D2.4 - **Technical, Operational and Interoperability Specifications and Architecture**

Work Package	Requirements Elicitation and Conceptual Framework Specification
Lead Author (Org)	NKUA
Contributing Author(s) (Org)	ECMWF, NVCR, CINECA, HES-SO, CMCC, DANAOS, KEMEA, EBOS, ENG, IES, MEEO, IVI
Due Date	30.11.2023
Date	20.11.2023
Version	V5.0

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)

Disclaimer

Versioning and contribution history

Version	Date	Author	Notes
0.1	31.07.2023	1st version; Structure of the deliverable defined	TOC and V0.1
0.2	29.08.2023	1st Input from the partners	
0.3	30.10.2023	1st Review; Additions and updates to the content and the sections	
0.4	12.11.2023	2nd Review;	
0.5	17.11.2023	Final Version	

Terminology

Terminology/Acronym	Description
AI	Artificial Intelligence
API	Application programming Interface
AUX	Auxiliary
CI/CD	Continuous Integration/Continuous Delivery
CLI	Command Line Interface
CNCF	Cloud Native Computing Foundation
CNI	Container Network Interface
CSA	Coordination and Support Action
CSI	Container Storage Interface
DDAV	Dashboard – Data Analytics Visualisation
DoW	Description of Work
DSL	Domain Specific Language
EC	European Commission
EO	Earth Observation
EOSC	European Open Science Cloud
FaaS	Function as a Service
FE	Fusion Engine
GA	Grant Agreement to the project
GUI	Graphical User Interface
HPC	High Performance Computing
IaaS	Infrastructure as a Service
IaC	Infrastructure as Code
KG	Knowledge Graph

Terminology/Acronym	Description
ML	Machine Learning
OCCM	OpenStack Cloud Controller Manager
OCI	Open Container Initiative
ONNE	Open Neural Network Exchange
PaaS	Platform as a Service
SSO	Single-Sign On
UI	User Interface
UML	Unified Modeling Language
UMM	User Management Model
VPN	Virtual Private Network
WFE	Work Flow Editor
XR	Extended Reality

-

Table of Contents

List of Figures

List of Tables

Executive Summary

The deliverable provides a first version of the EO4EU technical requirements that will form the basis for the definition of the Platform and EO4EU architecture and detailed design. In keeping with the overall project workflow, the requirements captured and synthesized here will be given to WP3 as input for the first development cycle.

Within the deliverable a description of the architecture to tiers is presented. An analytical representation of components and their class diagram is described. The interaction between the components is presented in sequence diagrams also.

1 Introduction

1.1 Scope of D2.4

The purpose of this document, "D2.4 Technical Operational and Interoperability Specifications and Architecture" is to define the design principles of the components through EO4EU architecture, to define the technical specification of the components with their respective interfaces based on the defined requirements coming from DoW, Consortium members and external end users. This deliverable also includes several sequence diagrams serving as reference of the interactions between the components.

This document is structured as follows:

- Chapter 2 presents the EO4EU architecture as a whole.
- Chapter 3 contains all the respective components in each tier. The technical description includes title, description, a list of the provided functionalities, relation to other components, related use cases presented in Chapter 4. The description is enhanced by Unified Modeling Language (UML) class diagram describing the main functionalities of the components. Chapter 3 provides a list of the requirements as finalised in the consortium (both functional and non-functional).
- Chapter 4 focus on the presentation of sequence diagrams that explain the interactions of the components and create a workflow of use.
- Chapter 5 contains the conclusions.

1.2 Relation to other deliverables

A detailed user requirement analysis was given in D2.2. Based on these requirements and the planned functionalities from the DoW, this architecture document was created.

2 Architecture

This chapter gives an overview of the architecture and its components.

