



# Blue-Cloud

Piloting innovative services for Marine Research & the Blue Economy

## ***D4.3: Interfacing EOSC Report (Release 1)***

<b>Work Package</b>	WP4 - Developing and operating the Blue Cloud VRE, its services and Virtual Labs
<b>Lead Partner</b>	DKRZ (EUDAT)
<b>Lead Author (Org)</b>	Merret Buurman (DKRZ), Pasquale Pagano (CNR), Heinrich Widmann (DKRZ)
<b>Contributing Author(s)</b>	Leonardo Candela (CNR)
<b>Reviewers</b>	Marco Lettere (NUBISware & CNR) Gilbert Maudire (IFREMER) Dick M.A. Schaap (MARIS)
<b>Due Date</b>	28.02.2021, M17
<b>Submission Date</b>	24.03.2021
<b>Version</b>	1.0

### Dissemination Level

- 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)



## 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



This work by Parties of the Blue-Cloud Consortium is licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). “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.

## VERSIONING AND CONTRIBUTION HISTORY

Version	Date	Authors	Notes
0.1	15.12.2020	Heinrich Widmann (DKRZ)	Initial draft.
0.2	25.02.2021	Merret Buurman (DKRZ), Leonardo Candela (CNR), Pasquale Pagano (CNR), Heinrich Widmann (DKRZ)	Finished first draft, ready for internal review.
0.3	26.02.2021	Marco Lettere (NUBISware & CNR)	Added some comments (internal review).
0.4	26.02.2021	Gilbert Maudire (Ifremer)	Added some comments (internal review).
0.5	02.03.2021	Merret Buurman (DKRZ)	Incorporated comments by reviewers. Checked and corrected figure numbers.
0.6	03.03.2021	Pasquale Pagano (CNR)	Reviewed comments by reviewers. Applied changes when requested.
0.7	03.03.2021	Merret Buurman (DKRZ)	Approved some changes, completed/revised parts of service list.
0.8	03.03.2021	Heinrich Widmann (DKRZ), Merret Buurman (DKRZ), Sara Pittonet (Trust-IT), Federico Drago (Trust-IT)	Final proof-reading and final version for upload
0.9	16.03.2021	Dick M.A. Schaap (MARIS)	Final review as Technical Coordinator
1.0	24.03.2021	Sara Pittonet (Trust-IT)	General formatting and submission to the EC Participant's Portal

## Contents

List of abbreviations and acronyms .....	5
1. Introduction .....	6
2. The Blue-Cloud Catalogue .....	7
○ 2.1. Description .....	7
○ 2.2. The Blue-Cloud onboarding procedure .....	10
○ 2.3. The Blue-Cloud Catalogue organisational model .....	12
3. The EOSC Catalogue and onboarding procedure .....	13
○ 3.1. Description .....	13
○ 3.2. Definition of EOSC Providers and EOSC Resources .....	13
○ 3.3. EOSC Profiles .....	13
○ 3.4. Preconditions for onboarding .....	14
○ 3.5. The onboarding procedure .....	14
■ 3.5.1. Initial onboarding of resource providers and services .....	15
■ 3.5.2. Onboarding/updating via EOSC Catalogue HTTP API .....	15
■ 3.5.3. Alternative: Bulk-onboarding the entire catalogue .....	16
4. Onboarding Blue-Cloud services into the EOSC Catalogue .....	17
○ 4.1. Integration of the Blue-Cloud Providers .....	17
○ 4.2. Integration of the services .....	20
○ 4.3. Services identified by the demonstrators .....	21
5. Conclusions and next steps .....	24
6. References.....	25

## Executive summary

The Blue-Cloud Service platform will feature a variety of services that can be used for undertaking world-class science via the **European Open Science Cloud (EOSC)** framework, by featuring leading operational marine research infrastructures and e-infrastructures. For that purpose, Blue-Cloud builds a service catalogue which will be integrated in the EOSC Portal. The objective is to identify the services enabled in the Blue-Cloud Virtual Research Environment (VRE), document these in a Blue-Cloud catalogue which then will be integrated in the EOSC service catalogue and made easily accessible to all EOSC users via the EOSC Portal Catalogue and Marketplace<sup>1</sup>.

This deliverable illustrates the thematic catalogue of blue services in the Blue-Cloud Virtual Research Environment (VRE) and reports on the technical implementation done to enable onboarding of resources of the Blue-Cloud catalogue into the EOSC Portal. It includes details on the activities performed to ensure that the configuration, formats and metadata of resources are compatible with the requirements of the EOSC Portal. This enables and simplifies the process of onboarding Blue-Cloud resources into the EOSC resource catalogue.

While interfacing with the EOSC Catalogue will be described in detail and in greater depth in the second release of the deliverable, this document provides an overview of the different procedures for onboarding 'providers' and 'services' from the Blue-Cloud catalogue into the EOSC Portal Catalogue.

---

<sup>1</sup> <https://marketplace.eosc-portal.eu/>

## List of abbreviations and acronyms

<b>Abbreviation</b>	<b>Signification</b>
API	Application programming interface
CMCC	Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici, Italy (Blue-Cloud partner leading the Marine Environmental Indicator demonstrator)
CNR	National Research Council of Italy, Blue-Cloud partner co-leading this deliverable.
DIVA	Data-Interpolating Variational Analysis, a spatial interpolation software developed and maintained by the Geohydrodynamics working group at the University of Liège in Belgium.
DKRZ	Deutsches Klimarechenzentrum GmbH / German Climate Computing Centre, Blue-Cloud partner leading this deliverable.
EMBL	European Molecular Biology Laboratory (Blue-Cloud partner leading the Plankton Genomics demonstrator)
EOSC	European Open Science Cloud
EOV	Essential Ocean Variables
FAO	Food and Agriculture Organization of the United Nations, Italy (Blue-Cloud partner leading the Fish and the Aquaculture Monitor demonstrator)
GIS	Geographical Information System
GUI	Graphical User Interface
HTTP API	API that uses Hypertext Transfer Protocol
OGC	Open Geospatial Consortium, an organisation for standards in the geospatial sciences.
REST API	HTTP API that follows specific concepts called Representational State Transfer
VLIZ	Flanders Marine Institute, Belgium (Blue-Cloud partner leading the Zoo- and Phytoplankton EOVS products demonstrator)
VRE	Virtual Research Environment
WFS	Web Feature Service, an OGC standard for accessing geographical vector data online.
WMS	Web Map Service, an OGC standard for accessing maps online.

# 1. Introduction

Blue-Cloud is a large pan-European project federating thematic data infrastructures and services from marine research institutions across Europe. This is achieved thanks to the Blue-Cloud platform, the architecture of which (Schaap et al. 2020) 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. The resulting Blue-Cloud VRE is made available via the Blue-Cloud gateway at <https://blue-cloud.d4science.org>. This gateway provides access to the set of services and Virtual Labs developed by the project (Assante et al, 2021).

