Blue-Cloud Position Paper on EOSC

prepared by Sara Garavelli, Trust-IT & Blue-Cloud Coordinator and Dick Schaap, MARIS & Blue-Cloud Technical Coordinator

endorsed by the Blue-Cloud Steering Committee & Consortium

"Europe is going to co-create a framework to allow the use of data. It should consist of a trusted pool of non personalized data that governments, businesses and other stakeholders can contribute to. And thereby, we open up data as a resource for innovation and bring new solutions to the market. And our scientists are already beginning to do this. We are creating a European Open Science Cloud (EOSC) now. It is a trusted space for researchers to store their data and to access data from researchers from other disciplines. We will create a pool of information leading to a web of research insight." Ursula von der Leyen, President of the European Commission.

On the 21st of January 2020, during the World Economic Forum Annual Meeting, Ursula von der Leyen, President of the European Commission pointed out the value of the European Open Science Cloud (EOSC) and the high importance of data as a key resource for innovation.

The recently funded H2020 Blue-Cloud[1] project fully shares these views by working towards the establishment of a thematic marine EOSC serving the Blue Economy, Marine Environment and Marine Knowledge agendas.

After the first communication on the EOSC initiative in 2016, several Calls have been published and several projects have been granted funding for developing (parts of) the EOSC environment. **Blue Cloud is the answer to the H2020 dedicated call[2] for 'The Future of Seas and Oceans Flagship Initiative'.**

At the opening of the All Atlantic Ocean Research Forum, 6-7 February 2020**, Mariya Gabriel, the European Commissioner for innovation, research, education and youth, highlighted Blue-Cloud a key instrument for the sustainable ocean strategy.**

[1] www.blue-cloud.org
[2] European Union's Horizon programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services.

Key messages for the EOSC Governance and the European Commission

- So far, the EOSC development has focused on e-infrastructures, with less attention placed on thematic research infrastructures. Blue-Cloud contributes to a more balanced approach by establishing cooperation between the blue thematic community and the horizontal e-infrastructures (including players such as: EUDAT, D4Science, OpenAIRE, EGI.eu, GÉANT, etc.). The requirements and perspectives of the blue thematic community will help to build a stronger, user oriented EOSC.
- Blue-Cloud is the equivalent of an INFRAEOSC04 cluster project specifically set up for the Blue Economy community. It is therefore of paramount importance that Blue-Cloud is part of the relevant discussions of the clusters together with the other thematic clouds (e.g. Food Cloud, Transport Cloud). Going forward, Blue-Cloud will play an increasingly relevant role in cross-disciplinary cooperation with the other EOSC-related clusters and thematic clouds.
- Blue-Cloud is piloting **5 actionable multi-disciplinary demonstrators** that address societal challenges. These demonstrators are used to showcase not only the potential of blue data but also the potential of EOSC to support thematic communities. The feedback collected by the scientists working on the demonstrators can support EOSC in better defining its value proposition and attract more users.
- In the European landscape of marine and ocean data management great progress has been made over three decades with **developing standards, services, and establishing dedicated infrastructures.** These infrastructures mainly provide data discovery and access services and focus on the associated challenge of making these services FAIR (Findable, Accessible, Interoperable, and Reusable). It's now the time to make a step forward by developing frameworks with Virtual Research Environments (VREs) for analytical processes by and for users to support their research.

The expectations of the Blue-Cloud community

Blue-Cloud supports the view of EOSC and expects EOSC to provide:

- Generic e-infrastructure services (e.g. AAI, PIDs, etc) and access to resources such as HPC, storage, cloud computing;
- Seamless access to other thematic clouds to support multi-disciplinary science;
- Easy to use APIs to make "Blue" services and data available to the target of the 1.7 Million researchers in Europe (and beyond) to accelerate FAIR science;
- A framework creating awareness on the importance and best practices for FAIR data and providing training on data management.

What Blue-Cloud brings to EOSC

- A pilot thematic-EOSC as a role model for the development of other thematic clouds. The cyber-platform of Blue-Cloud provides FAIR access to multidisciplinary data, analytical tools and computing and storage facilities that support research.
- Blue Cloud Services through pilot Demonstrators for oceans, seas and fresh water bodies for ecosystems research, conservation, forecasting and innovation in the Blue Economy, and making innovative use of seamless access to multidisciplinary data, algorithms, and computing resources - accelerating cross-discipline science.

