



Reinforcing the AI4EU Platform by Advancing Earth Observation Intelligence, Innovation and Adoption

D4.4: Testing and operation of bootstrapping services

Grant Agreement ID	101016798	Acronym	AI4COPERNICUS
Project Title	Reinforcing the AI4EU Platform by Advancing Earth Observation Intelligence, Innovation and Adoption		
Start Date	01/01/2021	Duration	36 Months
Project URL	https://ai4copernicus-project.eu/		
Contractual due date	30/6/2022	Actual submission date	30/06/2022
Nature	REPORT	Dissemination Level	PU
Author(s)	Marcin Ziółkowski (CF)		
Contributor(s)	-		
Reviewer(s)	George Stamouls (UoA), Michele Lazzarini (SatCen), Omar Barrilero (SatCen)		



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 101016798.

Document Revision History (including peer reviewing & quality control)

Version	Date	Changes	Contributor(s)
1	15/03/22	Initial version	Marcin Ziółkowski, CF
2	15/06/22	Added Sections	Marcin Ziółkowski, CF
3	23/06/22	Review	George Stamoulis UoA
4	28/6/22	Review	Michele Lazzarini, SatCen Omar Barrilero, SatCen
5	29/06/22	Finalising the document (pending review by EC)	Marcin Ziółkowski, CF

Draft Version

Executive Summary

This is the fourth deliverable of WP4 (D4.4 - Testing and operation of bootstrapping services). In this deliverable we present the architecture and services provided for testing and operation of bootstrapping services. This deliverable will describe the environment provided by CREODIAS to sustain and make bootstrapping services available. Tools needed to complete this deliverable were provided by CREODIAS DIAS Team.

Draft Version

Table of Contents

Introduction	6
Purpose and Scope	6
Approach for Work Package and Relation to other Work Packages and Deliverables	6
Organization of the Deliverable	7
Bootstrapping services Operation	8
Standard Management	9
Infrastructure Provision	10
Kubernetes Cluster	12
Docker Registry	12
Support	13
Standard support for the CREODIAS users	13
Dedicated Engineer support for the Bootstrapping Services	14
Dedicated Management support for the Bootstrapping Services	14
Testing	14

List of Figures

- Figure 1. CREODIAS Architecture
- Figure 2. High-level layout of the platform is presented.
- Figure 3. Table of cloud resources.
- Figure 4. Docker Registry placed within the AI4Copernicus integration Infrastructure

List of Terms & Abbreviations

Abbreviation	Definition
ML	Machine Learning
EO	Earth Observation
KG	Knowledge Graph
CM	Collaborative Management
AEM	AI4EU Experiments Management

1 Introduction

This is the fourth deliverable of WP4 (D4.4 - Testing and operation of bootstrapping services). In this deliverable we present the architecture of the provided DIAS solution to test and operate the bootstrapping services provided in WP5 of the project.

1.1 Purpose and Scope

The purpose of this deliverable is to report on the operation progress of the solution provided by CREODIAS and present what has been already achieved and what are the plans for future implementation. The presented services are based on CREODIAS Services and WEkEO elasticity which share the same platform as CREODIAS.

1.2 Approach for Work Package and Relation to other Work Packages and Deliverables

Work package WP4 started on M4 and ends on M24 of the project. It is led by partner CF with the collaboration of partners NCSR-D, UoA, TAS, ECMWF, UNITN. WP4 positions technically AI4Copernicus in the European AI and Copernicus ecosystems. In addition, it implements the software architecture of the project described in WP3.

The technical contribution of this task is the development of the software components for the project with specific emphasis to interfacing with the AI-on-demand platform and CREODIAS.