This deliverable describes the thematic catalogue of services in the Blue-Cloud Virtual Research Environment (VRE) and reports on the activities carried out in order to interface the services with the European Open Science Cloud (EOSC), which can be done in practice by making the Blue-Cloud services findable and accessible via the EOSC Portal Catalogue & Marketplace. To achieve this, the Blue-Cloud services have to be “onboarded” into the EOSC Portal Catalogue, while the actors serving as service providers have to be registered into the EOSC Portal according to the EOSC guidelines<sup>2</sup>. Previous to this, providers and services have to be identified, described and onboarded into the Blue-Cloud Thematic Service Catalogue. To onboard the providers and services, they have to be described by comprehensive metadata in an EOSC-compatible format.

Section 2 describes the Blue-Cloud Service Catalogue and its onboarding process. The EOSC Service Catalogue with its preconditions is described in Section 3, while its onboarding process is described in detail in Section 4.

The task of onboarding Blue-Cloud VRE services into EOSC takes place in two phases, namely from 11/2020 to 04/2021 (resulting in this Deliverable D4.3), and from 11/2021 to 04/2022 (resulting in the D4.5 Interfacing EOSC Report - Release 2) . Therefore, in this deliverable we focus mainly on the identification of services and on their onboarding into the Blue-Cloud catalogue (as phase 1 of the task is still ongoing, the service descriptions are not finished yet and will be included in the next deliverable). The handing over of these service catalogue entries into the EOSC Portal Catalogue & Marketplace is outlined. It will be carried out and described in the second phase and reported in deliverable D4.5 due in 04/2022.

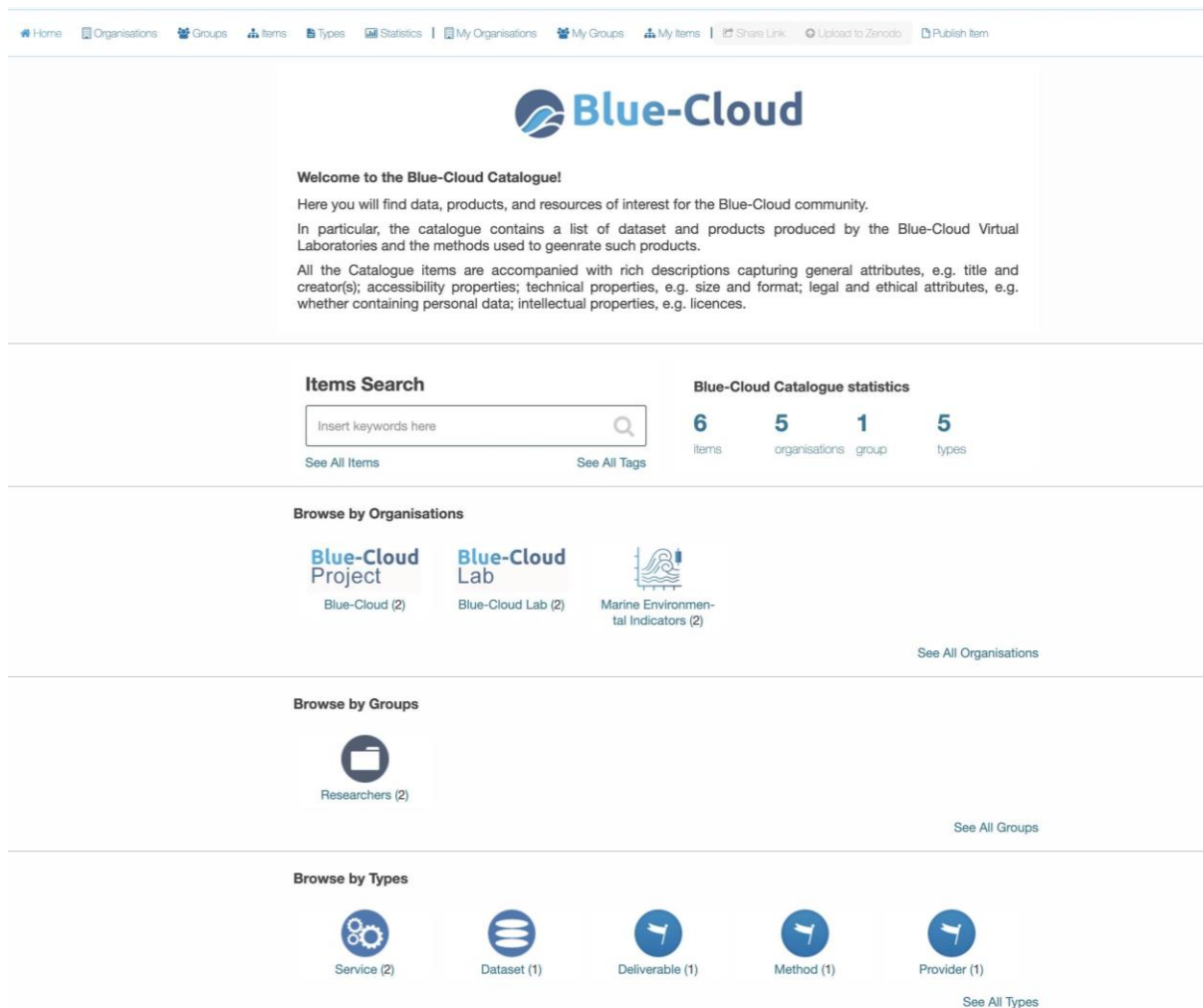
---

<sup>2</sup> <https://providers.eosc-portal.eu/becomeAProvider>

## 2. The Blue-Cloud Catalogue

### o 2.1. Description

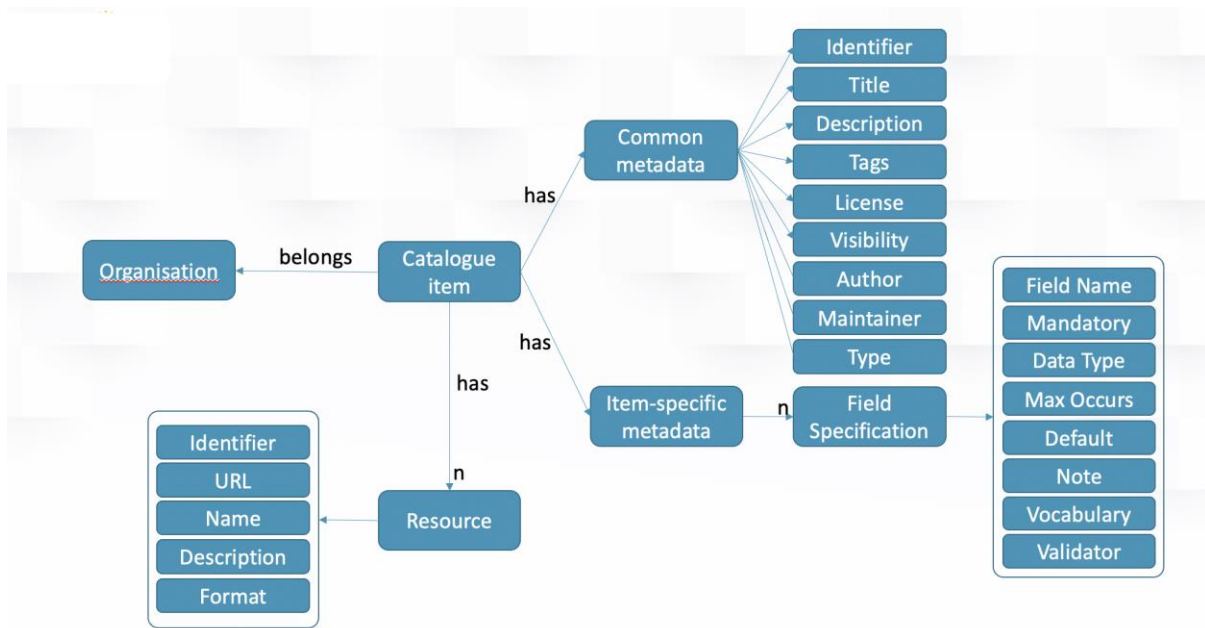
The Blue-Cloud Service Catalogue<sup>3</sup> lists all the resources and analytic services developed and provided in the scope of the Blue-Cloud project. A screenshot is available in Fig. 1.



