D4.1 Blue Cloud VRE Operation Report (Release 1)

Work Package	WP4, Developing and operating the Blue Cloud VRE, its services and Virtual Labs
Lead Partner	CNR
Lead Authors (Org)	Massimiliano Assante (CNR), Leonardo Candela (CNR), Pasquale Pagano (CNR)
Contributing Author(s)	Roberto Cirillo (CNR), Andrea Dell'Amico (CNR), Luca Frosini (CNR), Lucio Lelii (CNR), Francesco Mangiacrapa (CNR), Giancarlo Panichi (CNR), Fabio Sinibaldi (CNR)
Reviewers	Patricia Cabrera, Paul Focke (VLIZ), Gilbert Maudire, Cécile Nys (Ifremer) Dick M.A. Schaap (MARIS)
Due Date	28-02-2021, M17
Submission Date	17-03-2021
Version	1.0

Dissemination Level

X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Blue-Clou Innovative services for Marine Research & the Blue Economy - has received funding from the European Union's Horizon p call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, Grant Agreement n. 862409.

DISCLAIMER

"Blue-Cloud, Piloting Innovative services for Marine Research & the Blue Economy" has received funding from the European Union's Horizon programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, Grant Agreement n.862409.

This document contains information on Blue-Cloud core activities. Any reference to content in this document should clearly indicate the authors, source, organisation, and publication date.

The document has been produced with the funding of the European Commission. The content of this publication is the sole responsibility of the Blue-Cloud Consortium, and it cannot be considered to reflect the views of the European Commission. The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

COPYRIGHT NOTICE

VERSIONING AND CONTRIBUTION HISTORY

Version	*Date*	*Authors*	*Notes*
0.1	15/01/2021	CNR	Definition of the deliverable Table of Contents and planning of contributions
0.5	10/02/2020	CNR	The first version of planned contributions has been integrated into the overall document.
0.6	09/03/2021	Patricia Cabrera (VLiz)	Internal review
0.7	11/03/2021	CNR	Comments and suggestions proposed by the reviewer have been analysed and managed.
0.8	12/03/2021	Gilbert Maudire (Ifremer)	Internal review
0.9	15/03/2021	CNR	Comments and suggestions proposed by the reviewer have been analysed and managed.
0.10	16/03/2021	Dick M.A. Schaap (MARIS)	Final review as Technical Coordinator
1.0	17/03/2021	Sara Pittonet (Trust-IT)	Submission to the EC portal

Table of Contents

Table of Figures

Table of Tables

Executive summary

The Blue-Cloud project is developing a cyber platform bringing together and providing access to multidisciplinary data from observations and models, analytical tools, and computing facilities essential to support research to understand better and manage the many aspects of ocean sustainability. This is achieved by the Blue Cloud platform, whose architecture consists of two major families of components: (a) the *Blue Cloud Data Discovery and Access* service component to serve federated discovery and access to 'blue data' infrastructures, and (b) the *Blue Cloud Virtual Research Environment (VRE)* component to provide a Blue Cloud VRE as a federation of computing platforms and analytical services.

This deliverable presents the Blue Cloud Virtual Research Environment by complementing the picture described in Assante et al. (2020b), where the constituents have been discussed. In particular, this deliverable focuses on how the components have been exploited and operated to support the development of the Blue Cloud gateway https://blue-cloud.d4science.org, its underlying infrastructure, and the VLabs.

As of February 2021, ***9 Blue-Cloud VLabs*** were created and operated. In particular, the following VLabs are specifically conceived to support the developments of the Blue-Cloud Demonstrators: (i) The Aquaculture Atlas Generation VLab is developed in the context of the Demonstrator #5 - Aquaculture monitor; (ii) The Fisheries Atlas VLab is developed in the context of the Demonstrator #4 - Fish, a matter of scales; (iii) The GRSF_Pre VLab is developed in the context of the Demonstrator #4 - Fish, a matter of scales; (iv) The Marine Environmental Indicators VLab is developed in the context of the Demonstrator #3 - Marine Environmental Indicators; and (v) The Zoo and Phytoplankton EOV VLab is developed in the context of the Demonstrator #1 - Zoo- and Phytoplankton EOV products. The following VLabs are conceived to support project activities and/or to provide their users with development and demonstrative environments: (i) The Alien and Invasive Species VLab is a demonstration oriented VLab showcasing how a web-based, comprehensive, and collaborative working environment supporting decision-makers and scientists in predicting the spread of an invasive species (possibly alien) in a new environment can be developed; (ii) The Blue-Cloud Lab VLab is developed to provide the Blue-Cloud community with a working environment to experiment with Blue-Cloud facilities; (iii) The Blue-Cloud Project VLab was devised to support Blue-Cloud project activities and discussions; and (iv) The HealthyOcean VLab was developed to support the liaison between the REV Ocean team[1] and the Blue-Cloud team regarding several initiatives including the Ocean Data Platform.

These working environments are serving more than ***490 users*** in total spread across more than ***20 countries***.

Up to January 2021, a total of more than ***10,200 working sessions*** have been executed by these users, with an average of ***640 working sessions per month,*** since the start of the Blue-Cloud project in November 2019. Overall, a total of ***1573 analytics tasks*** have been executed, with an average of 98 tasks per month. These figures are expected to grow in the second period of the project, thanks to the development and availability of VLabs stemming from the demonstrators.

[1] REV Ocean website https://www.revocean.org/

During the reporting period, a total of 113 tickets (e.g., requests for support, tasks, incidents) have been created and managed in the specific area dedicated to Blue-Cloud (75% have been closed). Moreover, 44 tickets related to Blue-Cloud have been created in the D4Science overall context (79.5% have been closed).

1. Introduction

The Blue Cloud Architecture (Schaap et al. 2020) consists of two major families of components:

- the ***Blue Cloud Data Discovery and Access*** component to serve federated discovery and access to blue data infrastructures;
- the ***Blue Cloud Virtual Research Environment (VRE)*** component to provide a Blue Cloud VRE as a federation of computing platforms and analytical services.

The Blue Cloud Virtual Research Environment components range from services to promote collaboration among its users to services supporting the execution of analytics tasks embedded in a distributed computing infrastructure and services enabling the co-creation of entire Virtual Laboratories (Assante et al. 2020b). This part of the overall Blue Cloud platform is built on the D4Science infrastructure and the gCube open-source technology (Assante et al. 2019a, 2019b) and deployed in the Blue Cloud gateway (accessible at https://blue-cloud.d4science.org) to make the services and Virtual Laboratories available.

The Blue Cloud platform development is driven by the needs and requirements underlying the development of selected demonstrators (Maudire et al., 2020). An up-to-date picture of the development of the demonstrators is described in Nys et al. (2020). In reality, this is the result of a co-creation approach in VLab development (Assante et al. 2020a). The co-creation process, when exploited to promote the development of VLabs, suggests following a participatory process where the activities of software developers and service providers are intertwined with the activities of the VLab designated community that bring specific value to the VLab. In fact, VLabs consists of two complementary parts: a) the *community-agnostic part*, i.e., services offering basic or advanced functionality exposing a common behavior when instantiated in diverse contexts; b) the *community-specific part*, i.e., services offering a peculiar functionality or data, sometimes implemented by combining into specific workflows the community-agnostic part with context-specific services or data. The collaborative and participatory development of these two parts makes possible for (a) community-agnostic software developers and service providers to incrementally develop solutions matching the needs of diverse communities by promptly testing them in real usage scenarios; (b) VLab designated community members to actively contribute to the incremental development of the VLab by bringing in their peculiarity and diversity.

This deliverable documents how the components of the Blue Cloud platform have been exploited and operated to support the development of the Blue Cloud gateway https://blue-cloud.d4science.org, its underlying infrastructure, and the VLabs.

The deliverable is organised as follows. Section 2 describes the policies and procedures governing the planning and deployment of Virtual Laboratories. Section 3 describes the Virtual Laboratories that have been deployed and operated during the period. For each Virtual Laboratory, the deliverable describes the goal and the main facilities offered to their users. Section 4 concludes the report by describing some concluding remarks.

2. Blue-Cloud VRE and VLabs Planning and Procedures

The Blue-Cloud VRE is built on the D4Science infrastructure and the gCube open-source technology (Assante et al. 2019a, 2019b). From the end-user point of view, it manifests in the Blue Cloud gateway (accessible at https://blue-cloud.d4science.org), the access point to the services and Virtual Laboratories available to the Blue-Cloud.

The development of the Blue-Cloud VRE counts on the availability of new versions of the enabling technology that are made available by https://code-repo.d4science.org/gCubeCI/gCubeReleases. These versions are produced by taking into account the requirements (with the relative priority) formulated by the Blue-Cloud community via the specification of the demonstrators (Maudire et al. 2020, Nys et al. 2020) that might correspond to new facilities to be developed or request for enhancements of existing facilities as well as requests for resolving malfunctions.

The technology supporting the development of the Blue-Cloud VRE was included in the following 14 gCube open-source software releases that have been deployed into the production infrastructure: 4.16 (Nov. 2019), 4.17 (Dec. 2019), 4.18 (Dec. 2019), 4.19 (Feb. 2020), 4.20 (Feb. 2020), 4.21 (Mar. 2020), 4.22 (May 2020), 4.23 (Jun. 2020), 4.24 (Jul. 2020), 4.25 (Oct. 2020), 4.25.1 (Oct. 2020), 4.26 (Nov. 2020), 4.27 (Dec. 2020), and 4.28 (Feb. 2021).