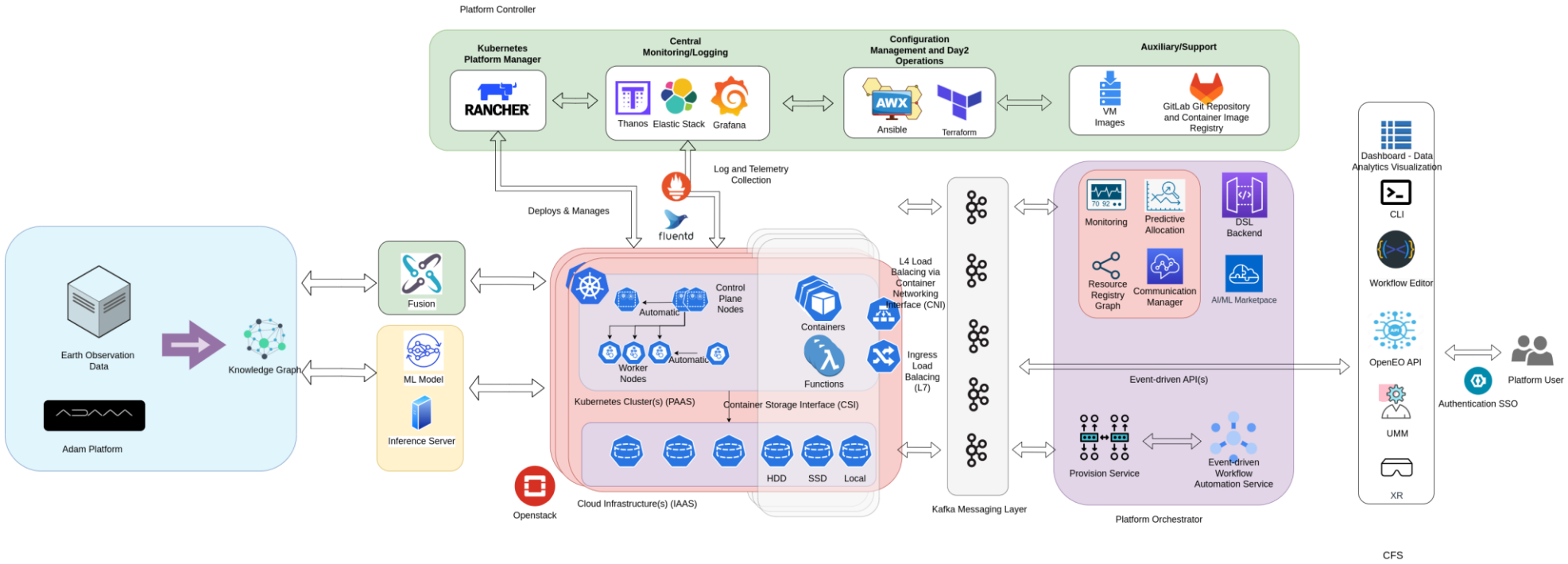

Figure 1 EO4EU Architecture

EO4EU architecture consists of five different tiers (see Figure 1):

- Data Tier (leftmost part of Figure 1): A set of data sources is the input of the platform. Heterogeneous data that need pre-processing. For this purpose, at this tier EO4EU provides a Knowledge Graph, a Fusion engine and a ML component to refine and homogenise data.
- Infrastructure as a Service (IaaS) Tier (bottom-centre part of Figure 1): Provides the necessary abstraction of the infrastructure, allowing PaaS components to efficiently consume compute, storage and networking resources dynamically, programmatically (via Openstack REST APIs).
- Platform as a Service (PaaS) Tier (central top and left part of Figure 1): Provides a higher-level cloud computing service compared to IaaS. PaaS builds upon the foundational infrastructure offered by IaaS and provides additional services and tools to simplify application development, deployment, and management.
- ML Tier (not clear where in Figure 1): Provides all machine learning models in a toolbox for the post processing of the retrieved or fused data.
- Front-end Tier (rightmost part of Figure 1): Provides multi-dimensional User Interface-UI (Web, XR, CLI, API) that enables the user to interact and control the platform.

A brief mapping of the components in the tiers can be found in Annex 1. Architectural design is the first layer of analysis of the components that are developed in the EO4EU ecosystem. Child components can be found in the development analysis of the Class Diagrams shown in the Chapter 3 and their technical details are described in depth in WP3 deliverables.
EO4EU aims to further improve access to the EU EO data offered by a variety of platforms and data repositories going beyond similar projects described in Section 5 - Related Projects. Data sources include Copernicus services and associate platforms like the Copernicus Data Access Service (C-DAS), Data and Information Access Services (DIAS[1]), and also upcoming platforms and initiatives like Destination Earth (DestinE[2]). Without prior knowledge about their structure and format, the platform shall be able to retrieve, process, fuse and deliver new datasets, supported by machine learning algorithms and advanced semantic annotations.