Blue-Cloud has received funding from the European Union's Horizon programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, Grant Agreement No.862409.

- A methodology for researchers interacting with e-infrastructure developers to establish a cyber platform with tools and services, which support multiple scientific challenges and are fit-for-purpose, while built upon generic core principles and services.
- A mechanism to easily access and discover *blue* data. Blue-Cloud partners manage important volumes of *blue* data (e.g. SeaDataNet, EMODnet, CMEMS, etc.) and links have been established with major European observing networks to increase the data volume.
- APIs to access *blue* services that will complement EOSC base services providing blue thematic functionalities.
- Dynamic examples on how a framework like Blue-Cloud can address one or several of the policy challenges defined in the Bioeconomy Strategy, the Circular Economy Strategy, the Blue Growth Strategy, the Common Fisheries Policy, the Maritime Spatial Planning Directive, the Marine Strategy Framework Directive, the International Ocean Governance Communication and the UN SDGs.
- A global Blue-Economy community close to the EOSC vision, including the marine and maritime industry.
- The opportunity of bringing EOSC in the Blue Economy long-term vision via the policy oriented Blue Cloud Roadmap to 2030 which seeks a series of EU Calls for further development and uptake of the Blue Cloud by multiple VRE applications and connecting additional marine data infrastructures.

About Blue-Cloud

Blue-Cloud[3] aims at **developing a marine thematic EOSC to explore and demonstrate the potential of cloud based open science for better understanding and managing the many aspects of ocean sustainability.**

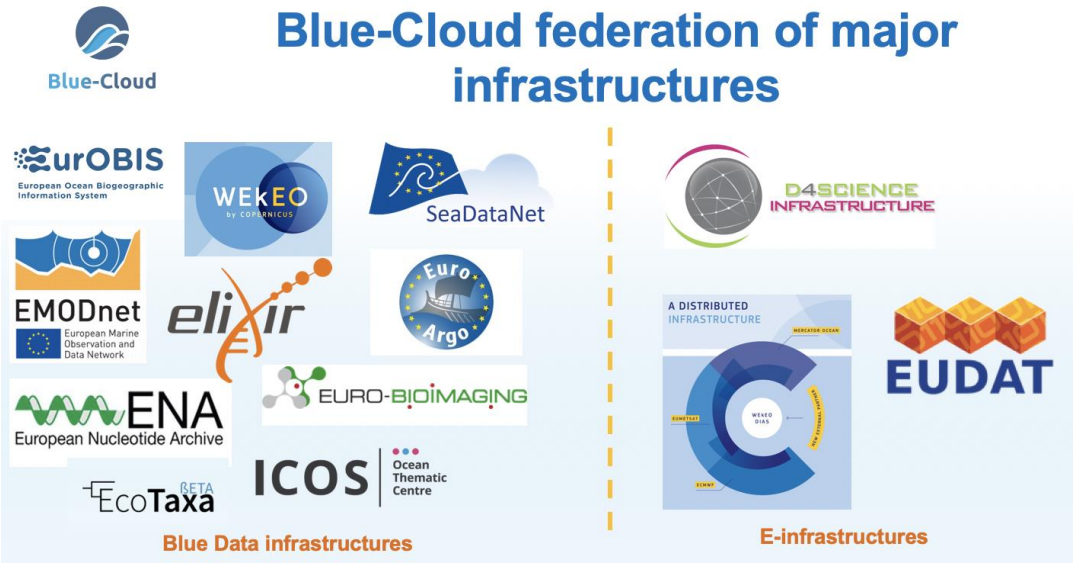
Blue-Cloud is the flagship project of the DG Research and Innovation Unit of the European Commission that will link the horizontal e-infrastructures supported by DG CNECT (e.g. EUDAT & D4Science) and DG GROW (e.g. Copernicus DIAS), long-term marine data initiative supported by DG MARE (e.g. EMODnet), research infrastructures supported by DG RTD and other recently funded thematic clouds (e.g. Food Cloud and Transport Cloud).