*Fig. 1: Screenshot of the Blue-Cloud service catalogue.*

This catalogue has been developed by instantiating the catalogue service technology (Assante et al, 2020). Apart from instantiating the technology, it is worth discussing how the data model is supported by the catalogue (Fig. 2) and has been exploited to serve the needs of the Blue-Cloud.

<sup>3</sup> <https://blue-cloud.d4science.org/group/bluecloud-gateway/data-catalogue>



*Fig. 2: Catalogue Data Model*

Catalogue items are published by “organisations”. An **organisation** represents the context/authority responsible for the publishing of the item. Organisations are usually paired with V Labs for the items stemming from them or with other existing contexts where the items pre-exist V Labs.

This catalogue serves all the virtual laboratories that are deployed in the Blue-Cloud VRE. In this initial phase, 5 organisations have been created:

- Blue-Cloud;
- Blue-Cloud Lab;
- Fisheries Atlas;
- Marine Environmental Indicators;
- Zoo-Phytoplankton EO.V.

In particular, the Blue-Cloud organisation is the organisation responsible for the publication of resources produced by the consortium for its operation, such as Deliverables and Training Materials; the Blue-Cloud Lab organisation is the organisation delivering general purpose services that are offered in open-access mode to researchers and scientists beyond the boundaries of the Blue-Cloud consortium. Later on, the Demonstrators will have their own environment to publish resources directly in the Blue-Cloud catalogue.

The Blue-Cloud catalogue supports five types of items (Provider, Dataset, Service, Deliverable and Method). In particular, Provider and Service are compliant with the profiles promoted by EOSC: the Provider catalogue item is compliant with the EOSC Provider profile<sup>4</sup> while Service catalogue item is compliant with the EOSC Resource profile<sup>5</sup>.

<sup>4</sup> <https://eosc-portal.eu/providers-documentation/eosc-provider-portal-provider-profile>

<sup>5</sup> <https://eosc-portal.eu/providers-documentation/eosc-provider-portal-resource-profile>



A *Service* describes an online technology that is used to transmit, process, store, create, display, share or exchange information. A *Service* is provisioned by a *Provider* that is the entity responsible for its operation. For example, the Blue-Cloud Lab is the *Organisation* providing access to the RStudio *Service* operated by the D4Science Infrastructure *Provider*.

A *Method* describes a computational script that requires a *Service* to be executed. A *Dataset* describes a collection of data either imported or generated via a *Service*. A *Deliverable* describes a report.

Every item of the catalogue is characterised by a set of **common metadata**, including a unique identifier, a title, a description, a list of tags (e.g., keywords, subjects), a licence, visibility (whether the item is publicly available or visible only to the members of a VLab), an author, a maintainer, and a type. Every Type is characterised by a specific set of attributes, controlled vocabularies and formats carefully describing the specific class of items.

*Service* and *Provider* typologies are inherited by EOSC; *Dataset* and *Method* are inherited by the SoBigData Infrastructure<sup>6</sup>; *Deliverable* has been defined by Blue-Cloud. The metadata of each of these types are summarised in the following tables.

Metadata item(s)	Information provided
Basic Information	ID, Name, Website
Marketing Information	Description, Logo
Classification Information	Scientific domains & subdomains, tags
Location Information	City, region, country
Contact Information	Main contact, public contact
Maturity Information	Lifecycle status, certifications
Other Information	ESFRI domain and type, MERIL classifications

Table. 1: **Provider Type** - 41 attributes (15 mandatories), 14 controlled vocabularies

Metadata item(s)	Information provided
Basic Information	ID, Name, Website, <b>Provider</b>
Marketing Information	Description, Tagline, Logo
Classification Information	Scientific domains & Subdomains, Categories
Localisation	Geographical & Language Availability; Resource Location
Contact Information	Main contact, public contact, helpdesk
Maturity Information	TRL, Lifecycle status, certifications
Dependencies & Attribution	Dependencies on other Services
Management Information	Helpdesk, ToU, access policy
Access & Order Information	Locator to access and order the service
Financial Information	Additional information about the costs of the service

Table. 2: **Service Type** - 61 attributes (22 mandatories), 15 controlled vocabularies

<sup>6</sup> SoBigData: European Research Infrastructure for Big Data and Social Mining, [www.sobigdata.eu](http://www.sobigdata.eu)

Metadata item(s)	Information provided
Basic Information	Support the identification of the resource
Coverage	Identify the extent of the resource
Access Modality	Report the modality of exploitation of the resource
Technical Details	Report detailed information about the resource
Rights Details	Report detailed information about the policies governing the exploitation of the resource
Attribution Details	Report detailed information about the text to use to acknowledge the usage of the resource resource
Data Protection Details	Report detailed information about the policy regulating the access and the usage of personal and sensitive data

*Table. 3: **Dataset Type** - 35 attributes (10 mandatories), 14 controlled vocabularies*

Metadata item(s)	Information provided
Basic Information	Support the identification of the resource
Coverage	Identify the extent of the resource
Access Modality	Report the modality of exploitation of the resource
Technical Details	Report detailed information about the resource
Rights Details	Report detailed information about the policies governing the exploitation of the resource
Attribution Details	Report detailed information about the text to use to acknowledge the usage of the resource resource

*Table. 4: **Method Type** - 22 attributes (5 mandatories), 6 controlled vocabularies*

Metadata item(s)	Information provided
Basic Information	Support the identification of the resource
Temporal Coverage	Date of publication

*Table. 5: **Deliverable Type** - 4 attributes (3 mandatories)*

## o 2.2. The Blue-Cloud onboarding procedure

This section describes in detail the procedure to register items into the Blue-Cloud catalogue.

Users having the ‘Catalogue Editor’ role are entitled to publish items into the catalogue. This role is assigned per VLab. For instance, one user could be authorised to publish items exclusively to the virtual laboratory “Fisheries Atlas”, whereas another user could be authorised to publish items to the two virtual laboratories “Blue-Cloud” and “Blue-Cloud Lab”.

Users having the ‘Catalogue Admin’ role are notified about new published items and they can review, edit and either approve or reject them. Also, this role is assigned per VLab ensuring complete autonomy in the management of the catalogue of each virtual environment.