All the requests are modelled and managed by an issue tracker operated by D4Science and available at https://support.d4science.org. For the needs of the Blue-Cloud community, a specific project has been created https://support.d4science.org/projects/blue-cloud (Figure 1) and configured to make it possible to create tickets for tasks, requests for support, incidents, VLabs creation, and request for specific services provisioning.

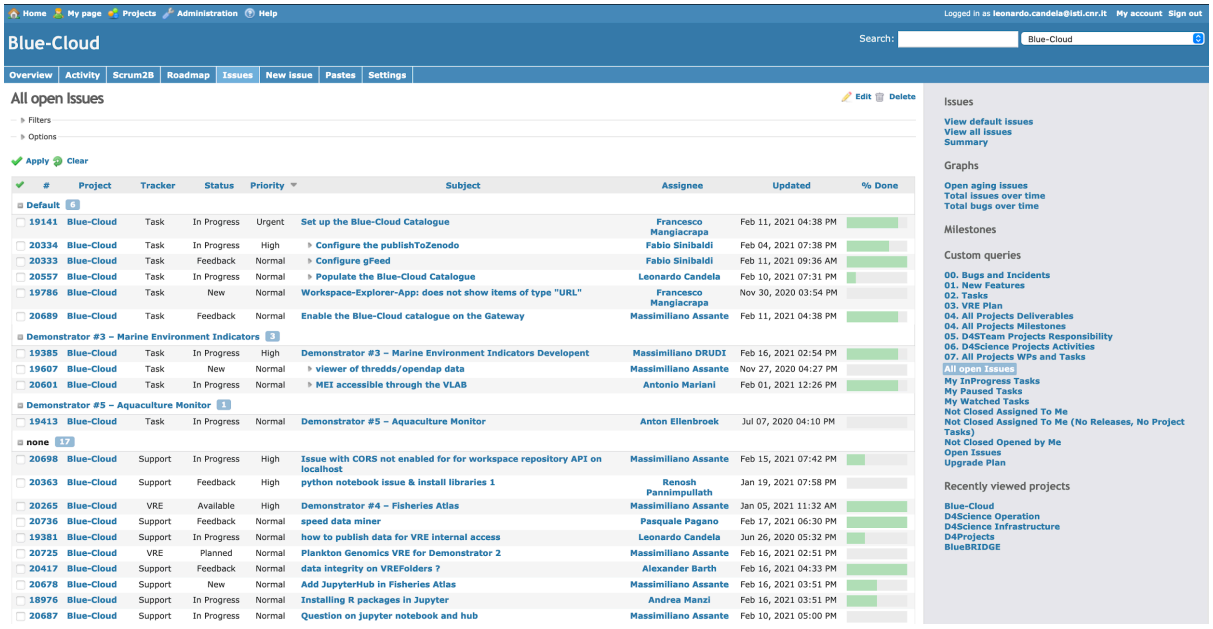

Figure 1. A screenshot of the Blue-Cloud issues tracker

2.1 Procedures

Deployment and operation of VLabs is a collaborative effort involving the WP4 team called to deploy and configure the technology to create VLabs expected by the work package working to jointly develop and onboard their technology, i.e., WP3.

The procedure leading to VRE and VLabs deployment is a consolidated one, i.e., it is the procedure inherited from the D4Science infrastructure and described in the D4Science Wiki:

https://wiki.d4science.org/index.php?title=Virtual_Research_Environments_Deployment_and_Operation

For the needs of Blue-Cloud, it was decided to support this activity by the project activity tracker. A specific VRE tracker has been created with the goal of capturing the entire process from specification to operation. The specification of the VRE/VLab is produced by the VRE/VLab designer/requester. This specification must contain:

- VLab name and abstract;
- Membership policy, i.e., whether the VLab is open or restricted, who is allowed to invite members; VRE expected datasets;
- VLab expected functionalities;
- VLab due date;

The following statuses are supported:

- ***Planned***: the WP4 team is fine with the specification, i.e., the specification contains enough details to proceed with the creation, and acknowledges that the creation of the VLab is feasible by the due date initially requested (or liaise with the designer/requester to find a mutually suitable date);
- ***Available***: the VLab is up and running and ready to be validated by the user designer/requester;
- ***Released***: the VLab has been validated, and the target community can start using it;
- ***Removed***: the VLab has been disposed as for the request of its manager;
- ***Rejected***: the requested VLab cannot be created as the requirements outlined for it cannot be satisfied.

3. Blue-Cloud VRE and VLabs Creation, Deployment, and Operation

This section briefly describes the facilities used by VLab creators for the actual deployment of VLabs, reports the complete list of deployed and operated VLabs, and offers a characterisation of each available VLab.

The act of definition and deployment of a new VLab is supported by a wizard (cf. Figure 2) that enables authorised users to transform the opened requests according to the procedure described in Sec. 2 into an actual specification and then, automatically, into a working VLab made available by the Blue-Cloud gateway. Through the wizard, the user is requested to specify: (i) the descriptive information characterising the expected VLab (i.e., name, description, duration), and (ii) the functionalities and datasets to be made available in the specific VLab by selecting among the available ones. The resulting list of functionalities is derived from the feasible functionalities created thanks to the software version and services hosted by the underlying infrastructure.

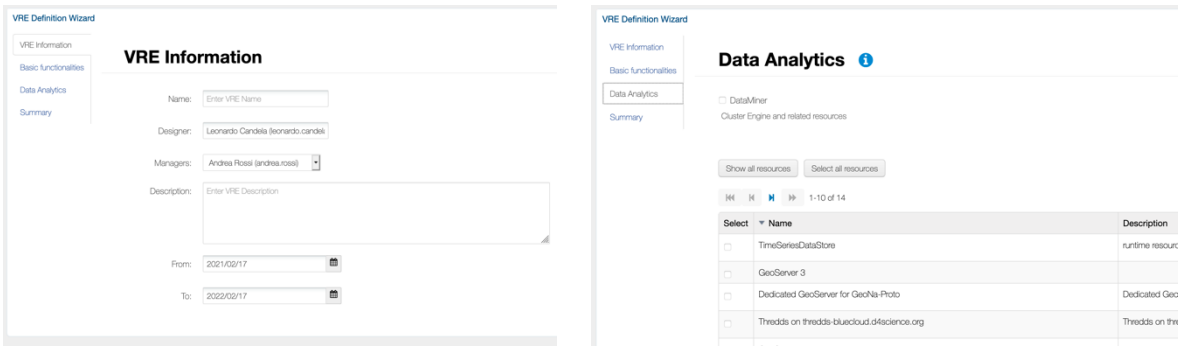

Figure 2. VLab Creation Wizard Screenshots

A total of one ***9 VLabs*** were created and/or operated to serve the needs arising in the context of the Blue-Cloud project (the complete list is in Table 1). In particular, the following VLabs are specifically conceived to support the developments of the Blue-Cloud Demonstrators:

(i) The Aquaculture Atlas Generation VLab is developed in the context of the Demonstrator #5 - Aquaculture monitor (cf. Sec. 3.2);

(ii) The Fisheries Atlas VLab is developed in the context of the Demonstrator #4 - Fish, a matter of scales (cf. Sec. 3.5);

(iii) The GRSF_Pre VLab is developed in the context of the Demonstrator #4 - Fish, a matter of scales (cf. Sec. 3.6);

(iv) The Marine Environmental Indicators VLab is developed in the context of the Demonstrator #3 - Marine Environmental Indicators (cf. Sec. 3.8); and

(v) The Zoo and Phytoplankton EOV VLab is developed in the context of the Demonstrator #1 - Zoo- and Phytoplankton EOV products (cf. Sec. 3.9).

The following VLabs are conceived to support project activities and/or to provide their users with development and demonstrative environments:

i. The Alien and Invasive Species VLab is a demonstration oriented VLab showcasing how a web-based, comprehensive, collaborative working environment supporting decision-makers and scientists in predicting the spread of an invasive species (possibly alien) in a new environment can be developed (cf. Sec. 3.1);
ii. The Blue-Cloud Lab VLab is developed to provide the Blue-Cloud community with a working environment to experiment with Blue-Cloud facilities (cf. Sec. 3.3);
iii. The Blue-Cloud Project VLab was devised to support Blue-Cloud project activities and discussions (cf. Sec. 3.4); and
iv. The HealthyOcean VLab was developed to support the liaison between the REV Ocean team[2] and the Blue-Cloud team regarding several initiatives including the Ocean Data Platform (cf. Sec. 3.7).

In Figure 3, the number of VLabs operated per month is reported. During the first months of the project, available VLabs include those inherited by previous initiatives (namely, BlueBRIDGE) and those created for supporting project activities. From April '20, new VLabs began being deployed to serve the needs of Blue-Cloud demonstrators (Nys et al. 2020).