An innovative ML-based learned compression algorithms shall enhance the accessibility of the data sources by reducing the respective data volumes needs to be transferred over the network, reducing the footprint of storage capacity and the network bandwidth requirements.

The control and core plane components of the EO4EU platform adapt an event-driven and microservices-based architecture, hosted on the PaaS Tier. PaaS platforms often integrate with Kubernetes to simplify microservices development, scaling, and management. This combination accelerates the adoption of microservices by abstracting complexities, offering pre-built services, and streamlining deployment and scaling processes.
User Experience are further enhanced with a set of visualization services and interfaces, including a multi-layered user interface (GUI) for visual analytics coupled with a Workflow Editor, a Command Line Interface (CLI), and a respective Application Programming Interface (API), and an extended reality (XR) interface to further boost the usability and the adoption of the platform, combining traditional access methods with cutting-edge technology stack.

All platform communications are handled through a message-based middleware (via a Communication Manager and its equivalent Message Bus). This provides a coherent communication model with distribution, replication, reliability, availability, redundancy, backup, consistency, and services across distributed heterogeneous systems. This Message Bus communication system

[1] https://www.copernicus.eu/en/access-data/dias
[2] https://destination-earth.eu/

interconnects all components and all tiers. It can be used for asynchronous notifications and asynchronous method calls / response handling.

3 Components Overview

This chapter describes at a high level the components and the interactions between them and their interfaces. The respective deliverables in WP3 will give the development details of the components in all tiers. Chapter 4 provides representative use cases and corresponding sequence diagrams depicting interactions among the various components.

3.1 Data Tier

Component	**Knowledge Graph**
Responsible partner	NVCR
Parent Component	Direct interconnection with Earth Observations (EO) data providers. Specifically, either in services which enable users to consume EO data through a dedicated API such as ADS[3], CDS[4], ADAM, or services that output raw/processed EO data through a download link.
Description	The Knowledge Graph (KG) is providing users with an extended search capacity that supports free text queries, enabling accessing EO data based on semantic search. By these means it allows non-expert users to access, uptake and work with previously unknown and undiscoverable resources. The KG semantically processes available datasets (set of processed raw EO data) from various sources, such as Copernicus Services (ADS, CDS, Land[5], Marine[6]), as well as third-party platforms such as ADAM. It gathers the textual descriptions (metadata) of the datasets, created through the embeddings of their descriptions. It also accepts input from either the users or an automatic request from another component through a dedicated API. This request is being transformed into an internal representation (vector). The KG identifies the similarity between the internal representation of the query of the user (or the component) with the vector produced for each dataset derived from each different service to provide a list of results to the user.
Provided functionalities	1. The user performs a search based on dedicated queries. The response leads the request to the most similar datasets in terms of the description of the dataset, but not by source, in the form of a list that also provides the name of the dataset and its origin. The user chooses one of the outputs, which is considered more relevant. 2. The user sends a second request to the KG, and it gets as a response a dataset that is prioritized based on similarity in terms of products, features or options.

[3] https://ads.atmosphere.copernicus.eu/cdsapp#!/home

[4] https://cds.climate.copernicus.eu/#!/home

[5] https://land.copernicus.eu/en/dataset-catalog

[6] https://data.marine.copernicus.eu/products

	3. The user selects the items of interest and sends a third and final request to the system, and as a response it gets the code corresponding to the previous step, alongside the appropriate API call.
Relation to other components	The APIs produced will be consumed by the visualization dashboard and the fusion engine.
Related user stories section	4.1