The publication of the items can be performed via either the Workspace user interface, the Catalogue user interface, or the gCat REST service APIs<sup>7</sup>.

The first two modalities open a wizard organised in three sections.

The first section of the Wizard is common to all different types and allows the user to specify the common metadata (see Fig. 3).

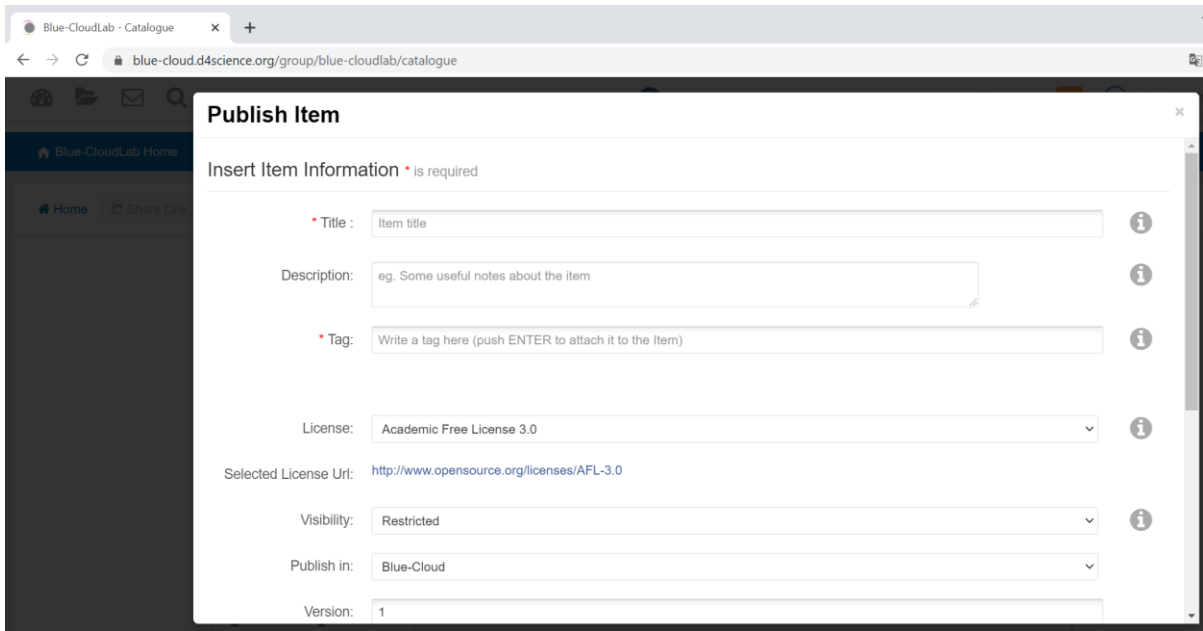


Fig. 3: Publish an item - Common metadata.

The second section is specific to the selected Type (see Fig. 4).

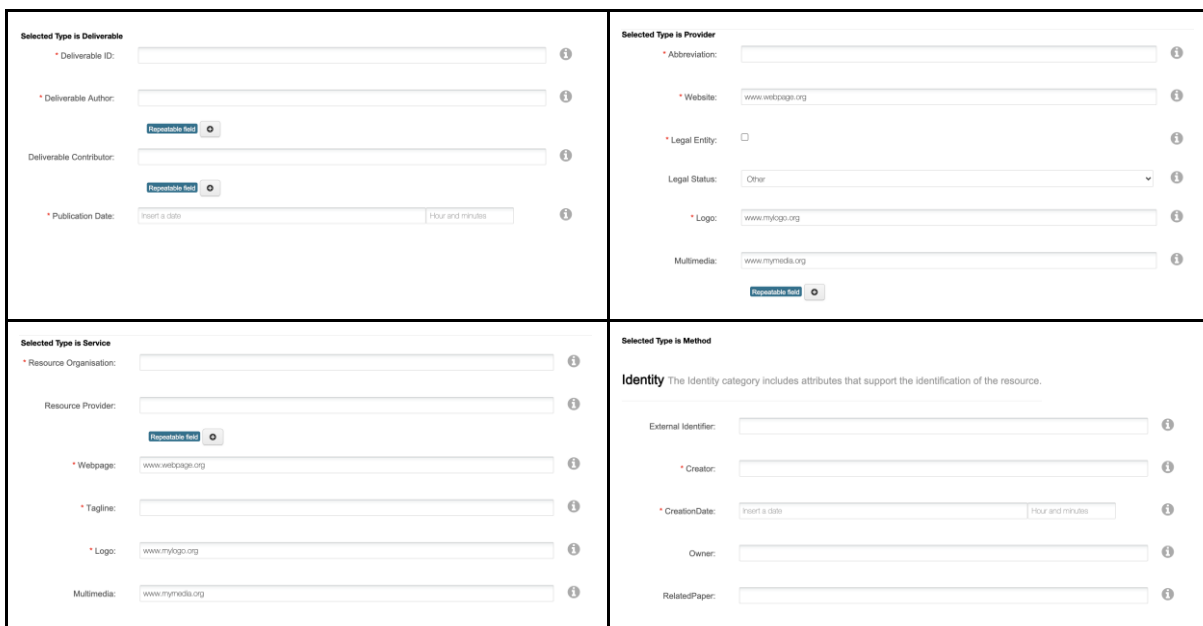


Fig. 4: Publish an item - Type specific metadata.

<sup>7</sup> gCat REST APIs - [https://wiki.gcube-system.org/gcube/GCat\\_Service](https://wiki.gcube-system.org/gcube/GCat_Service)

The third section of the Publication wizard allows the user to add one or more manifestations of the resource. A manifestation is the payload of the item and it is characterised by the identifier, a name, a description, a format, and, most importantly, the URL pointing to the content. The definitions of the metadata models are quite abstract, which implies that people can have different interpretations. The staging process with a final review and editing is needed in practice to guarantee a homogeneous approach and documentation.

### o 2.3. The Blue-Cloud Catalogue organisational model

The Blue-Cloud project federates and pilots innovative services for the marine research and the Blue economy communities. Its catalogue includes services and resources that pre-exist Blue-Cloud and are operated by third-parties e-infrastructures as well as new services and new resources that are generated by the Blue-Cloud project consortium. To model this dichotomy in the Blue-Cloud catalogue, different Providers will be registered. We started from the following two providers:

- the Blue-Cloud Provider is the entity that manages or delivers the services and resources generated by the project, or that coordinates the delivery of those services and resources in a federated scenario;
- the D4Science Provider is the entity that manages and delivers the enabling services and technology exploited by the virtual laboratories. For example, the services offered by the Blue-Cloud Lab - namely RStudio, JupyterNotebook, Analytics Engine and the Software Importer - pre-exist the Blue-Cloud project but their delivery is coordinated by the project.

As soon as the services of other providers, e.g. WEkEO, EUDAT, will become available through the Blue-Cloud VRE, we will register those as e-infrastructure Providers.

Similarly, we will register the data infrastructures federated by Blue-Cloud, e.g. EMODnet, EurOBIS, ICOS - Marine, etc, as soon as their datasets will become accessible through the Blue-Cloud Discovery & Access service.