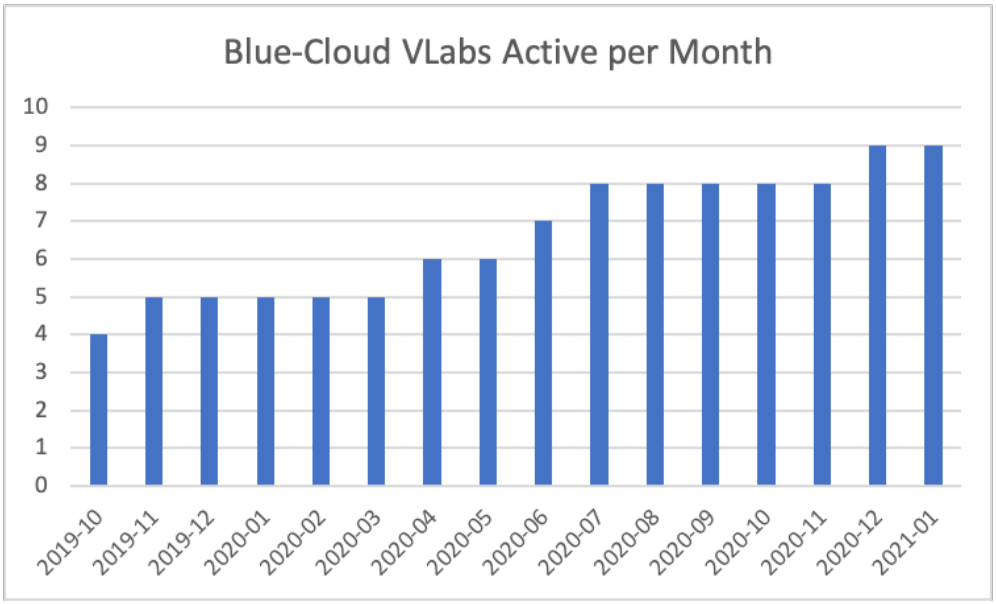

Figure 3. Number of VLabs operated per month (Oct. '19 - Jan. '21)

In Figure 4 it is reported the distribution of VLabs with respect to the membership policy. It can be observed that 56% of the VLabs are "restricted", i.e., users can request to join, yet the requests have to be explicitly approved by the managers. This is because the developed environments are still under development or conceived to serve the needs of known communities of practices. The 22% circa of the deployed VREs / VLabs is "restricted", i.e., membership is by invitation only. Finally, 22% of VLabs are "open", i.e., any user can request to join, and no approval is needed.

[2] REV Ocean website https://www.revocean.org/

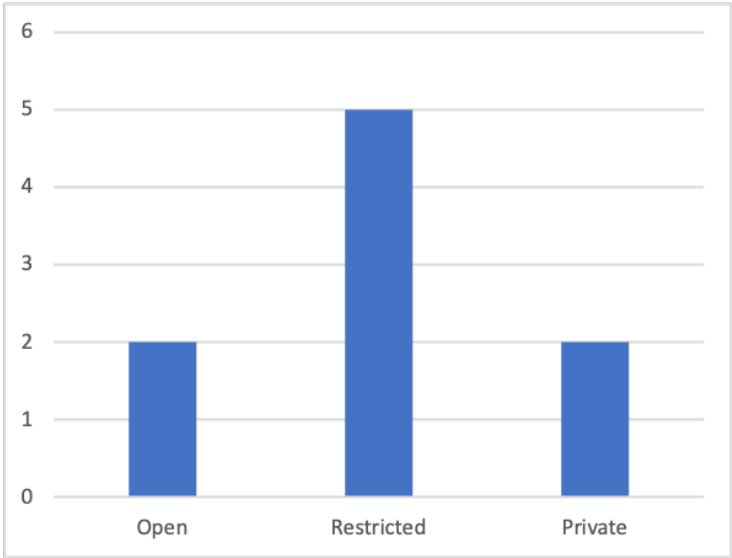

Figure 4. number of VLabs by membership typology

In Figure 5, the overall number of users benefitting from the facilities offered by the existing VLabs is reported, i.e., in January '21, the 9 existing VLabs are serving more than 490 users. Detailed figures per VLab are reported in the VLab dedicated sections of this document.

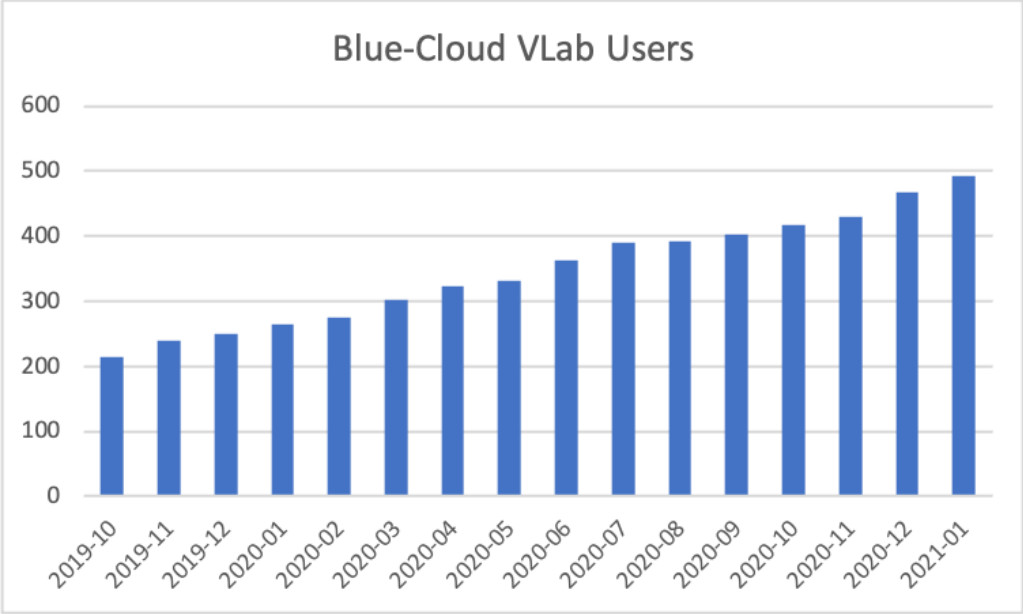

Figure 5. Number of users served by Blue-Cloud VLabs

Figure 6 reports the distribution of users with respect to VLab membership policy: the 46% circa of the served users (227 out of 492) is exploiting "open" environments, the 37% circa (183 out of 492) is exploiting "restricted" environments, and the 17% circa (82 out of 492) is exploiting "private" environments.

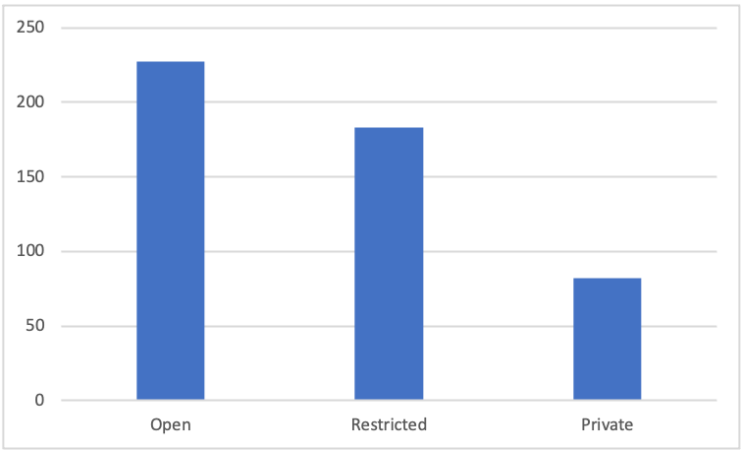

Figure 6. Number of users per VLab membership typology

By analysing the email addresses of the users (which is what they are using to log in), it can be observed that: 56% of the users are exploiting an email address that can be attributed to national domains (e.g., .it, .fr, .be) while the remaining 44% of the users are exploiting email addresses provided by commercial providers (e.g., google.com) or international organizations (e.g., fao.org). The users exploiting an email address that can be assimilated to national domains are spread across 20 countries. The top 3 countries are Italy (18%), France (11%), and Belgium (8%).

Figure 7 reports the overall number of working sessions initiated per month via the Blue-Cloud VRE. Up to January 2021, a total of more than 10,200 working sessions have been executed by the users, with an average of 640 working sessions per month.

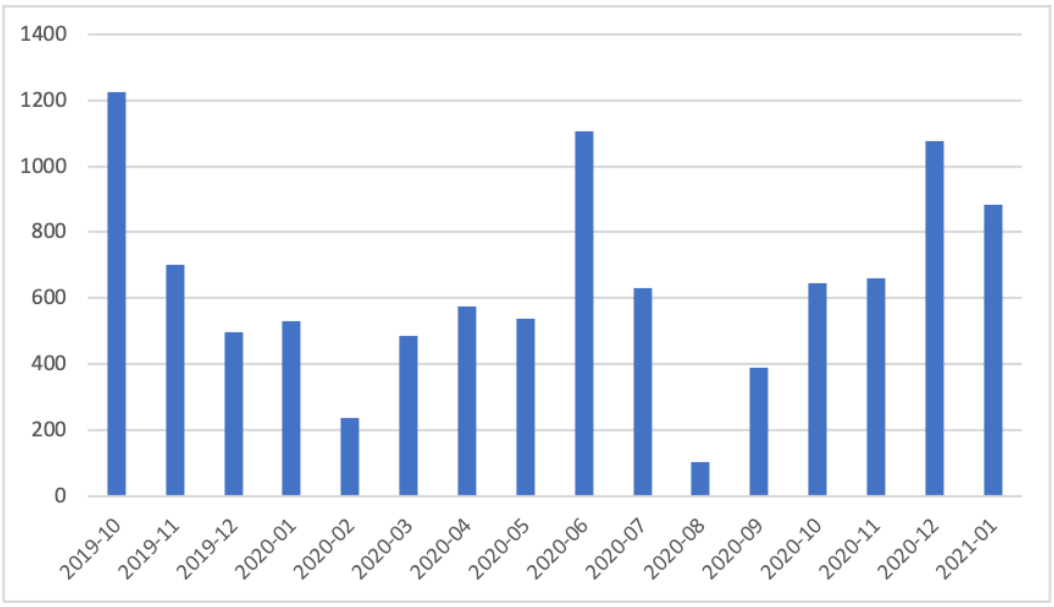

Figure 7. Accesses per month

The operation of the VRE and the VLabs requires the management of requests for support, issues, and malfunctions. A total of 113 tickets have been created and managed (Table 1) in the specific area dedicated to Blue-Cloud (75% have been closed). Moreover, 44 tickets related to Blue-Cloud have been created in the D4Science overall context (79.5% have been closed).