WP4 has the following five tasks:

- Task 4.1 Integration of AI4EU platform with CREODIAS/WEkEO [M4-M12, lead: CF, contributor: TAS]. Configuration of the environment to accommodate the requirements identified in the WP2.
- Task 4.2 Integration of tools for transformation, querying, interlinking and federating big linked geospatial data [M4- M12, lead: UoA] This task will integrate the linked data suite (developed by UoA) to the platform.
- Task 4.3: Implementation of the semantic catalogue and the semantic search and discovery functionality [M4-M12, lead: UoA, contributor: NCSR-D] This task will implement the semantic catalogue designed in Task 3.2.
- Task 4.4 Machine learning models for EO [M4-M24, lead: UNITN, contributors: NCSR-D, ECMWF] Different supervised machine learning techniques and models will be identified and integrated taking into account the inputs from WP2.
- Task 4.5 Testing and operation of bootstrapping services [M7-M18, lead: CF, contributors: NCSR-D, UoA]

Task 4.5 Testing and operation of bootstrapping services for AI4Copernicus will be based on dedicated environments which will be made available for the use cases. Access to these resources

was covered by the relevant budget for equipment. The resources include: VMs, storage, standard management and monitoring tools and related user support services. The resources can be on WEKEO -elasticity and on CREODIAS.

This deliverable does not focus on the description of the Bootstrapping Services which has been done in D5.1.

1.3 Organization of the Deliverable

The rest of the deliverable is organized as follows: Introduction chapter gives the information about scope of this deliverable, followed by “Bootstrapping service operations” chapter which describes the key elements provided by DIAS.

It includes description of:

- infrastructure
- management
- Docker Registry
- support

Creodias testing - touches the issue of testing with the reference to bootstrapping services.

Draft Version

2 Bootstrapping services Operation

The AI4Copernicus consortium provides a set of services and resources for the Security, Agriculture, Energy and Health domains to be utilized by the open calls winners. The development of these bootstrapping services aims to reduce the time and resources of the bidders in different tasks such as data access (EO and ancillary data), pre-processing, labelling datasets, ML algorithm definition. The AI4Copernicus consortium support to the bidders allows to address open calls winner's effort on the development of innovative services based on AI. Each service is documented in the section of the domain responsible for its deployment: following a cooperative approach, each service can be used by different domains, if relevant. The description of the services (e.g. the purpose, the input needed and the output produced) referred to the application deployed in the AI4Copernicus infrastructure is responsibility of WP5.

To support these bootstrapping services we first identified the following key elements:

- Standard Management
- Provision of the infrastructure
 - Kubernetes Cluster
 - Docker registry
- Support Management
- Internal Monitoring

Chapters below present how these elements were implemented within the CREODIAS Architecture, that supports the services with the environment needed to operate.