Blue-Cloud objectives are to:

- **build and demonstrate a Pilot Blue Cloud** by combining distributed marine data resources, computing platforms, and analytical services
- **develop services for supporting research to better understand & manage the many aspects of ocean sustainability** and promote innovation in the blue economy
- develop and validate **demonstrators of relevance for marine societal challenges**
- formulate a **roadmap for expansion and sustainability of the Blue Cloud infrastructure and services**.

[3] Blue-Cloud has received funding from the European Union's Horizon programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, Grant Agreement No.862409.

The project will **federate leading European blue data management infrastructures** (SeaDataNet[4], EurOBIS[5], Euro-Argo[6], Argo GDAC[7], EMODnet[8], ELIXIR-ENA[9], EuroBioImaging[10], CMEMS[11], C3S[12], and ICOS-Marine[13]), and **horizontal e-infrastructures** (EUDAT[14], DIAS[15], D4Science[16]) to capitalise on what already exists and to deliver the "Blue Cloud" framework. **The federation will be at the levels of data resources, computing resources and analytical service resources**.



Some of the infrastructures behind Blue-Cloud

A Blue Cloud **data discovery and access service** will be developed as an overarching service to facilitate smart sharing of multi-disciplinary datasets with human and machine users. A Blue Cloud **Virtual Research Environment (VRE) will orchestrate the computing and analytical services in specific integrated and managed applications that through the same VRE use the federated Blue-Cloud data resources in addition to external data resources.**

The infrastructures federated in Blue-Cloud are developed and operated by research, governmental, and industry organisations from European states, and in close interaction with several European and international initiatives. Each of these data infrastructures has established

[4] www.seadatanet.org
[5] www.eurobis.org
[6] www.euro-argo.eu
[7] www.argodatamgt.org
[8] www.emodnet.eu
[9] www.embric.eu/ELIXIR
[10] www.eurobioimaging.eu
[11] marine.copernicus.eu
[12] climate.copernicus.eu
[13] www.icos-ri.eu
[14] eudat.eu
[15] www.copernicus.eu/en/access-data/dias
[16] www.d4science.org

Blue-Cloud has received funding from the European Union's Horizon programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, Grant Agreement No.862409.

links to data originators and their data collection, facilitating to oversee and engage in the process from collection to validation to storage and distribution. Several are also increasingly involved in generating data products and knowledge by developing and deploying analytical workflows with various algorithms and numerical models. They are also mostly complementary to each other, dealing with other data types and/or data originators and/or different stages in the processing chains from data acquisition to data products to knowledge.

The D4Science e-infrastructure will provide the core platform for the Blue Cloud Virtual Research Environment, making optimal use of the major experiences gained in the BlueBRIDGE project with developing and operating a generic VRE with many core services for building and running multiple Virtual Labs, each dedicated to specific research targets. D4Science has proven solutions for connecting to external computing platforms and means for orchestrating distributed services. This will be instrumental for the smart federation with the EUDAT and DIAS e-infrastructures.

The Blue-Cloud innovation potential will be explored and unlocked by developing five dedicated Demonstrators as Virtual Labs together with excellent marine researchers.

The five demonstrators are focusing on:

1. **Zoo- and Phytoplankton EOV products**, led by the Marine Flanders Institute (VLIZ): The proposed demonstrator will build upon a range of oceanographic data from multiple streams, made interoperable and integrated through Blue Cloud services, in order to produce unique 3D and 4D synergistic zooplankton and phytoplankton products that will be made available to various end-users. The virtual EOVs to be produced, will contribute to improve knowledge and quantitatively reduce uncertainty regarding the present state of the marine plankton ecosystems and their response to ongoing and future climate change.
2. **Plankton Genomics**, led by the European Molecular Biology Laboratory (EMBL): The demonstrator will showcase a deep assessment of plankton distributions, dynamics and fine-grained diversity to molecular resolution, focusing on species and functions discovery and exploration of genetic and morphological markers of plankton diversity and abundance.
3. **Marine Environmental Indicators**, led by the Euro-Mediterranean Center on Climate Change (CMCC): The demonstrator will develop an online service with associated cloud based analytical computing framework and dedicated web interface to provide and display indicators and information on the environmental quality of the ocean.
4. **Fish, a matter of scales**, coordinated by the Food and Agriculture Organisation of the UN (FAO): The demonstrator will improve data management and analytic capabilities for fisheries by building a global vertically integrated toolset that will manage public fisheries statistical data from ingestion, through harmonization, to publication.
5. **Aquaculture Monitor**: coordinated by the Food and Agriculture Organisation of the UN (FAO): The demonstrator will provide a robust and replicable environment for monitoring aquaculture in marine cages and in coastal areas making extensive use of Copernicus data, and combines AI with in-situ datasets to obtain inventories at regional level.