Finally, each VLab will register the services implemented as project demonstrator by specifying Blue-Cloud as Resource Organisation and either D4Science, WEkEO, or the proper e-infrastructure as Resource Organisation.

## 3. The EOSC Catalogue and onboarding procedure

### o 3.1. Description

The EOSC Service Catalogue is a listing of services targeted towards open research. These resources can be browsed in the EOSC Portal Catalogue & Marketplace, an “*integrated platform that allows easy access to lots of resources for various research domains along with integrated data analytics tools*”<sup>8</sup>.

Although EOSC also aims at making research data accessible across Europe, datasets may not be found directly in the marketplace. However, the marketplace lists services that allow the discovery of research datasets. As this deliverable focuses on the interfacing of services and methods, the discovery of Blue Cloud data is covered in a different task (2.3) and deliverable.

### o 3.2. Definition of EOSC Providers and EOSC Resources

Central concepts in this and the next sections are those of an EOSC Provider and an EOSC Resource. These concepts are defined as follows:

*An EOSC Provider<sup>9</sup> is an EOSC System User responsible for the provisioning of one or more Resources to the EOSC. EOSC Providers are organisations, a part of an organisation or a federation that manages and delivers Resources to End-Users. EOSC Providers can be: Resource Providers, Service Providers, Data (Source) Providers, Service Developers, Research Infrastructures, Distributed Research Infrastructures, Resource Aggregators, Thematic Clouds, Regional Clouds, etc.*

*An EOSC Resource<sup>10</sup> is an asset made available by means of the EOSC system and according to the EOSC Rules of Participation to EOSC End-Users to perform a process useful to deliver value in the context of the EOSC. EOSC Resources include Services, Data Sources and any other asset. A Resource Profile describes the information requested to onboard Resources into the EOSC Provider Portal.*

### o 3.3. EOSC Profiles

The services listed in the EOSC Service Catalogue as well as their providers are described by structured and well-defined metadata. The metadata items are listed and defined in the EOSC Profiles<sup>11</sup>. Their documentation provides examples and guidance on every metadata item.

---

<sup>8</sup> <https://marketplace.eosc-portal.eu/>

<sup>9</sup> <https://eosc-portal.eu/providers-documentation/eosc-provider-portal-provider-profile>

<sup>10</sup> <https://eosc-portal.eu/providers-documentation/eosc-provider-portal-resource-profile>

<sup>11</sup> <https://eosc-portal.eu/sites/default/files/EOSC-Profiles-v3.00.pdf>

For describing a service provider, the EOSC Provider Data Model must be used:

<https://eosc-portal.eu/providers-documentation/eosc-provider-portal-provider-profile>

It lists the mandatory and optional metadata, as well as controlled vocabularies (code lists, taxonomies, classifications) for items such as legal statuses, areas of activity, countries, scientific domains, consortia, etc. Currently, 15 items are listed as mandatory and 36 are optional.

For describing a service, the EOSC Resource Data model must be used:

<https://eosc-portal.eu/providers-documentation/eosc-provider-portal-resource-profile>

Currently, 22 items are listed as mandatory and 39 are listed as optional.

### o 3.4. Preconditions for onboarding

There are a number of preconditions that services must comply with, before they can be part of the EOSC. First of all, as Open Science is the primary goal of the EOSC, all onboarded services have to enable and enhance and comply with the principle of Open Science.

EOSC defined a set of Rules of Participation that the services and their providers have to comply with<sup>12,13</sup>. These rules mainly address high-level general concepts such as openness, FAIR-ness, policies and legislation. The services also have to be technically mature and be classified at least Technology Readiness Level 7 (as defined by the European Commission<sup>14</sup>), which means that a prototype demonstrator of the system must be running in an operational environment. Also, the services have to be properly and correctly described with service metadata. More details on the preconditions can be found in the EOSC-Hub Integration Handbook (Sipos et al. 2020)<sup>15</sup>, on the EOSC documentation for providers<sup>16</sup> and also in the EOSC wiki<sup>17</sup>.

### o 3.5. The onboarding procedure

There are various ways of onboarding providers and services into the EOSC Catalogue and for updating the information once it is integrated: Using a form/GUI, using an HTTP API, or by bulk-onboarding an existing compatible catalogue.

---

<sup>12</sup> <https://www.eosc-hub.eu/key-exploitable-results/eosc-rules-participation-rop>

<sup>13</sup> <https://op.europa.eu/en/publication-detail/-/publication/a96d6233-554e-11eb-b59f-01aa75ed71a1/language-en/format-PDF/source-184432576>

<sup>14</sup> [https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014\\_2015/annexes/h2020-wp1415-annex-g-tr1\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-tr1_en.pdf)

<sup>15</sup> <https://www.eosc-portal.eu/sites/default/files/EOSC-hub%20Integration%20Handbook%20for%20Service%20Providers.pdf>

<sup>16</sup> <https://eosc-portal.eu/providers-documentation/eosc-provider-portal-inclusion-criteria>

<sup>17</sup> <https://wiki.eosc-hub.eu/display/EOSC/Criteria+for+possible+inclusion+in+the+EOSC+Service+Portfolio>

### ■ 3.5.1. Initial onboarding of resource providers and services

At first, a provider has to apply for becoming an EOSC Provider. This process is documented in detail at <https://eosc-portal.eu/for-providers> (see EOSC-Hub Integration Handbook, Sipos et al. 2020)<sup>18</sup>. A new resource provider can be registered using the form at <https://providers.eosc-portal.eu/provider/add>. Once this information is reviewed by EOSC, detailed *Service Description Templates* are created and have to be filled and handed in by the Blue-Cloud service providers. The information contained in these Service Descriptions are used by EOSC to determine whether the provider's services offered fulfil the preconditions, such as the Rules of Participation (RoP)<sup>19</sup>. It is also used to populate the EOSC Service Portfolio. After review and approval by both sides - EOSC and the service providers - the descriptions are published and are now part of the public EOSC Service Catalogue.

### ■ 3.5.2. Onboarding/updating via EOSC Catalogue HTTP API

EOSC provides an open HTTP API to manage the provider and service descriptions in the Service Catalogue. The API is documented here: <https://providers.eosc-portal.eu/openapi>. In this deliverable, we refer to the version 3.0.0 of the API (as of 2021-02-22).

The API provides endpoints to update provider information, and to create and update service information. In the following subsections, we will describe the various actions that can be taken.

The providers still have to be onboarded beforehand, as their integration has to be reviewed and approved by the EOSC onboarding team. However, integrating the individual services and keeping their descriptions up-to-date in the EOSC catalogue becomes easier and more automatic with the HTTP API. An automatic tool could be deployed that feeds changes in the metadata automatically into the EOSC catalogue in near-real time. Obviously, the metadata about the services still has to be created by the service providers and it has to be stored someplace - ideally the provider's own service catalogue - for such a tool to pick it up and push it into EOSC catalogue.

For any of the operations described below, the client performing them has to be authorized. The authorization mechanism adopted in the EOSC Catalogue API is the API key mechanism<sup>20</sup>, where the client provides a token in the request header of an HTTP request. This token is supplied by the operator of the API, i.e. by EOSC.