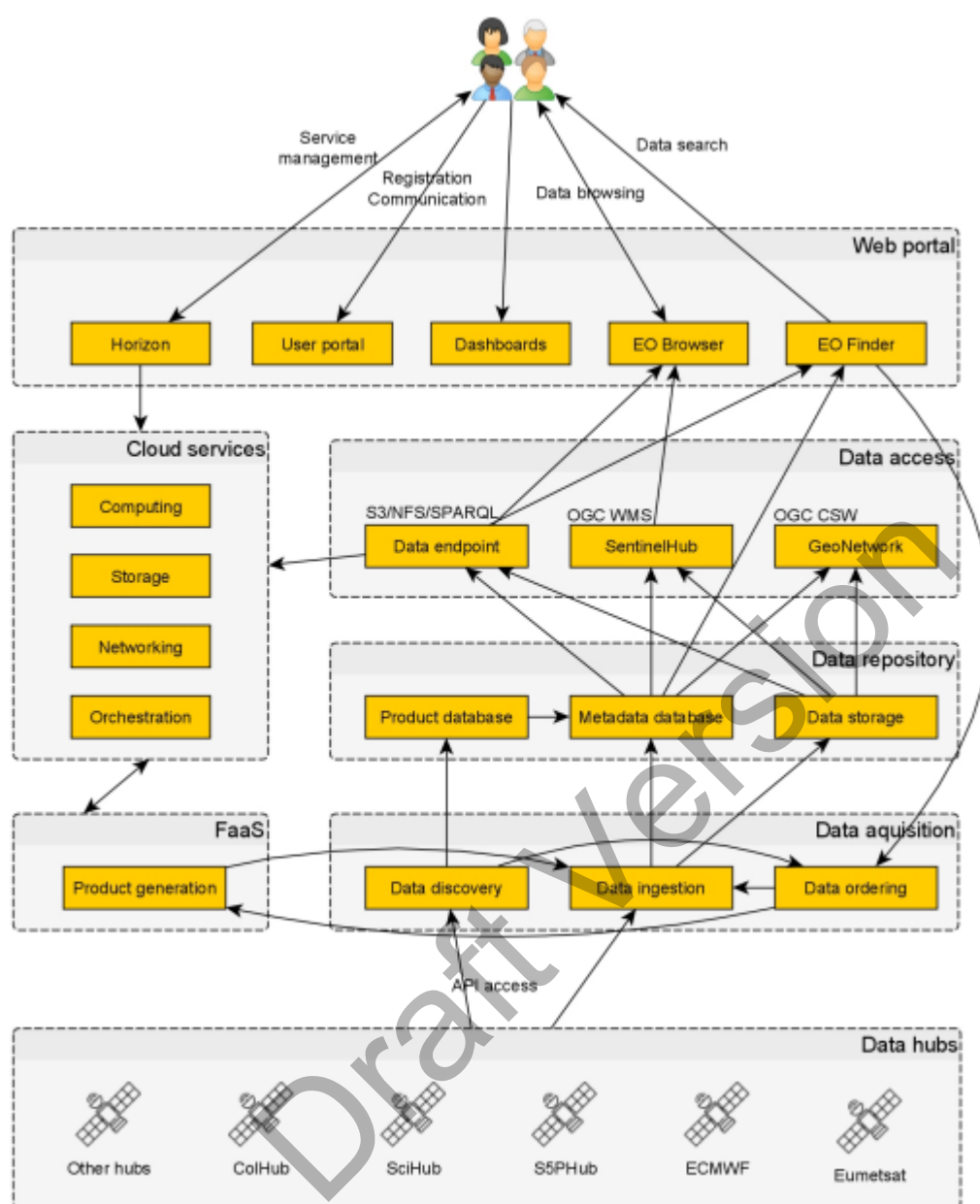


Figure 1. CREODIAS Architecture

2.1 Standard Management

CREODIAS is a seamless environment that allows processing of Earth Observation data (EODATA - EO DATA Free Archive). The platform hosts most of Copernicus Sentinel satellites data and Services, Envisat and ESA/Landsat data and other EODATA. Its design allows Third Party Users to prototype and build their own value-added services and products with a set of pertinent tools to guarantee simplicity, scalability and repeatability of any services' value chain.

A high-level layout of the platform is presented below.

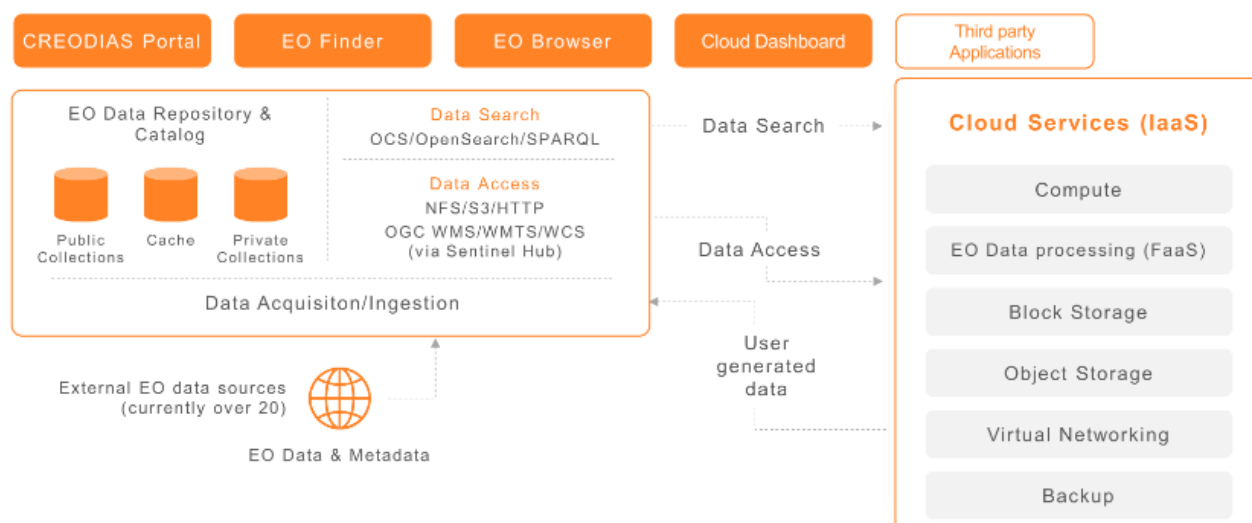


Figure 2. High-level layout of the platform is presented.