Table 1. Tickets created in the Blue-Cloud project

Ticket type	*Total number*	*Open / in progress*	*Closed*
Incident	11	1	10
Support	57	12	45
Task	34	10	24
VLab	7	5	2
Other	4	1	3

Table 2 reports the complete list of VLabs created and/or operated during the reporting period (from October 2019 up to January 2021).

Table 2. List of Blue-Cloud VLabs

VLab name	**Start date**	**Membership**	**#Users[3]**
Alien and Invasive Species	Oct. '19	Open	159
Aquaculture Atlas Generation	Oct. '19	Private	70
Blue-Cloud Lab	Oct. '19	Open	68
Blue-Cloud Project	Oct. '19	Restricted	112
Fisheries Atlas	Dec. '20	Restricted	15
GRSF_Pre	Jun. '20	Restricted	12
HealthyOcean Lab	Nov. '19	Restricted	8
Marine Environmental Indicators	Apr. '20	Restricted	36
Zoo and Phytoplankton EOV	Jul. '20	Private	12

A brief description of each available VLab is reported in the following sections.

All the VLabs are equipped with:

- A ***shared workspace*** to enable every user to store and organise the information objects he/she is interested in working with. In addition to that, the user is allowed to collaborate with other users by sharing objects and messages;
- A ***user management facility*** to enable authorized users (i.e., VLab Managers) to manage other users using or wanting to access the VLab. VLab Managers can (i) authorize users for

[3] Number of members of the VLab in January 2021.

access to the VLab, (ii) assign or withdraw roles to users, (iii) remove users, and (iv) send communications to the current users;

- A ***social networking facility*** to enable users to use the common facilities typical of social networks – e.g., posting news, commenting on posted news – yet adapted to the settings of working environments like those characterising Blue-Cloud. Users can post news as well as applications;
- A ***notification facility*** to alert users on relevant activities as they happen. These notifications offer a sense of anticipation and create a productivity boost. Users receive an alert (through a priori selected channels, e.g., email, web portal, Twitter) notifying them when something of interest has happened in their VLab(s);
- A ***members facility*** to provide users with a list of VLab co-workers, i.e., the list of members partaking in the VRE and contributing to it;
- A ***messaging facility*** to provide users with a cloud-based common email environment. The distinguishing feature is represented by its integration with the rest, e.g., it is possible to send any information object residing in the workspace (regardless of how "big" and "complex" it may be) as an attachment without consuming bandwidth.

3.1 Alien and Invasive Species

The Alien and Invasive Species VLab is a web-based, comprehensive, collaborative working environment supporting decision-makers and scientists in predicting the spread of an invasive species (possibly alien) in a new environment. The VLab hosts' examples of suitable habitat maps produced for today and 2050 in new areas for more than 11,000 species. It also provides models and workflows to combine environmental data with species observations in their habitats to predict their future spread.

This VLab is available at https://blue-cloud.d4science.org/web/alienandinvasivespecies

This VLab stems from previous initiatives (BlueBRIDGE project). From the Blue-Cloud point of view, it has been in operational status since October '19, i.e., when the Blue-Cloud gateway was released. The uptake of this VLab from Blue-Cloud was due to the willingness to have a developed working environment to be used for demonstrative purposes for Demonstrator designers and implementers. It is currently serving 159 users.

A screenshot of the VLab is available in Figure 8. It shows the home page and the menu items for accessing the VLab facilities.

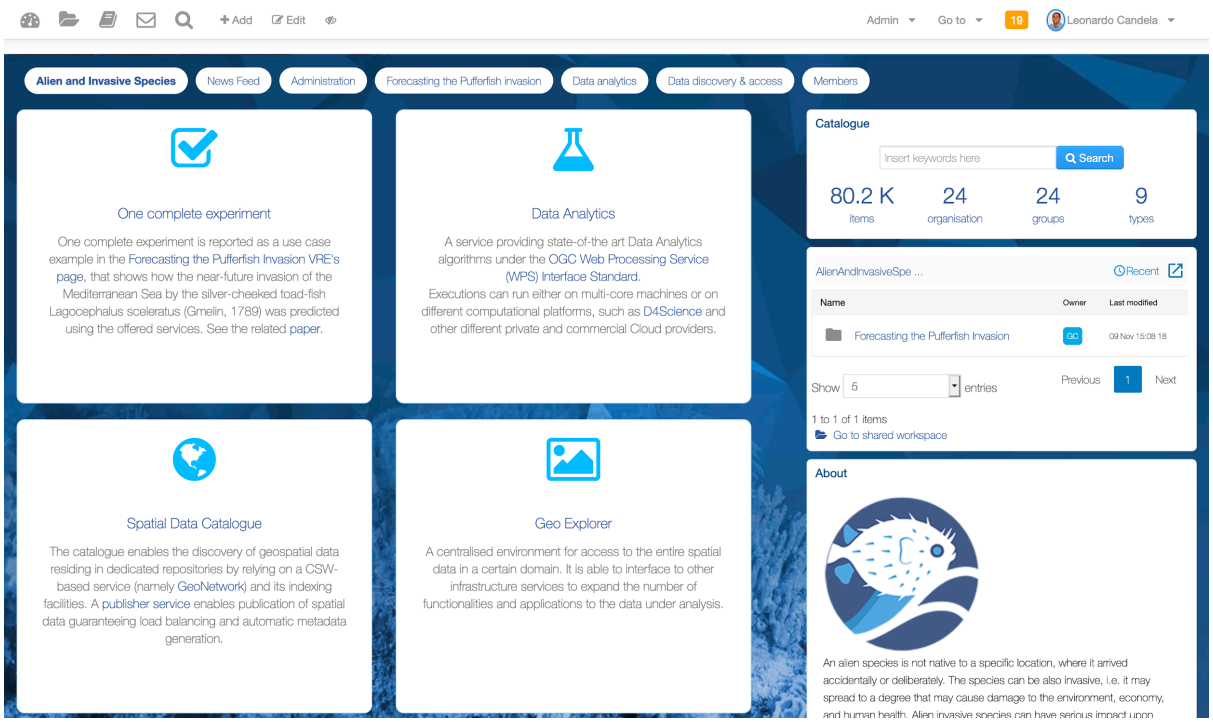

Figure 8. A screenshot of the Alien and Invasive Species VLab

In addition to the basic functionality, e.g., a workspace for sharing objects of interest and a user management facility for managing membership, this VLab is specifically equipped with the following functionalities:

- a ***data analytics facility*** enabling users to benefit from the offerings of the Data Miner service (Assante et al. 2019a, Assante et al. 2019b) and interactively execute a large array of data analytics tasks on datasets. As of January '21 the facility offers ***66 ready to use method implementations*** ranging from Bayesian methods to maps comparison and data clustering;
- a ***geospatial data discovery facility*** enabling users to have access to species distribution maps;
- a ***species data discovery facility*** enabling users to have seamless access to species data (including occurrence points) across several data providers;
- a ***catalogue facility*** enabling users to have access to more than ***70k objects*** including datasets, research objects, training material, code lists, and series;

3.2 Aquaculture Atlas Generation

The Aquaculture Atlas Generation VLab is developed in the context of the Demonstrator #5 - Aquaculture monitor (Nys et al. 2020).

This VLab is available at https://blue-cloud.d4science.org/web/aquacultureatlasgeneration

This VLab actually stems from the BlueBRIDGE project and thanks to the demonstrator implementation it will be extended. From the Blue-Cloud point of view, it has been in operational

status since October ’19, i.e., when the Blue-Cloud gateway was released. It is currently serving 70 users.

A screenshot of the VLab is available in Figure 9. It shows the home page and the menu items for accessing the VLab facilities.

