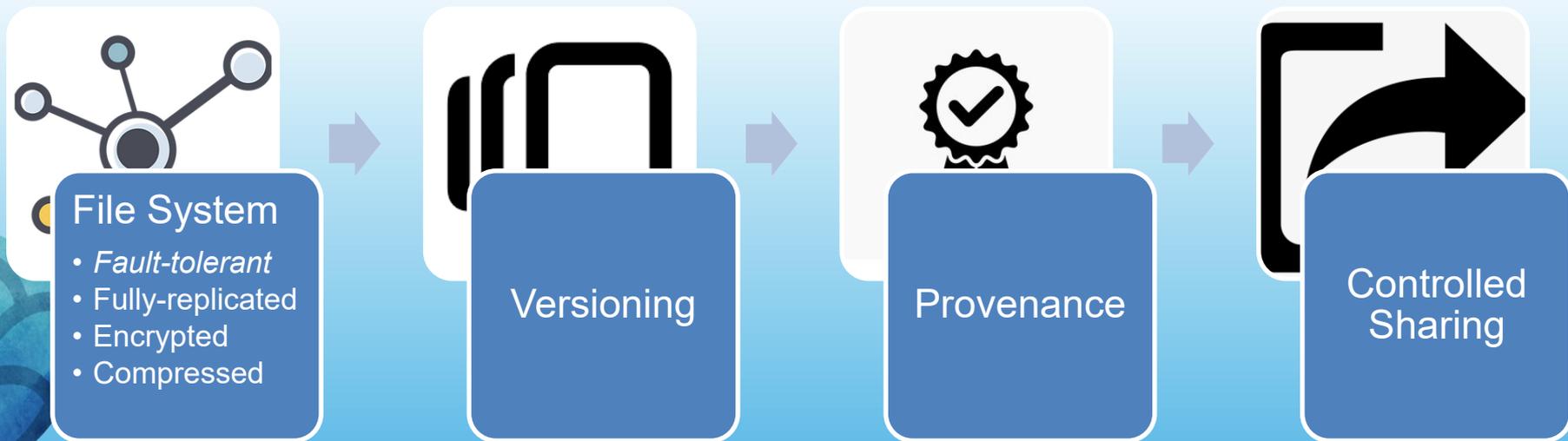
Retrieve the datasets by downloading from the Dashboard



Push datasets to the Blue-Cloud VRE Data Pool

One place for storing, sharing, and preserving datasets

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



Storage Layer

Persistent Storage

Workspace

Pros

Cons

Fault-tolerant

Replicated Distributed

Compressed

High Latency

Slower

Dataspace

Pros

Cons

Large Volume (1TB)

Faster

Low Latency

Single Site

Linked only to JupyterHub

Volatile Storage

Volatile

Pros

Cons

Faster

Single Site

Low Latency

Deleted after 24 hours

Accessible only via URIs

One place for executing analysis and processes

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

 <p>JupyterHub</p> <p>JupyterHub enables the exploitation of computational environments and resources without burdening users with installation and maintenance tasks. This JupyterHub environment is <i>(i)</i> preconfigured with libraries and packages to ease the execution of common data analytics tasks, and <i>(ii)</i> provides access to the Workspace enabling sharing of resources with other members much easier.</p>	 <p>RStudio</p> <p>RStudio provides an integrated development environment for R. It includes a console and a syntax-highlighting editor and it enables code execution. Tools for plotting are also included. This RStudio environment is <i>(i)</i> preconfigured with libraries and packages to ease the execution of common data analytics tasks, and <i>(ii)</i> provides seamless access to the Workspace enabling sharing of resources with other members much easier.</p>
 <p>Analytics Engine</p> <p>Analytics Engine (DataMiner) permits the execution of an array of analytics methods by transparently relying on distributed computing infrastructure. Executions can run either on multi-core machines or on different computational platforms, such as D4Science and other different private and commercial Cloud providers. New software can be integrated by using the dedicated Software Importer (SAI).</p>	 <p>Catalogue</p> <p>Catalogue contains a list of dataset and products produced by the Blue-Cloud Virtual Laboratories and the methods used to generate such products. All the Catalogue items are accompanied with rich descriptions capturing general attributes, e.g. title and creator(s); accessibility properties; technical properties, e.g. size and format; legal and ethical attributes, e.g. whether containing personal data; intellectual properties, e.g. licences.</p>

Development and integration environment for R, Python, and other supported software languages



- It is powered by a cluster of DataMiner servers, each with 16 cores and 32 GB RAM.
- It is powered by a cluster of RStudio servers, each with 16 cores and 32 GB RAM.
- It is powered by JupyterHub with a maximum of 8 cores and 32 GB RAM per notebook. Jupyter



One place to access tailored services



Biodiversity

Zoo and Phytoplankton EOY products



Genomics

Plankton Genomics



Environment

Marine Environmental Indicators



Fisheries

Global Record of Stocks and Fisheries



Aquaculture

Aquaculture Monitor



JERICO Core VLab



Joint Framework for Ocean Noise in the Atlantic Seas

JONAS VLab



One common place



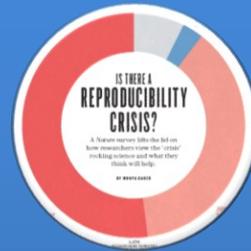
Collaboration



Sharing



Reuse



Reproducibility

Blue Cloud VRE promotes Open Science



Blue-Cloud

Unlocking
Open Science
in support of the
EU Green Deal

Website: www.blue-cloud.org

E-mail: info@blue-cloud.org

Twitter: [@BlueCloudEU](https://twitter.com/BlueCloudEU)

LinkedIn: [Blue-Cloud Org](https://www.linkedin.com/company/blue-cloud-org)

FINAL CONFERENCE