- **Updating a provider**

The HTTP endpoint `/provider` allows the client to update the information about the provider, by uploading a JSON document containing the updated information via HTTP PUT. The JSON document must contain the updated metadata in a specific format. For examples and documentation, please refer to <https://providers.eosc-portal.eu/openapi>. The types and values are described in the "Provider Data Model", presented in section on

---

<sup>18</sup> <https://www.eosc-portal.eu/sites/default/files/EOSC-hub%20Integration%20Handbook%20for%20Service%20Providers.pdf>

<sup>19</sup> <https://op.europa.eu/en/publication-detail/-/publication/a96d6233-554e-11eb-b59f-01aa75ed71a1/language-en/format-PDF/source-184432576>

<sup>20</sup> <https://swagger.io/docs/specification/2-0/authentication/api-keys/>

EOSC Profile (section 3.3). The initial integration of a provider is not possible using this API and has to be done as described in section 3.5.1.

- **Creating a resource**

The HTTP endpoint `/resource` allows the providers to create a new resource, e.g. a service, by uploading a JSON document containing the new information via `HTTP POST`. The JSON document must contain the metadata in a specific format. For examples and documentation, please refer to <https://providers.eosc-portal.eu/openapi>. The types and values are described in the “Resource Data Model”, presented in section on EOSC Profile (section 3.3).

- **Updating or validating a resource**

The HTTP endpoint `/resource` allows the providers to update an existing resource/service based on its id, by uploading a JSON document containing the updated information via `HTTP PUT`. It is also possible to validate an existing resource (without changing it) via the endpoint `/resource/validate`, using `HTTP POST`.

- **3.5.3. Alternative: Bulk-onboarding the entire catalogue**

There are plans to enable the onboarding of entire Service Catalogues if their resources are compatible with the EOSC Service Catalogue:

*"The EOSC Portal team (...) is implementing a method to import and share whole service catalogues from in one-go [sic!], simplifying the service onboarding for provider consortia."*

(EOSC-Hub Integration handbook for service providers, Sipos et al. 2020).

The resource provider still has to be onboarded first, as its integration has to be reviewed and approved by the EOSC onboarding team. However, the integration of the individual services becomes easier and more automatic. Of course, the services have to be fully described in the original catalogue for the descriptions to be imported into the EOSC catalogue.



## 4. Onboarding Blue-Cloud services into the EOSC Catalogue

After having described the EOSC onboarding procedure in detail, this section will explain how it will be used to integrate the Blue-Cloud services into the EOSC catalogue.

The creation of the EOSC-compatible Blue-Cloud catalogue was an important basis for the EOSC integration. Based on it, it will be easy to transfer the created resources into the EOSC catalogue. This avoids having to (re-)create metadata twice, separately for both catalogues.

First, the service providers have to be registered in the EOSC catalogue. Once the providers are integrated, the next steps to be performed in this task will focus on the integration of the services.

### o 4.1. Integration of the Blue-Cloud Providers

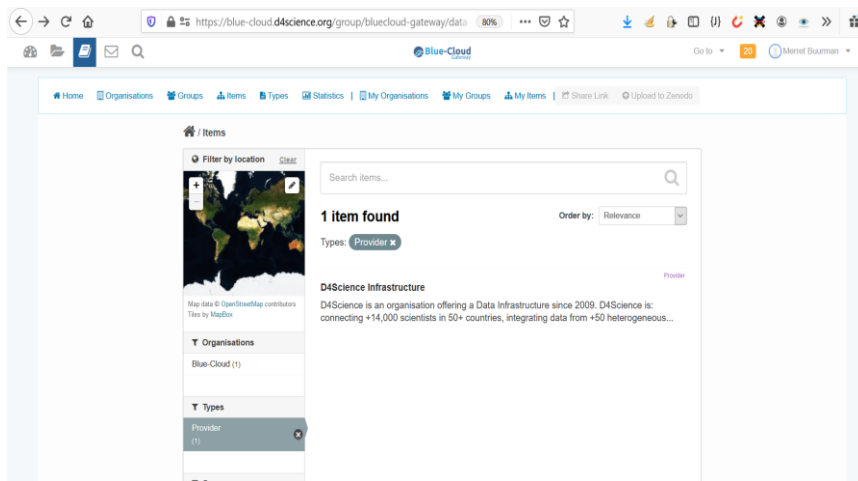
As mentioned in Section 2.3, the services generated in the Blue-Cloud project will be provided by the Provider *Blue-Cloud*, while *D4Science* is the provider that manages and delivers the enabling services and technology exploited by the virtual laboratories.

D4Science is already present as a Provider in the Blue-Cloud catalogue as you can see in Figures 5 and 6. As you can see in Fig. 7, D4Science is also already integrated into the EOSC catalogue as a provider for various non-Blue-Cloud services<sup>21</sup>, so there is no need to integrate D4Science again. Blue-Cloud itself is not yet integrated as a Provider in EOSC. So, the first step will be to create and integrate the Blue-Cloud service provider, which will be the main service provider in the Blue-Cloud project besides D4Science.

Services of other providers, e.g. WEkEO, EUDAT, may become available through the Blue-Cloud VRE and in that case will be integrated into the Blue-Cloud service catalogue. Some of them are already integrated into EOSC previously (e.g. EUDAT services). We encourage the integration of the remaining Providers and their services into EOSC as well.

---

<sup>21</sup> [https://marketplace.eosc-portal.eu/services?geographical\\_availabilities=&order\\_type=&providers%5B%5D=30&providers-all=&providers-filter=&rating=&related\\_platforms-all=&related\\_platforms-filter=&scientific\\_domains-all=&scientific\\_domains-filter=&target\\_users-all=](https://marketplace.eosc-portal.eu/services?geographical_availabilities=&order_type=&providers%5B%5D=30&providers-all=&providers-filter=&rating=&related_platforms-all=&related_platforms-filter=&scientific_domains-all=&scientific_domains-filter=&target_users-all=)



*Fig. 5: D4Science is listed as a provider in the Blue-Cloud catalogue.*

Home / Organisations / Blue-Cloud / D4Science Infrastructure

### D4Science Infrastructure

Followers  
**0**


[+ Follow](#)

**Rating**  
 ☆☆☆☆☆ (0)

**Your rating**  
 ☆☆☆☆☆  
 no rating given

---

**Organisation**



**Blue-Cloud**  
 The Blue-Cloud Project environment is designed to support the Blue-Cloud project activities and discussions. It is equipped with the following facilities:  
 Workspace: for sharing...  
[read more](#)

---

**License**

European Union Public License 1.1 [OPEN DATA](#)

Item
Groups

### D4Science Infrastructure

D4Science is an organisation offering a Data Infrastructure since 2009. D4Science is: connecting +14,000 scientists in 50+ countries, integrating data from +50 heterogeneous providers, executing +50,000 data analysis/month, providing access to over a billion quality records in repositories worldwide, operating with 99,8% service availability. D4Science hosts +150 Virtual Research Environments (VREs) to serve the biological, ecological, environmental, social mining, culture heritage, and statistical communities world-wide.