By default to operate the environment, a dedicated default Openstack dashboard, Horizon, is available for the users to allow compute and storage nodes creation and management. It provides administrators with a summary view of resources such as number of instanced VMs, allocated RAM, CPU and disk resources.

2.2 Infrastructure Provision

CREODIAS provides resources necessary to support provision of the bootstrapping services. Those resources are needed from the perspective of WP4 and WP5.

CREODIAS can provide for the use of the project machines built in dedicated cloud environment:

Virtual Machines				
Available VMs	#vCores	RAM (GB)	SSD Network Storage (GB)	SSD Local Storage (GB)
eo1				
eo1.xsmall	1	1	8	-
eo1.small	2	2	16	-
eo1.xmedium	1	2	8	-
eo1.medium	2	4	16	-
eo1.large	4	8	32	-
eo2				
eo2.medium	1	4	16	-
eo2.large	2	8	32	-
eo2.xlarge	4	16	64	-
eo2.2xlarge	8	32	128	-
eo2.3xlarge	16	64	256	-
eo2.4xlarge	32	128	512	-
eo2a*				
eo2a.medium	1	4	16	-
eo2a.large	2	8	32	-
eo2a.xlarge	4	16	64	-
eo2a.2xlarge	8	32	128	-
eo2a.3xlarge	16	64	256	-
eo2a.4xlarge	32	128	512	-
eo2a.5xlarge	64	236	1024	-
hm				
hm.medium	2	16	64	-
hm.large	4	32	128	-
hm.xlarge	8	64	256	-
hm.2xlarge	16	128	384	-
hm.3xlarge	32	256	384	-
hm.4xlarge	48	484	348	-
ds				
ds.3xlarge	48 (24 cores)	496	128	2x 1920

Figure 3. Table of cloud resources.

Within the CREODIAS cloud, users can use the data for free via rich multiple access interfaces like object S3, NFS, zipped download. Remote Transfer for EODATA with S3 protocol that could be configured within the kubernetes environment is also possible after some configuration. Another special feature is a dedicated Kubernetes environment is available to configure (either created by the user or directly from a template) on cloud environment.

A lot of additional information and dedicated manual's are accessible within CREODIAS' FAQ¹.

Few key elements that are fundamentals of integration architecture will be provided within DIAS infrastructure.

2.2.1 Kubernetes Cluster

Kubernetes is a portable, extensible, open-source platform for managing containerized workloads and services, that facilitates both declarative configuration and automation. It has a large, rapidly growing ecosystem. Kubernetes is often used when deploying data science models as it's important to be able to reproduce the same environment used for training and the environment used for serving. Kubernetes provides a tool for building reproducible environments, and they are much lighter weight than alternative approaches such as virtual machines.

We prepared an easy to follow guide² that describes how to install Kubernetes on CREODIAS OpenStack cloud, with support for adding/removing nodes, persistent volumes and load balancing. This deployment method uses Terraform and Ansible playbooks from the upstream Kubespray project with prepared Ansible configuration to enable required features. We believe it describes the best way to run Kubernetes smoothly and effectively.

CREODIAS provides instructions on how to run and operate the environment in FAQ.

2.2.2 Docker Registry

The docker registry is a docker repository designed for AI4Copernicus to facilitate the use of docker images by open call winners and to support the upload of the created solutions to the AI4EU Experiments environment .

Implementation of the registry is based on Harbor, an open source registry that secures artifacts with policies and role-based access control. This ensures images are scanned and free from vulnerabilities, and signs images as trusted. Harbor, a CNCF Graduated project, delivers compliance, performance, and interoperability to help you consistently and securely manage artifacts across cloud native compute platforms like Kubernetes and Docker.