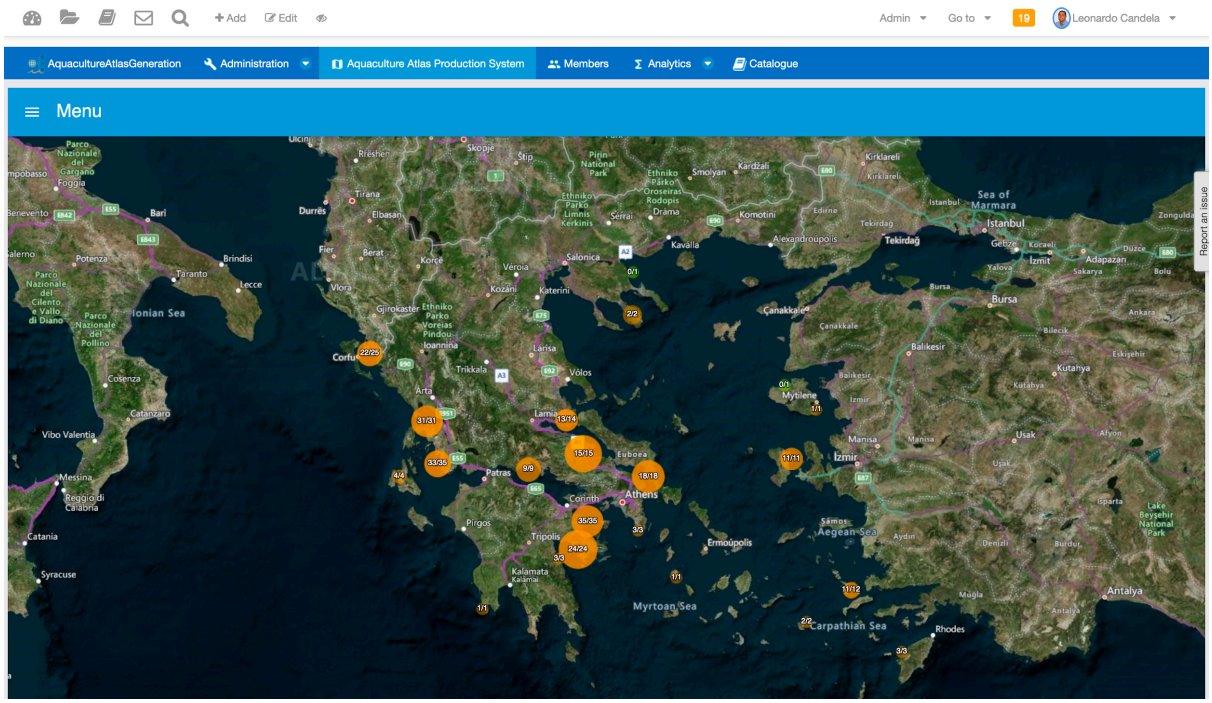

Figure 9. A screenshot of the Aquaculture Atlas Generation VLab

In addition to the basic functionality, e.g., a workspace for sharing objects of interest and a user management facility for managing membership, this VLab is specifically equipped with the following functionalities:

- the ***Aquaculture Atlas Production Systems*** map-viewer enabling users to identify and assess (for authorised users only) aquaculture farms. In particular, the tool visualizes the results of fish cage detection algorithms based on VHR optical images (ACUITY toolbox) and allows users to edit map features and to generate output as FAO NASO Maps;
- the ***data analytics*** facility enabling users to benefit from the offerings of the Data Miner service (Assante et al. 2019a, Assante et al. 2019b) and interactively execute a large array of data analytics tasks on datasets. It has been configured to serve ***4 specific algorithms***, i.e., those underlying the Aquaculture Atlas Production System facility (Geographic Proximity Tool, AAPS Public Publisher, AAPS Staging Publisher, AAPS NASO Publisher);
- the ***RStudio*** facility enabling users to use a fully-fledged RStudio® working environment directly from the VRE. This environment is integrated with the rest of VLab facilities, e.g., it is possible to use files from the workspace and to store new files within the workspace;
- the ***Catalogue*** facility enabling users to publish, search and browse datasets and other products of interest for the specific use case. This environment has been configured to give access to datasets stemming from the Greece and Indonesia cases.

3.3 Blue-Cloud Lab

The Blue-Cloud Lab VLab is developed to provide the Blue-Cloud community with a working environment to experiment with Blue-Cloud facilities.

This VLab is available at https://blue-cloud.d4science.org/web/blue-cloudlab

This VLab has been in operational status since October '19, i.e., when the Blue-Cloud gateway was released. It is currently serving 68 users.

A screenshot of the VLab is available in Figure 10. It shows the home page and the menu items for accessing the VLab facilities.

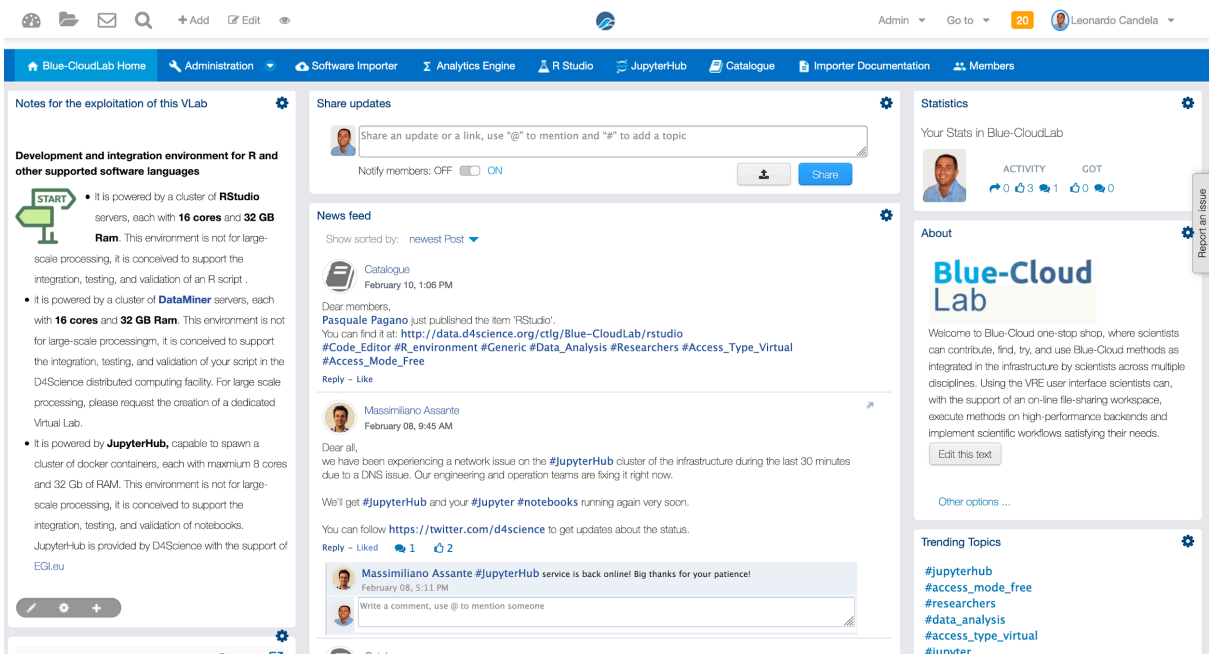

Figure 10. A screenshot of the Blue-Cloud Lab VLab

In addition to the basic functionality, e.g., a workspace for sharing objects of interest and a user management facility for managing membership, this VLab is specifically equipped with the following functionalities:

- a ***data analytics facility*** enabling users to benefit from the offerings of the Data Miner service (Assante et al. 2019a, Assante et al. 2019b) and interactively execute a large array of data analytics tasks on datasets. As of January '21 the facility offers ***39 ready to use method implementations*** ranging from Bayesian methods to maps comparison and data clustering;
- the ***RStudio*** facility enabling users to use a fully-fledged RStudio® working environment directly from the VRE. This environment is integrated with the rest of VLab facilities, e.g., it is possible to use files from the workspace and to store new files within the workspace;
- the ***JupyterHub*** facility enabling users to develop and execute Jupyter notebooks. This environment is integrated with the rest of VLab facilities, e.g., it is possible to use files from the workspace and to store new files within the workspace;

- the ***Catalogue*** facility enabling users to publish, search and browse datasets and other products of interest for the specific use case.

3.4 Blue-Cloud Project

The Blue-Cloud Project VLab was devised to support Blue-Cloud project activities and discussions. Only members of the Blue-Cloud consortium have access to it.

The Blue-Cloud Project VRE is available at https://blue-cloud.d4science.org/web/blue-cloudproject

This VLab has been in operational status since October '19 and it is currently serving 113 users, namely the Blue-Cloud Consortium members.

A screenshot of the VRE is provided in Figure 11. It shows the home page and the menu items for accessing the VLab facilities.