**Tags**

Data Infrastructure
Generic
Virtual Research Environment

**Data and Resources**


**D4Science Developers**  
 What you can accomplish with D4Science  
[Go to resource](#)

**D4Science Help Center**  
 Request a new Functionality on existing VRE - Request support for a...  
[Go to resource](#)

**D4Science Portfolio**  
 Supported communities and initiatives  
[Go to resource](#)

**Item URL**

[https://data.d4science.org/ctlg/Blue-CloudProject/d4science\\_infrastructure](https://data.d4science.org/ctlg/Blue-CloudProject/d4science_infrastructure)



**Basic Information**  
 Description: Basic Information

Field	Value
Abbreviation	D4Science
Legal Entity	true
Legal Status	Public Legal Entity
Website	www.d4science.org

**Marketing Information**  
 Description: Marketing Information

Fig. 6: Information about the provider D4Science in the Blue-Cloud catalogue.

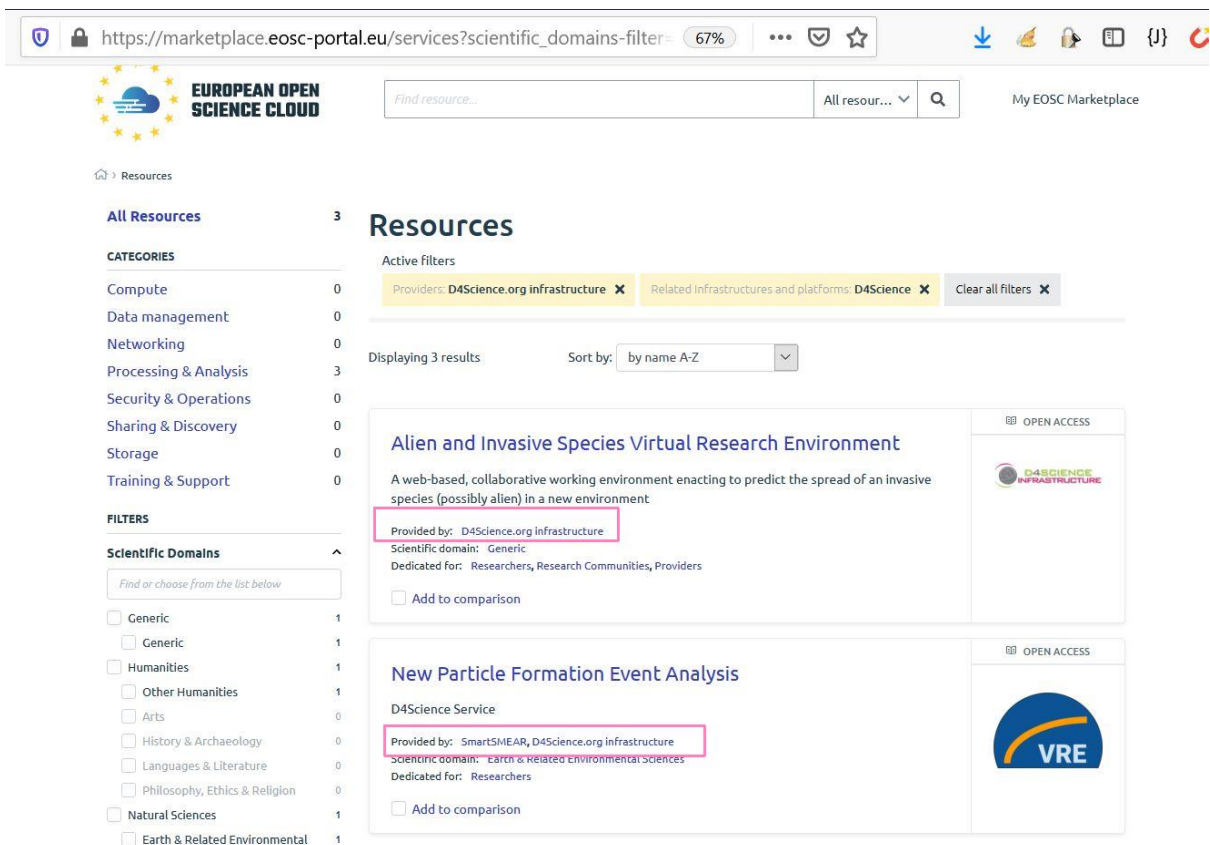


Fig. 7: D4Science is already a Provider in the EOSC Catalogue.

## o 4.2. Integration of the services

The main goal of the EOSC integration is to make the Blue-Cloud processing services visible and usable to a wider audience. Thus, the integration of the services into EOSC is the main activity to be completed in this task (task 4.5). For this, all relevant Blue-Cloud services must be identified. A list of services has been established and is listed in section 4.3.

As the Blue-Cloud catalogue adopted the same metadata standard as the EOSC catalogue, we will adopt a two-step approach:

**First**, the services will be thoroughly described according to the EOSC Profiles (see section 3.3) and included in the Blue-Cloud service catalogue (see section 2). This work heavily depends on the readiness of the services developed in the Blue-Cloud consortium as well as of the service owners' cooperation in providing service metadata.

**Then**, the descriptions of those services that are ready to be offered to a broader user base and comply to the EOSC preconditions (described in section 3.4, i.e. classified at least Technology Readiness Level 7, etc.) will be transferred into the EOSC catalogue.

The preferred method for this transfer would be the bulk-import described in section 3.5.3. As the Blue-Cloud catalogue is using the EOSC Profiles to describe their provider and services, the required compatibility is provided. Thus, if this functionality is being provided in time, Blue-Cloud could - and would like to - make use of it and directly onboard the entire Blue-Cloud catalogue (or the subset of

services which are considered mature enough). Obviously, as mentioned before, the services have to be fully described in the Blue-Cloud catalogue for the descriptions to be imported into EOSC.

If the content of the Blue-Cloud service catalogue cannot be onboarded in bulk-mode, we will set up a small service that will leverage the EOSC HTTP API (see section 3.5.2) to regularly push changes to service and provider descriptions that were made in the Blue-Cloud catalogue to the EOSC catalogue.

The bulk option would be the most simple and convenient option to integrate services, should it be available in time. It still has to be clarified whether there will be a filter functionality to limit integration to a subset of the services, in case not all services comply with the EOSC preconditions described in section 3.4. If this functionality is not viable, the HTTP REST API will be the preferred option for the service integration.

The technical side of these tasks is easily solved. However, there is a non-technical dimension - as in every undertaking that involves metadata -, which is making sure that the metadata is kept up-to-date in the first place. This challenge remains and may have to be addressed on the long run by the Blue-Cloud sustainability working group (WP6). However, this mainly concerns the Blue-Cloud catalogue, as metadata which is updated there will automatically be propagated to the EOSC catalogue.