CREODIAS plans to implement this and provide it as a service for the needs of the project. This is not a standard function provided by CREODIAS and will be customized to the needs of the potential users and managed by users themselves.

¹ <https://creodias.eu/faq>

² https://creodias.eu/faq-other/-/asset_publisher/SIs09LQL6Gct/content/how-to-configure-kubernetes

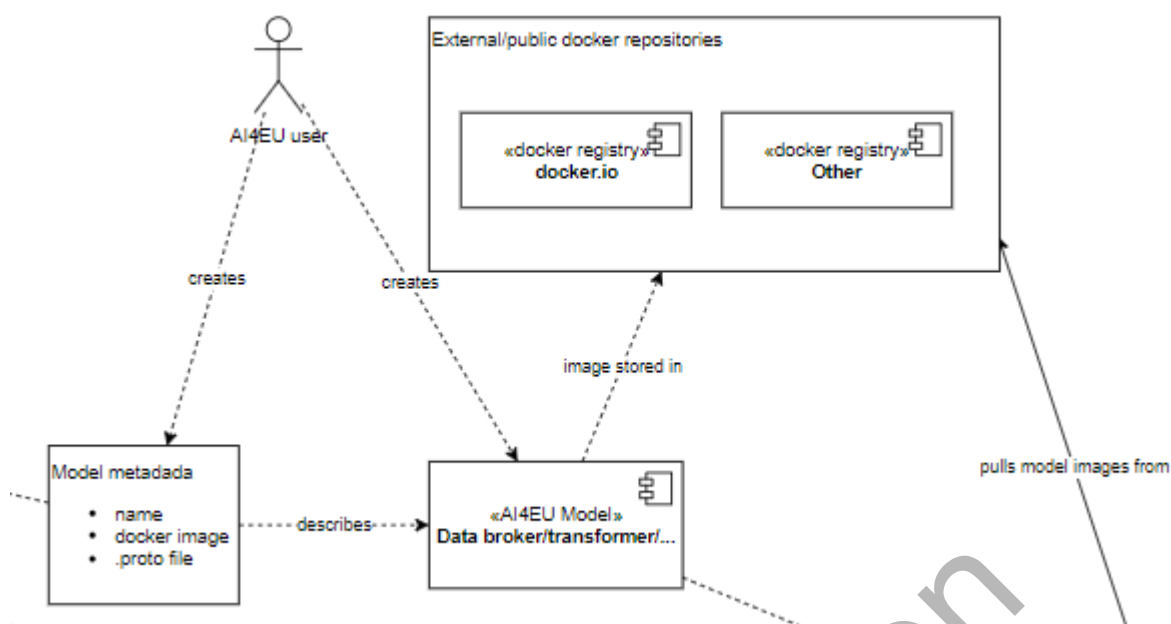


Figure 4. Docker Registry placed within the AI4Copernicus integration Infrastructure

2.3 Support

CREODIAS can support its users with the usage of the provided environment through CloudFerro support team. For the scope of the AI4Copernicus project and Bootstrapping Services few levels of support were presented:

■ Standard support for the CREODIAS users

CREODIAS always welcomes feedback from our users. To contact, there are multiple forms which include:

- dedicated contact form
- dedicated emails for particular tasks
- dedicated contact point as a manager for the provision of the product

If a user would like to proceed with a contact form, it is also a possible way to contact CREODIAS support team to respond with any relevant information.

Support	Main contact with our technical support. Doubts and questions related to the use of our service.	support@creodias.eu
Billing and Accounting	Any queries concerning settlements, invoices and accounting matters.	accounts@creodias.eu

Technical problems	Reporting problems and critical failures requiring a quick response, such as major service unavailability or severe performance issues.	problems@creodias.eu
--------------------	---	--

A contact form is also available for registered users through the Openstack environment and creates a ticket in an automated way.

■ Dedicated Engineer support for the Bootstrapping Services

CloudFerro as a provider of the infrastructure actively supported AI4Copernicus with the provision of the kubernetes cluster.

Dedicated DevOps engineers performed a short training video on how to set-up the kubernetes environment which was distributed between consortium partners.