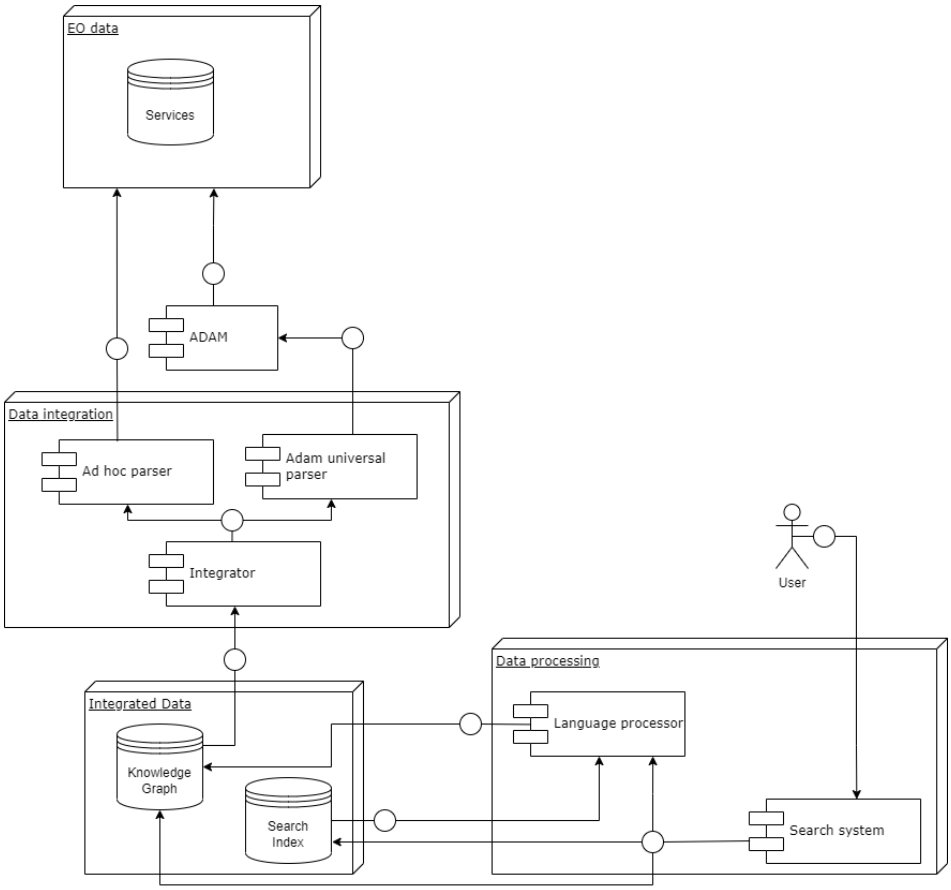

Figure 2 KG Class Diagram

Component	**List of Datasets integrated**
Responsible partner	NVCR
Parent Component	Knowledge Graph
Description	The sources identified to query and aggregate data from are: - CAMS forecast O3 - CAMS forecast CO - CAMS forecast SO2 - CAMS forecast NO2 - CAMS forecast PM10 - CAMS forecast PM2.5 - CAMS forecast birch - CAMS forecast olive - CAMS forecast grass

	- CAMS forecast ragweed - CAMS forecast alder - CAMS forecast mugworth - "Sensory operational information (speed, draft, cons., power, fuel cons., etc)" - weather data (NOAA) - AgERA5 2m temperature - AgERA5 Precipitation flux - AgERA5 2m relative humidity - AgERA5 10m wind speed - AgERA5 Solar radiation flux - Sentinel-1 GRD - Sentinel-2 L1C - L2A - ESA World Cover - Sicily CTR (Regional Technical Chart) - ISTAT crop and production - EURO-CORDEX (historical simulations and projections) - Temperature - EURO-CORDEX (historical simulations and projections) - Precipitation - EURO-CORDEX (historical simulations and projections) - Relative humidity - EURO-CORDEX (historical simulations and projections) - wind speed - EURO-CORDEX (historical simulations and projections) - Solar radiation - CMCC VHR (historical simulations and projections) - Temperature - CMCC VHR (historical simulations and projections) - Precipitation - CMCC VHR (historical simulations and projections) - Dew point temperature - CMCC VHR (historical simulations and projections) - Wind speed - CMCC VHR (historical simulations and projections) - Solar radiation - ERA5-Land - Temperature - ERA5-Land - Precipitation - ERA5-Land; - Radiation - ERA5-Land- VPD - CERRA - Temperature - CERRA- Precipitation - CERRA - Radiation - CERRA - VPD - EURO-CORDEX (historical simulations and projections) - VPD - SOILGRIDS - ERA5-Land - Available Soil Water - CERRA - Available Soil Water - ESDB- Available Soil Water - EU-DEM - PROFUND - CO2 Content - Point scale gauges (digitized observations) - UERRA (reanalysis) - E-OBS (interpolation)