### o 4.3. Services identified by the demonstrators

Before the Blue-Cloud services can be integrated into any catalogue, they have to be identified and their maturity evaluated. In the Blue-Cloud project, the following five demonstrators are developing services as well as exploiting services to solve specific marine-related use cases:

1. Zoo- and Phytoplankton EOY products (led by VLIZ);
2. Plankton Genomics (led by EMBL);
3. Marine Environmental Indicators (led by CMCC);
4. Fish, a matter of scales (led by FAO);
5. Aquaculture Monitor (led by FAO);

So far, a number of services have been identified from the deliverables D3.3 and D3.1; they are listed below. This list is preliminary and likely to undergo some refinements and revisions. Some of the services are specific to one of the demonstrators, while several others are common to all of the demonstrators. Once implemented and released, services are going to be integrated into the Blue-Cloud service catalogue. Ideally, all of them should be transferred into the EOSC catalogue eventually. However, as EOSC has some specific requirements and preconditions, it cannot be guaranteed that the entire set will be included right away.

Demonstrator #1 – Zoo- and Phytoplankton EOY products (led by VLIZ)

- Workflow to create interpolated maps of zooplankton abundances, using DIVA interpolation software, provided as a set of containerised Jupyter Notebooks (see Deliverable D3.3 for usage details). As of D3.3, the workflow is available to Blue-Cloud users;

- Workflow to create phytoplankton products using artificial neural networks. As of D3.3, this service is not publicly available yet.

#### Demonstrator #2 – Plankton Genomics (led by EMBL)

- Workflow to allow scientists to correlate unannotated raw imaging of unknown taxa and molecular data with known taxonomies, based on Jupyter Notebooks. According to D3.3, this service should become available in February 2021 and was not available yet at the time of writing.

#### Demonstrator #3 – Marine Environmental Indicators (led by CMCC)

- Marine Environmental Indicator App, an application that allows users to subset data from larger datasets (e.g. by time, indicator) and perform simple aggregating methods on them (e.g. monthly means, annual means), thus leading to more lightweight datasets for further analyses. For user's convenience, this app provides a GUI, while the processing is performed via a Web Processing Service (WPS);
- Oceans Pattern Indicator Model Development Service, which allows users to train a model on a custom dataset, and store the model as a NetCDF file;
- Oceans Pattern Indicator Prediction Service, which allows to use a model generated with the above service, and apply it to a study area to receive and display a predicted dataset;
- Storm Severity Index (SSI) Calculation Service, which calculates an SSI map for a specific study area and a custom time period based on either daily or monthly grid data provided by the user. The outcome is provided as a NetCDF file.

#### Demonstrator #4 – Fish, a matter of scales (led by FAO)

- This demonstrator is a rich environment of services that interact at various levels, e.g. the Fisheries Atlas with a focus on Statistical Data Management and the Global record of Stocks and Fisheries with a focus on stocks and fisheries Information management;
- Tuna / Fisheries Atlas, an online overview of harmonized time-series of catch and effort accessible through a map viewer, metadata and data services, analytical and reporting tools, and R Shiny, Jupyter and Markdown reporting services;
- Global record of Stocks and Fisheries, the global reference repository for stocks and fisheries accessible through a set of APIs that allow retrieving particular information for stocks and fisheries in a programmatic manner.

#### Demonstrator #5– Aquaculture Monitor (led by FAO)

- Aquaculture Cage Atlas: an online overview of satellite data derived maps of cages and cage clusters, delivered through an ISO-OGC compliant map viewer, allowing editing features of the detected cages and cage clusters (only for authorized users) and possibly providing estimates of cage activity over a production season. To be deployed for Blue-Cloud users in 2021;
- Aquaculture Ponds Atlas: Preliminary test service similar to the aquaculture cage atlas, but for coastal ponds. The result will be a coastal land-use classification map based on Copernicus remote sensing data. As of D3.3, this was going to be deployed as a test service for Blue-Cloud users in 2021.

Deliverable D3.1 lists a number of required services which are common between the demonstrators and which may thus be provided in a generic way. Like the list above, this list is likely to undergo corrections and revisions until the final list of services is established.

- Efficient access to common datasets using protocols and access mechanisms such as OpenDAP, Intake, OGC WMS/WFS, Data Cubes, Virtual File Systems/Cloud Storage such as Ceph;
- Visualisation tools for data (generic charts, charts on maps, bar chart, pie chart, time series, etc.) which are able to operate on multi-dimensional data tables, and GIS functionalities (online visualisation services, etc.);
- Data analysis tools based on Artificial Intelligence methods such as machine learning and/or deep learning for image pattern detection, data analyses and generation of indicators, etc.

## 5. Conclusions and next steps

In this deliverable, we described the Blue-Cloud Service Catalogue and the EOSC Portal Catalogue & Marketplace, and for both catalogues, we described the onboarding process for new providers and services as well as its update procedures. Among the service providers in the Blue-Cloud project - the Blue-Cloud project and the D4Science infrastructure - the latter has already been onboarded as a provider in both the Blue-Cloud Catalogue and the EOSC Portal Catalogue, while the former is still pending.

We described several ways how the providers and services can be integrated into the catalogues:

- By manual integration (via the EOSC Portal website and the Service Description Template provided by EOSC on request). For providers, this is currently the only way to register initially;
- By integration via the HTTP REST API interface provided by the EOSC Portal (for providers, only updates are possible this way, not the initial integration), or
- By bulk-integrating the entire Blue-Cloud service catalogue. Please note that this functionality is not currently available, but may become available at some point.

The third option would be the simplest and convenient option to integrate services, should it be available in time. As the Blue-Cloud Catalogue is using the EOSC Profiles to describe their provider and services, the required compatibility is provided. It still has to be clarified whether there will be a filter functionality to limit integration to a subset of the services, in case not all services comply with the EOSC preconditions described in section 3.4.

If this functionality is not viable, the HTTP REST API will be the preferred option for the service integration.

For each of these methods, the services must be thoroughly described by detailed metadata using the EOSC Resource Data Model. Furthermore, to be included in EOSC, the services must be at least at Technology Readiness Level (TRL) 7. A number of services have been identified and listed in this deliverable. The next step is to evaluate their TRL and describe and integrate them using one of the above manners. Furthermore, a technical solution to ensure that metadata updates in the Blue-Cloud catalogue are escalated into the EOSC catalogue will be developed and deployed.



## 6. 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](https://doi.org/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](https://doi.org/10.1016/j.future.2019.05.063)
- M. Assante, L. Candela, P. Pagano, A. Dell'Amico, L. Frosini, L. Lelii, M. Lettere, F. Mangiacrapa, G. Panichi, F. Sinibaldi (2020) ***Blue Cloud VREs Common Facilities (Release 1)***. D4.2 Blue-Cloud Deliverable, November 2020
- M. Assante, L. Candela, P. Pagano, R. Cirillo, A. Dell'Amico, L. Frosini, L. Lelii, M. Lettere, F. Mangiacrapa, G. Panichi, F. Sinibaldi (2021) ***Blue Cloud VREs Operation Report (Release 1)***. D4.1 Blue-Cloud Deliverable, February 2021
- 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
- G. Sipos, T. Zamani, K. Koumantaros, A. Coveney, I. D. Álvarez, N. Tonello, G. Koloventzos, O. Appleton, N. Liampotis (2020) ***EOSC-hub Integration Handbook for Service Providers***. 10.5281/zenodo.3826907