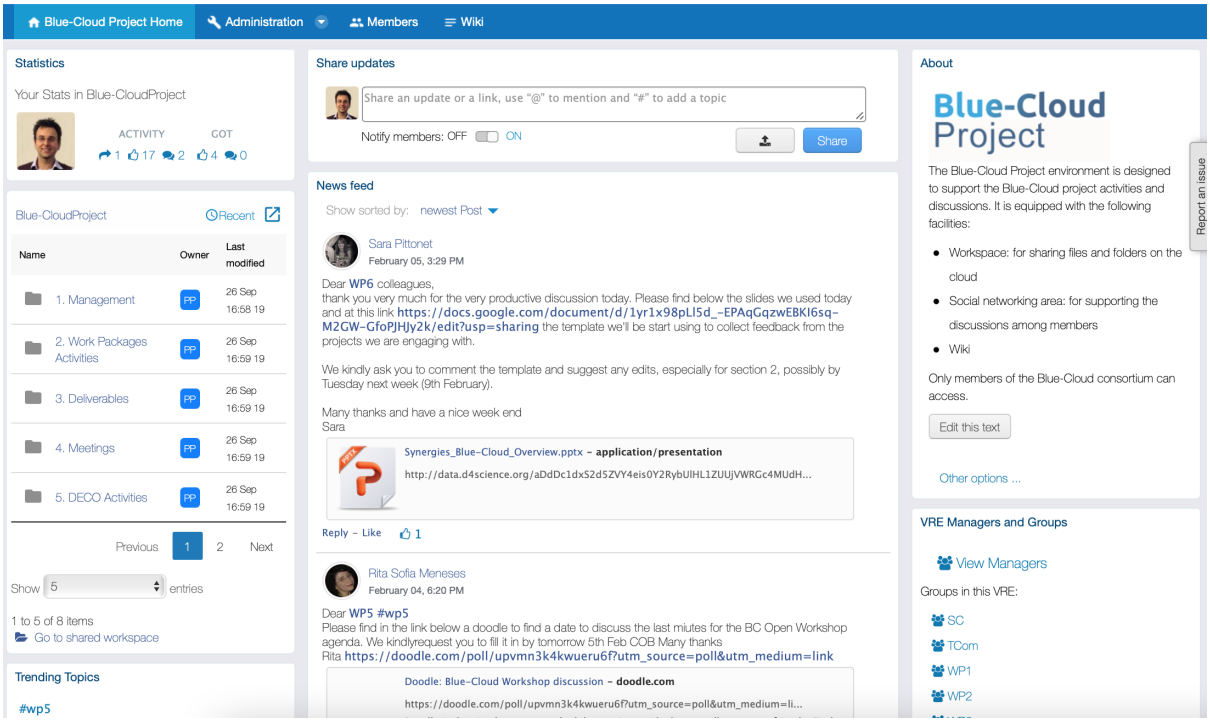

Figure 11. A screenshot of the blue-cloud Project VLab

In addition to the basic functionality, as social networking area for supporting the discussions among members and a user management facility for managing membership, this VLab is specifically equipped with the following capabilities:

- A ***Wiki*** facility enabling each VRE member to create wiki pages by using either the wiki or markdown syntax.
- the ***Catalogue*** facility enabling users to publish, search and browse products of interest for the specific community. In particular, this catalogue is planned to be populated with ***project deliverables*** and ***service descriptions*** characterising the service portfolio developed by the project.

3.5 Fisheries Atlas

The Fisheries Atlas VLab is developed in the context of the Demonstrator #4 - Fish, a matter of scales (Nys et al. 2020). This environment is conceived to provide its users with facilities supporting the development of an online overview of harmonized time-series of catch and effort.

This VLab is available at https://blue-cloud.d4science.org/web/fisheriesatlas

This VLab has been in operational status since December 2020. It is currently serving 15 users.

A screenshot of the VLab is available in Figure 12. It shows one specific app of this VRE (the map viewer) and the menu items for accessing the VLab facilities.

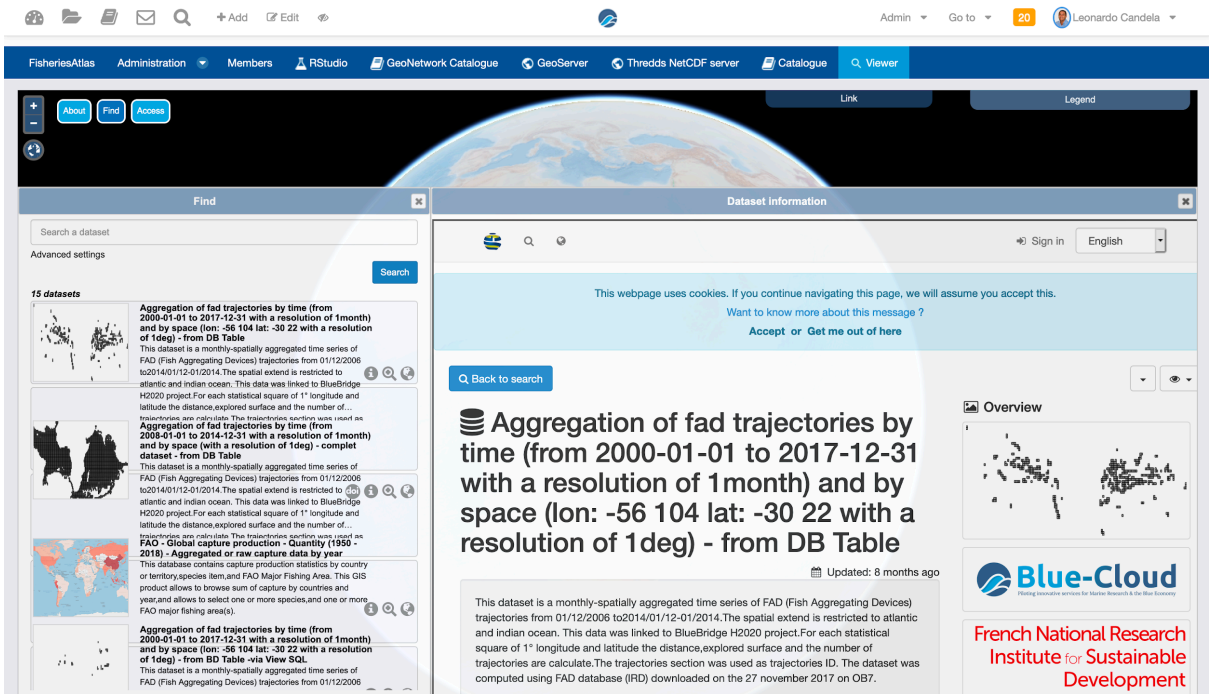

Figure 12. A screenshot of the Fisheries Atlas VLab

In addition to the basic functionality, e.g., a workspace for sharing objects of interest and a user management facility for managing membership, this VLab is specifically equipped with the following functionalities:

- the ***RStudio*** facility enabling users to use a fully-fledged RStudio® working environment directly from the VRE. This environment is integrated with the rest of VLab facilities, e.g., it is possible to use files from the workspace and to store new files within the workspace;
- a ***geospatial data catalogue*** facility enabling users to search and access ***more than 3500 datasets, services and maps*** of interest for the community;
- two ***repositories for geospatial datasets***, one based on GeoServer technology and another one based on THREDDS Data Server.
- a ***map viewer*** providing users with a seamless data discovery and visualization of information including FAO and IRD produced Tuna Atlas layers; FAO Productions and trade national statistics; Selected Regional Fisheries Organizations data; Selected EMODNet

layers; Selected CMEMS products; Selected species distribution maps; Selected FAO and VLIZ layers;

3.6 GRSF_Pre

The GRSF_Pre VLab is developed in the context of the Demonstrator #4 - Fish, a matter of scales (Nys et al. 2020). This environment is conceived to provide its users with a working environment to validate new content in the GRSF Knowledge Base (Tzitzikas et al. 2019) and thus it is not a public service. In 2020, the GRSF team used this environment to validate data harvests from 3 global GRSF sources that are now published into the GRSF Admin[4] and GRSF VREs[5].

This VLab is available at https://blue-cloud.d4science.org/web/grsf_pre

This VLab has been in operational status since June 2020. It is currently serving 12 users.

A screenshot of the VLab is available in Figure 13. It shows the catalogue facility displaying the available records.

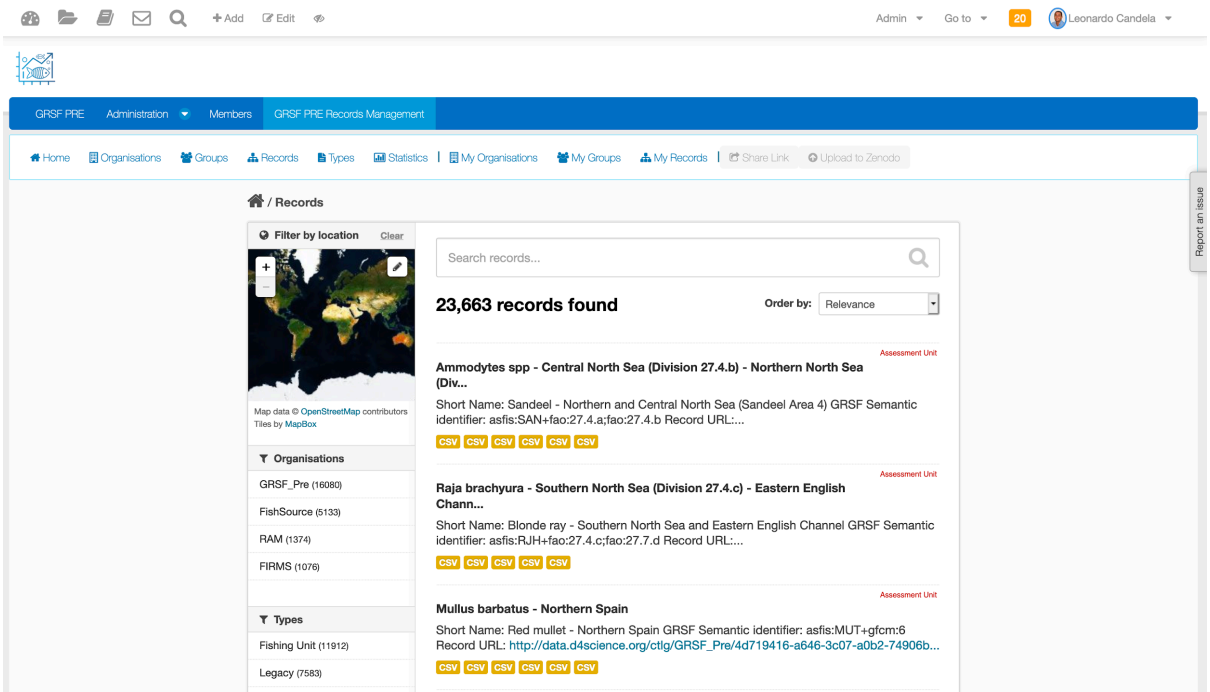

Figure 13. A screenshot of the GRSF_Pre VLab

[4] The GRSF Admin environment (https://i-marine.d4science.org/web/grsf_admin) provides authorized users with an environment and tools for building an integrated catalogue on stocks and fisheries information, i.e. for defining the authoritative version of the stocks and fisheries records made available by the GRSF official catalogue.