	- "Climate (sub-)seasonal forecasts - (TBD if and whIch will be used)" - Copernicus Land vegetation products - Copernicus Land land cover products - soil products (SOILGRIDS and/or ESDB) - OpenStreetMap - roads and railways - ERA5 Land LAI (leaf area index) - FAO locust location reference swarms - FAO locust location reference hoppers - Seasonal weather forecast - temperature - Seasonal weather forecast - solar radiation - Seasonal weather forecast - soil water content - ENDVI10_metop LST - ENDVI10_metop NIR SWIR - ENDVI10_metop NIR SWIR archive - ENDVI10_metop NDVI - ENDVI10_metop NDVI archive - Eumetsat LSA-SAF FRMv2 - FIRMS Fire HotSpots
Provided functionalities	n/a
Relation to other components	n/a
Related user stories section	All

3.2 IaaS Tier

Component	**Openstack**
Responsible partner	ECMWF
Parent Component	Cloud Infrastructure (IaaS)
Description	Openstack[7] is an open-source cloud computing platform that provides IaaS capabilities. It offers foundational cloud infrastructure services such as computing, storage (block and object), and networking. Openstack is the foundation of the infrastructure layer for higher-level services like Kubernetes[8] and PaaS tier, making it a versatile choice for building and managing cloud environments.
Provided functionalities	Openstack provides the foundation of the infrastructure layer for higher-level services like Kubernetes and PaaS tier in EO4EU, making it a versatile choice for building and managing cloud environments. EO4EU adopted a multi-cloud strategy and capitalized on the capabilities of WEkEO and CINECA Openstack infrastructures to architect a robust and resilient environment for its diverse set of applications. Through subnetting and assigning separate networks Security Classes and unique Kubernetes pod and service CIDR blocks across the clouds, the platform ensured smooth routing and interconnectivity between cloud environments. A key feature of this setup is the detailed separation of workloads into different subnets,

[7] https://www.openstack.org/
[8] https://kubernetes.io/

	known as the Defense-in-Depth (security zones) strategy, boosting security by keeping workloads isolated by their domains (e.g., operational, development, and Kubernetes). Via employing the Defense-in-Depth strategy, EO4EU embraced a multi-layered security approach, similar to a castle with several concentric walls. This means, that even if a security breach occurs at one level, additional security layers are present to thwart or slow down potential attacks, safeguarding the infrastructure and data across various levels and thereby ensuring consistent and robust cybersecurity, which is essential for public clouds. Furthermore, this architecture facilitates streamlined and effective monitoring across the clouds, enhancing administrative oversight and service management. EO4EU adopted Infrastructure as Code (IaC) via Terraform[9] and Ansible[10] to ensure the infrastructure is not only replicable but also inherently flexible to accommodate evolving needs and potential expansions via IaC templates. These IaC templates enabled the precise and efficient replication of the setup, ensuring that the EO4EU platform remains cloud-agnostic and maintains its versatility and reliability across diverse cloud environments.
Relation to other components	3.3
Related user stories section	