■ Dedicated Management support for the Bootstrapping Services

CloudFerro within the project provides a dedicated Manager for the support of the activities mentioned above. This direct point of contact can operate on a different level of communication just to speed up the communication process.

3 Testing

The project involves several stages of tests that must be performed for the needs of the developed solution.

The first stage of tests of bootstrapping services was carried out in the CREODIAS environment on a dedicated k8s cluster by the WP5 team. This kubernetes cluster served as a test platform for WP5 users.

Cloudferro has provided the kubernetes cluster for the project needs. Cluster tests were performed in accordance with internal arrangements, based on the following steps:

Kubernetes cluster test lists	
ID	Proposed plan
1	<ul style="list-style-type: none">• Connect to cluster with config-file provided
2	<ul style="list-style-type: none">• Create namespace
3	<ul style="list-style-type: none">• Create storageclass, pvc and pod with a yaml
4	<ul style="list-style-type: none">• Access to pod to do processing
5	<ul style="list-style-type: none">• Delete pod and recreate it confirming that data in persistent volume is kept
6	<ul style="list-style-type: none">• Access to s3 catalogue and copy a product to persistent volume
7	<ul style="list-style-type: none">• Connect directly in s3
8	<ul style="list-style-type: none">• Executed s1 slc preprocessing using copied S1 data

The next stage of tests are related to checking the key elements that enable the integration of AI4EU with the CREODIAS environment.

The components necessary to test this are:

- Docker Repository - docker images repository
- AI4EU Integration Tests (docker Repository, Data Broker, CREODIAS k8s Environment).

Docker Repository is a key element that will allow for AI4Copernicus users to upload and make available the bootstrapping services.

Assumptions for the tests:

- Docker Repository will be placed in CREODIAS environment
- AI4C users will be able to upload images,
- Images have to be publicly available,
- Images can be downloaded to users private environment,

To prove that the solution is working we need to follow few steps

Docker Repository Test Plan	
STEP ID	Description
1	Cloudferro will create accounts for AI4Copernicus users
2	AI4Copernicus users will check the possibility of logging into the application
3	AI4Copernicus user will upload docker images to Docker Registry
4	The users will share the images publicly
5	The user will try to download images from the registry to his own environment
6	User will share the results with AI4Copernicus

Structure and Idea of docker registry was described in Chapter 2.2.2

Docker Repository is one of the key elements of the integration between AI4EU and CREODIAS. It is an element that acts as a repository for models that are uploaded into the Acumos environment (the url for images will be available from the Docker Repository). Therefore, after checking the correctness of Docker Repository operation, the next step is to check its interaction with AI4EU Experiments and with the Kubernetes cluster in the CREODIAS environment.

The assumptions are as follows:

- - Docker Repository is responsible for uploading and hosting images for AI4EU Experiments
- - AI4EU Experiments is responsible for creating models by creating a solution zip
- - Kubernetes cluster should enable the launch of the created solution

Bootstrapping services will also require to perform integration test which could be described as follows:

Integration Tests Steps	
Step ID	Description
1.	Models from Docker Registry will be available to deploy on Acumos
2.	User will create a pipeline using an elements like bootstrapping service (and for example a DataBroker)
2a	Bootstrapping service will be uploaded to Acumos (needed Docker registry to work).
2b	DataBroker will be available to use.
3	User will create a pipeline using the Bootstrapping service and DataBroker
4	User will create solution.zip based on the created pipeline.
4a	User will upload solution.zip manually to kubernetes environment in CREODIAS
4b	User will use dedicated deployer (button to deploy solution.zip in CREODIAS)
5	User has to check the results: <ul style="list-style-type: none">- Uploaded model should work in kubernetes environment- Results should be available and accessible / Shared Volume

4 Conclusions

Within the scope of the AI4Copernicus Project, CREODIAS as a main provider for DIAS infrastructure was responsible for provision of environment for the operation of the bootstrapping services. In this deliverable, we presented the activities we carried out in M1-M18 of AI4Copernicus in order to provide an environment for the project services. We have tested the most important elements of the infrastructure.

Draft Version