[5] The GRSF environment https://i-marine.d4science.org/web/grsf implements the "Global Record of Stocks and Fisheries" by providing users with an environment and tools for accessing stocks and fisheries information collated from three database sources: (i) Fisheries and Resources Monitoring System (FIRMS): http://firms.fao.org, (ii) RAM Legacy Stock Assessment Database: http://ramlegacy.org, and (iii) FishSource: http://www.fishsource.com

In addition to the basic functionality, e.g., a workspace for sharing objects of interest and a user management facility for managing membership, this VLab is specifically equipped with the following functionalities:

- a ***catalogue facility*** enabling users to visualise the GRSF candidate records stemming from the integration facility and manage their approval. This catalogue has been populated with ***more than 23,500 records*** including Fishing Unit, Assessment Unit, and Marine Resource;

3.7 HealthyOcean Lab

The HealthyOcean VLab was developed to support the liaison between the REV Ocean team[6] and the Blue-Cloud team regarding several initiatives including the Ocean Data Platform.

This VLab is available at https://blue-cloud.d4science.org/web/healthyocean-lab

This VLab has been in operational status since November 2019. It is currently serving 8 users.

A screenshot of the VLab is available in Figure 14. It shows the home page and the menu items for accessing the VLab facilities.

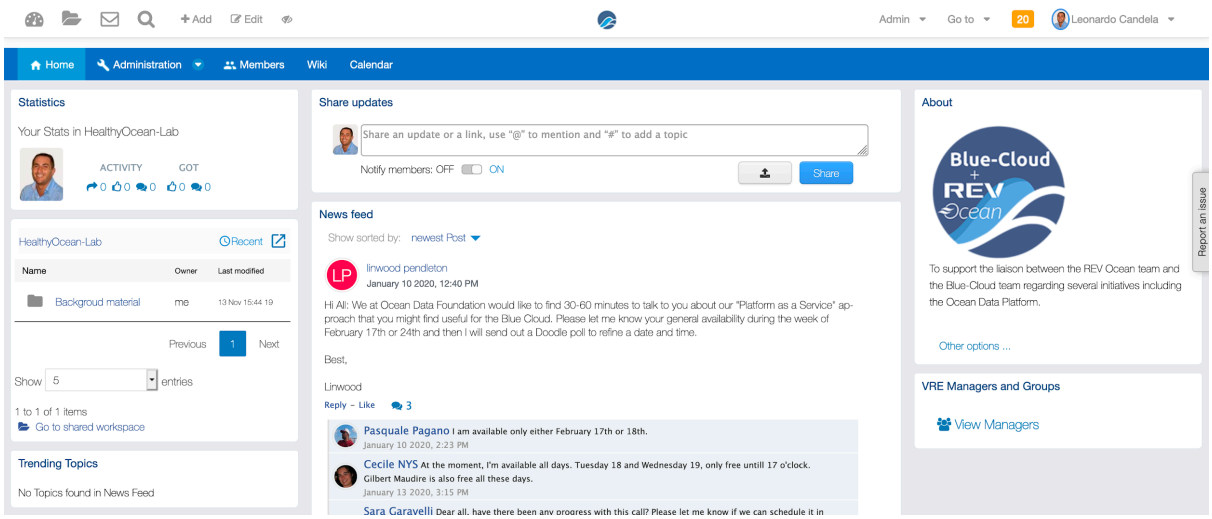

Figure 14. A screenshot of the HealthyOcean VLab

In addition to the basic functionality, e.g., a workspace for sharing objects of interest and a user management facility for managing membership, this VLab is not equipped with specific facilities since it was mainly conceived to support the cooperation between the two teams.

3.8 Marine Environmental Indicators

The Marine Environmental Indicators VLab is developed in the context of the Demonstrator #3 - Marine Environmental Indicators (Nys et al. 2020). It is conceived to provide its users with facilities to (i) identify environmental quality indicators in selected marine regions/areas; (ii) obtain new added-value data applying big data analysis and machine learning methods on the multi-source data

[6] REV Ocean website https://www.revocean.org/

sets, and (iii) to perform online and on the fly operations such as selecting portion of a dataset, to perform statistical analysis or display the data.

This VLab is available at https://blue-cloud.d4science.org/web/marineenvironmentalindicators

This VLab has been in operational status since April 2020. It is currently serving 36 users.

A screenshot of the VLab is available in Figure 15. It shows the MEI Generator app specifically developed for this community and the menu items for accessing the VLab facilities.

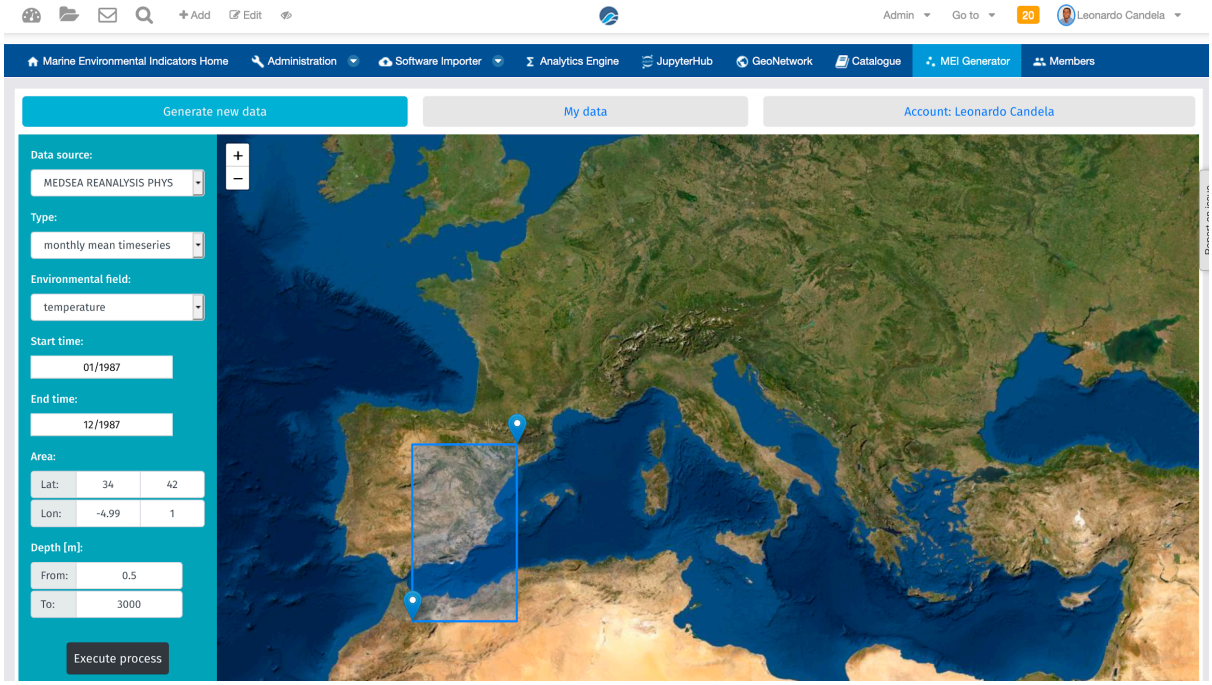

Figure 15. A screenshot of the Marine Environmental Indicators VLab

In addition to the basic functionality, e.g., a workspace for sharing objects of interest and a user management facility for managing membership, this VLab is specifically equipped with the following functionalities:

- a ***software importer facility*** enabling users to onboard existing software and algorithm implementations into the Data Miner platform (Assante et al. 2019a, Assante et al. 2019b);
- a ***data analytics facility*** enabling users to benefit from the offerings of the Data Miner service (Assante et al. 2019a, Assante et al. 2019b) and interactively execute a large array of data analytics tasks on datasets. As of January '21 the facility offers ***19 ready to use method implementations*** including preliminary versions of algorithms for computing the Storm Severity Index (dataset providing users insights on atmospheric wind/storm circumstances that impact the circulation of seas such as the Mediterranean Sea) and the seastat algorithm (i.e., the analytics process behind the MEI Generator app described below);
- the ***JupyterHub*** facility enabling users to develop and execute Jupyter notebooks. This environment is integrated with the rest of VLab facilities, e.g., it is possible to use files from the workspace and to store new files within the workspace. Two specific notebooks have been developed and shared: a ***Model Development notebook*** to design, optimize and train

a model; and a ***Prediction notebook*** to use a trained model and classify the profiles of a dataset into the different classes.