[9] https://www.terraform.io/
[10] https://www.ansible.com/

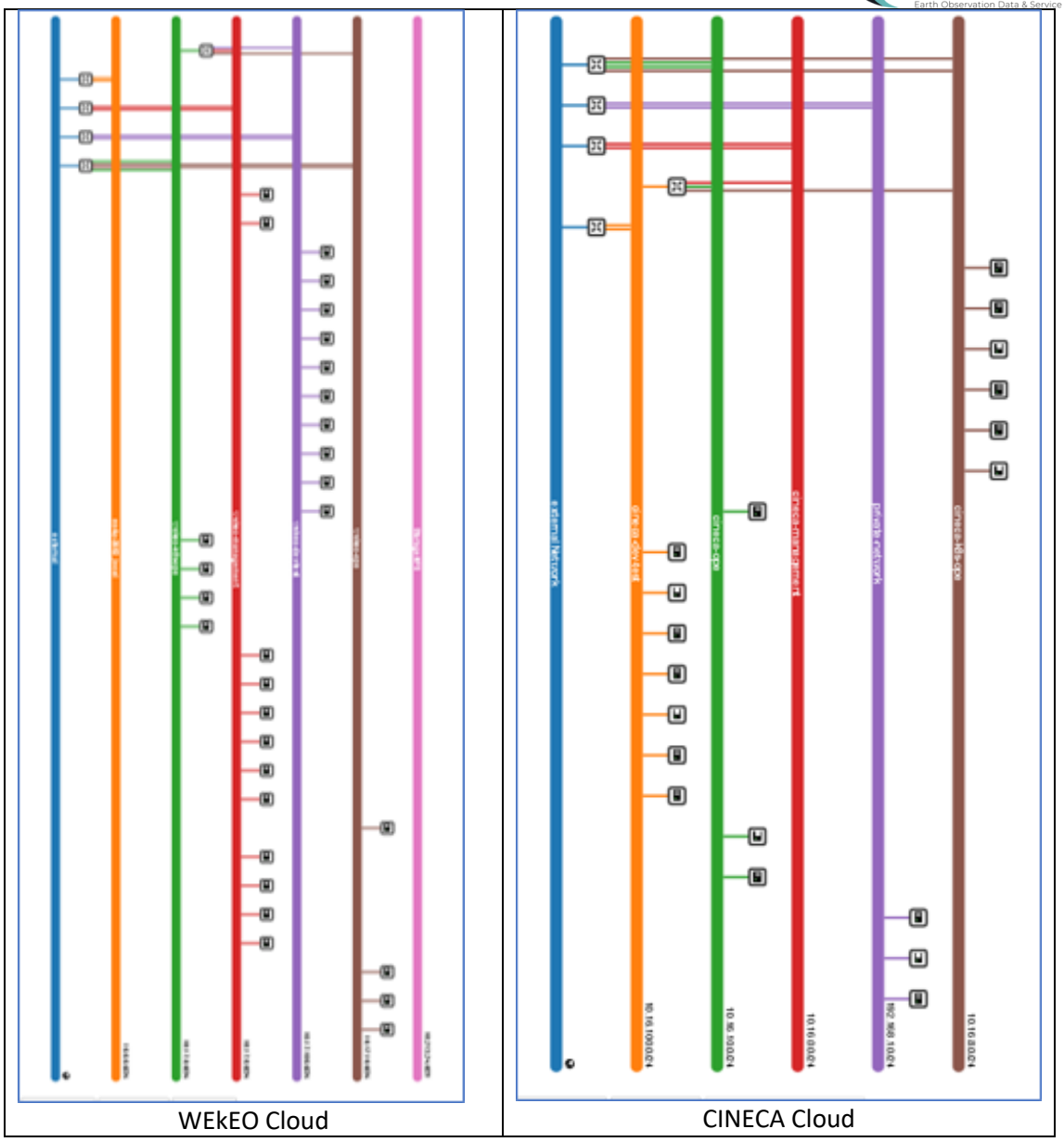

Figure 3 Openstack Defence-in-Depth Networking Setup

3.3 PaaS Tier

Component	**Kubernetes Platform Manager**
Responsible partner	ECMWF
Parent Component	Platform Controller
Description	Rancher Cluster Manager[11] provides a robust platform for deploying, managing, and orchestrating Kubernetes, with native integration with Openstack via ClusterAPI, and unified access across multiple clusters, providing a GUI and an extensive API.
Provided functionalities	PaaS architecture has been built upon the multi-cloud IaaS platform of EO4EU, which utilizes CINECA and WEkEO Openstack Cloud Infrastructures. Leveraging Rancher Cluster Manager and RKE2[12] for