There is already a large portfolio of services managed by the Blue Cloud founders which will be integrated. The modular architecture of the VRE does allow scalability and sustainability for

Blue-Cloud has received funding from the European Union's Horizon programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, Grant Agreement No.862409.

further expansions, such as connecting additional infrastructures, implementing more blue analytical services, configuring more dedicated Virtual Labs, and targeting more (groups of) users.

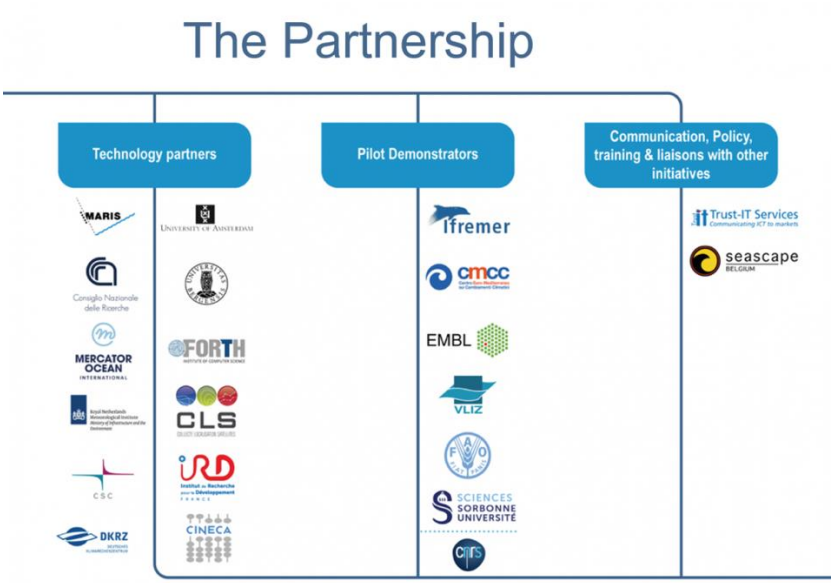
Finally, Blue-Cloud will deliver a pragmatic, policy oriented **Blue Cloud Roadmap to 2030** which seeks a series of EU Calls for further development and uptake of the Blue Cloud by multiple VRE applications and connecting additional marine data infrastructures.

Complementarity to the ENVRI-FAIR Cluster

The main objective of the ENVRI-FAIR cluster is to further develop common standards, protocols and policies for the data life cycle. Several Blue-Cloud partners are members of ENVRI-FAIR to make their discovery and access services FAIRer, e.g. by improving metadata richness, further uptake of controlled vocabularies, and developing or upgrading machine-to-machine services. The FAIR analyses developed by the cluster are directly beneficial to the Blue-Cloud development.
However, the Blue-Cloud goes beyond ENVRI-FAIR as the Blue-Cloud will establish an operational smart discovery and access service, a Virtual Research Environment with services for building, deploying, and operating Virtual Labs, and with solutions for interacting with other data, computing, and analytical infrastructures.

More information on Blue-Cloud are available at www.blue-cloud.org
Twitter @BlueCloudEU
LinkedIn https://www.linkedin.com/company/blue-cloud-org/
YouTube https://www.youtube.com/channel/UChYrsWsZuIrjVCAUsBfw8PA

Blue Cloud consortium



Blue-Cloud has received funding from the European Union's Horizon programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, Grant Agreement No.862409.