- the ***geospatial data catalogue*** facility enabling users to search and access geospatial products of interest for the community. At the time of writing this deliverable the catalogue has not been populated yet with products of interest;
- the ***Catalogue*** facility enabling users to publish, search and browse datasets and other products of interest for the specific use case;
- the ***MEI Generator app*** facility, i.e., a specific web application aiming at providing its users with a user interface facilitating the steps of generating new data starting from the existing ones and access and visualize the data previously calculated by the user.

One important feature of this VLab is that the environment was configured to serve two main classes of users: "VLab developers" and "VLab users". The former class of users has full access to all the facilities discussed above. The latter class of users has access only to the JupyterHub and MEI Generator app.

3.9 Zoo and Phytoplankton EOV

The Zoo and Phytoplankton EOV VLab is developed in the context of the Demonstrator #1 - Zoo- and Phytoplankton EOV products (Nys et al. 2020). This environment is conceived to support a methodology to generate: (i) zooplankton products based on in situ observations of abundance of different zooplankton species in a region encompassing the North-East Atlantic; (ii) global ocean three-dimensional (3D) products of chlorophyll-a (Chla) concentration, that is a proxy for total phytoplankton biomass, based on Argo vertical profiles matched up with satellite imagery; (iii) a mechanistic model using near real-time data to quantify the relative contributions of the bottom-up and top-down drivers in phytoplankton dynamics.

This VLab is available at https://blue-cloud.d4science.org/web/zoo-phytoplankton_eov

This VLab has been in operational status since July 2020. It is currently serving 12 users.

A screenshot of the VLab is available in Figure 16. It shows the Data Miner GUI and the menu items for accessing the VLab facilities.

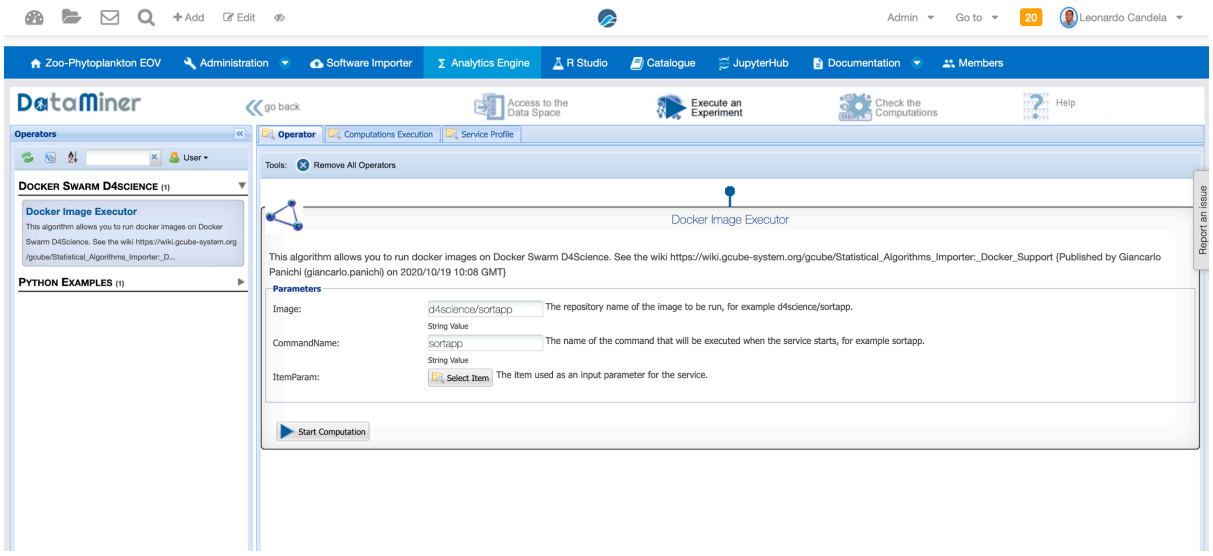

Figure 16. A screenshot of the Zoo and Phytoplankton EOV VLab

In addition to the basic functionality, e.g., a workspace for sharing objects of interest and a user management facility for managing membership, this VLab is specifically equipped with the following functionalities:

- a ***software importer facility*** enabling users to onboard existing software and algorithm implementations into the Data Miner platform (Assante et al. 2019a, Assante et al. 2019b);
- a ***data analytics facility*** enabling users to benefit from the offerings of the Data Miner service (Assante et al. 2019a, Assante et al. 2019b) and interactively execute a large array of data analytics tasks on datasets. As of January '21, the facility offers two ready-to-use method implementations. In particular, the ***Docker Image Executor*** algorithm that was specifically conceived to enable users to execute any algorithm packaged as a docker image. This algorithm is exploited to support the execution of the ***DIVAnd algorithm*** (Data-Interpolating Variational Analysis in n dimensions)[7];
- the ***RStudio*** facility enabling users to use a fully-fledged RStudio® working environment directly from the VRE. This environment is integrated with the rest of VLab facilities, e.g., it is possible to use files from the workspace and to store new files within the workspace;
- the ***Catalogue*** facility enabling users to publish, search and browse datasets and other products of interest for the specific use case;
- the ***JupyterHub*** facility enabling users to develop and execute Jupyter notebooks. This environment is integrated with the rest of VLab facilities, e.g., it is possible to use files from the workspace and to store new files within the workspace. Specific notebooks have been developed and shared: a notebook to create monthly fields of the vertical distribution of SOCA-Chl and a notebook to produce plots.

[7] https://github.com/gher-ulg/DIVAnd.jl

4. Concluding Remarks

This deliverable presented the Blue Cloud Virtual Research Environment by complementing the picture previously described in Assante et al. (2020b) where the constituents have been discussed. In particular, this deliverable described how the constituents have been deployed, exploited, and operated to support the development of the Blue Cloud gateway https://blue-cloud.d4science.org, its underlying infrastructure, and the VLabs.

A total of ***9 Blue-Cloud VLabs*** were created and operated. Five of these VLabs are conceived to support the developments of the Blue-Cloud Demonstrators, while four are conceived to support project activities and/or to provide their users with development and demonstrative environments.

These working environments are serving more than ***490 users*** spread across more than 20 countries. Up to January 2021, the served users performed a total of more than ***10,200 working sessions*** with an average of ***640 working sessions per month***. Overall, a total of ***1573 analytics tasks*** have been executed, with an average of 98 tasks per month. These figures are expected to grow in the second period of the project, thanks to the development and availability of VLabs stemming from the demonstrators.

References

M. Assante, L. Candela, D. Castelli, R. Cirillo, G. Coro, L. Frosini, L. Lelii, F. Mangiacrapa, V. Marioli, P. Pagano, G. Panichi, C. Perciante, F. Sinibaldi (2019a) ***The gCube system: Delivering Virtual Research Environments as-a-Service***. Future Gener. Comput. Syst. 95: 445-453 10.1016/j.future.2018.10.035

M. Assante, L. Candela, D. Castelli, R. Cirillo, G. Coro, L. Frosini, L. Lelii, F. Mangiacrapa, P. Pagano, G. Panichi, F. Sinibaldi (2019b) ***Enacting open science by D4Science***. Future Gener. Comput. Syst. 101: 555-563 10.1016/j.future.2019.05.063

M. Assante, L. Candela , D. Castelli, R. Cirillo, G. Coro, A. Dell'Amico, L. Frosini, L. Lelii, C. Magliozzi, F. Mangiacrapa, P. Pagano, G. Panichi, T. Piccioli, F. Sinibaldi (2020a) ***Co-creation of Virtual Research Environments: the D4Science Experience***. In Proceedings of 12th International Workshop on Science Gateways (IWSG 2020), 10-12 June 2020

M. Assante, L. Candela, P. Pagano, Andrea Dell'Amico, Luca Frosini, Lucio Lelii, Marco Lettere, Francesco Mangiacrapa, Giancarlo Panichi, Fabio Sinibaldi (2020b) ***Blue Cloud VREs Common Facilities (Release 1)***. D4.2 Blue-Cloud Deliverable, November 2020

G. Maudire, C. Nys, E. Blondel, G. Cochrane, M. Drudi, A. Ellenbroek, G. Everaert, A. Gentile, R. Lecci, S. Pesant, L. Schepers, L. Tyberghein (2020) ***Demonstrator general technical requirements***. D3.1 Blue-Cloud Deliverable, February 2020

C. Nys, G. Maudire, J. Augot, J. Barde, E. Blondel, P. Cabrera, G. Cochrane, P. Debeljak, M. Drudi, A. Ellenbroek, A. Gentile, E. Lavergne, R. Lecci, Y. Marketakis, S. Pesant, L. Schepers, L. Tyberghein (2020) ***Blue Cloud Demonstrator Users Handbook V1***. D3.3 Blue-Cloud Deliverable, December 2020

D. M. A. Schaap, P. Thijsse, P. Pagano, M. Assante, E. Boldrini, M. Buurman, M. D'Antonio, C. Ariyo, G. Maudire, C. Nys (2020) ***Blue Cloud Architecture (Release 1)***. D2.6 Blue-Cloud Deliverable, July 2020

Tzitzikas Y. et al. (2019) ***Methods and Tools for Supporting the Integration of Stocks and Fisheries***. In: Salampasis M., Bournaris T. (eds) Information and Communication Technologies in Modern Agricultural Development. HAICTA 2017. Communications in Computer and Information Science, vol 953. Springer, Cham. https://doi.org/10.1007/978-3-030-12998-9_2