[11] https://ranchermanager.docs.rancher.com/getting-started/overview

[12] https://docs.rke2.io/

	Kubernetes distribution, the platform was designed to provide seamless access to the Kubernetes APIs, ensuring optimal use and management of Kubernetes resources. To achieve efficient and standardized resource provisioning and management, Terraform templates were developed, in line with the IaC best practices. This approach not only automates and standardizes infrastructure provisioning but also ensures replicability and consistency across deployment life cycles. By adopting the GitOps approach and using the Git repository as the single source of truth, the platform enables automated and auditable changes, improving both reliability and operational efficiency. The integration of modern tools such as OpenStack Cloud Controller Manager (OCCM), Kubernetes Operators, and Helm Charts allows the effective deployment and management of cloud resources and higher-level services, providing them seamlessly as a service (X-as-a-service model). Several key design decisions have been made to improve cluster life-cycle management, which includes introducing backups to S3 object storage, adopting Calico[13] as the Container Network Interface (CNI), and Cinder as the Container Storage Interface (CSI) provider. Incorporating OpenFaas[14] into the platform enables the provision of functions-as-a-service, supporting a dynamic, event-driven infrastructure and applications. Another priority was placed on ensuring comprehensive observability within the environment. By incorporating monitoring, auditing, logging, and tracing capabilities with Cloud Native Computing Foundation[15] (CNCF) and industry-standard tools like Prometheus[16], Thanos[17], and the Opensearch[18] stack, consistent performance, audit, and issue/bottleneck identification was ensured, allowing for timely identification and resolution of issues.
Relation to other components	3.2
Related sequence diagrams sections	4.2, 4.3, 4.4, 4.5

[13] https://www.tigera.io/project-calico/
[14] https://www.openfaas.com/
[15] https://www.cncf.io/
[16] https://prometheus.io/
[17] https://thanos.io/
[18] https://opensearch.org/

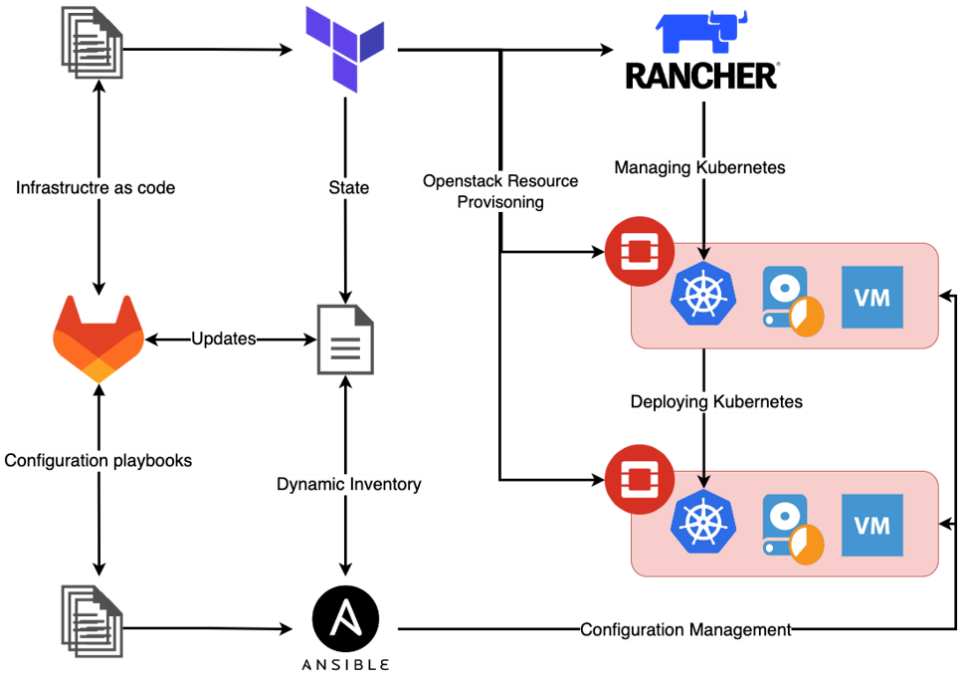

Figure 4 IaC and GitOPS Pipeline Design

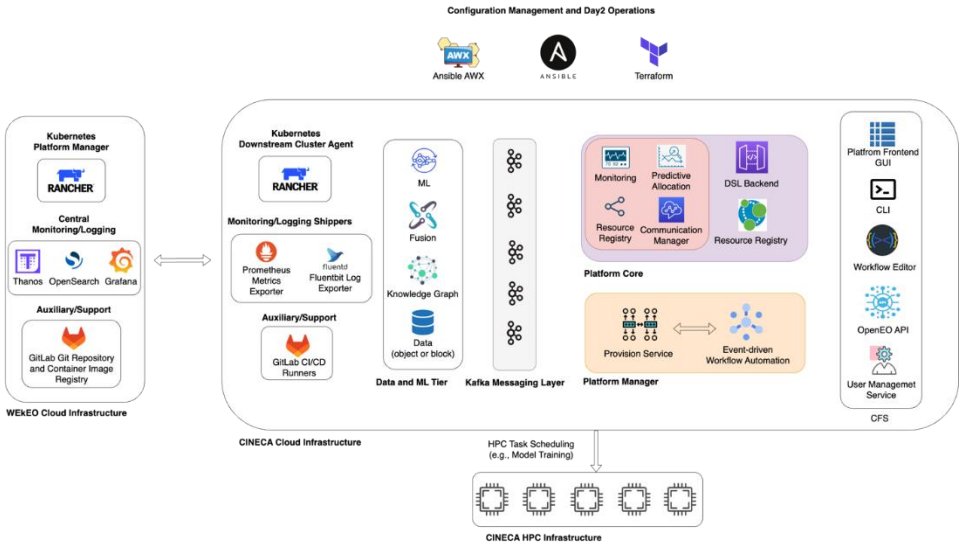

Figure 5 EO4EU Multi-cloud Infrastructure High-level Design

Component	**Auxiliary/support - GitLab Container Image Registry**
Responsible partner	ENG, NKUA
Parent Component	Platform Controller
Description	GitLab Container Image Registry allows users to securely store and distribute container images, and other Open Container Initiatives (OCI) compatible artefacts (e.g, Helm Charts) within the GitLab environment. Through a private registry, CI/CD integration, and precise access control, it facilitates the management of container images during the application development and deployment process.

Provided functionalities	- **Centralized Repository for Codebase and Container Images:** GitLab Container Image Registry allows users to centrally store and maintain container images within the same code repository, allowing for controlled access and version tracking. - **Access and permission control:** GitLab Container Image Registry allows managing access rights and permissions for container image management. You can define who can view, download, upload or delete images. This level of control helps ensure the security and proper handling of container images. - **Private Registry:** GitLab Container Image Registry offers the ability to create a private registry of Docker[19] images. This means that images can only be stored and deployed within the specific GitLab environment, providing greater security and control over the container images used in the development and deployment processes. - **CI/CD Integration:** GitLab Container Image Registry is fully integrated with GitLab's CI/CD module. This means that container images created during CI/CD pipelines can be automatically pushed into the registry and used in later stages of the deployment process. This integration greatly simplifies the development and deployment workflow for container-based applications.
Relation to other components	Configuration Management and Day2 Operations
Related sequence diagrams sections	4.2, 4.3, 4.4, 4.5

Component	**Auxiliary/support - GitLab Git Repository and CI/CD**
Responsible partner	ENG, NKUA
Parent Component	Platform Controller (Auxiliary/support)
Description	GitLab and Gitlab CI/CD (Continuous Integration/Continuous Delivery) provide developers a comprehensive and flexible platform to automate the software development lifecycle, enabling faster and more reliable delivery of applications from a web-based DevOps platform for version control, issue tracking, and continuous integration and deployment.
Provided functionalities	**GitLab**: - **Git Repository Hosting and Version Control:** Provides a centralized platform for hosting and managing Git repositories, making it easy to collaborate and track code changes. - **Issue Tracking:** Includes built-in issue tracking and project management tools, enabling teams to organize tasks, track progress, and resolve issues effectively.

[19] www.docker.com/

	- **Security:** Offers security scanning and vulnerability management to identify and address security issues in code. - **Terraform Integration:** Supports integration with Terraform, a popular IaC tool, allowing for the management and automation of cloud infrastructure deployments. **GitLab CI/CD :** - **Automated Build:** Automates the process of building software applications from source code, ensuring consistency and repeatability. - **CI/CD:** Automatically runs tests and checks code quality with each code change, catching issues early in the development process and streamlining the deployment of applications to various environments, from staging to production, reducing manual errors and speeding up release cycles. - **Pipeline Configuration:** Allows developers to define CI/CD pipelines using a .gitlab-ci.yml configuration file, specifying stages and jobs. - **Self-Hosted GitLab Runners:** Allows to set up and use their own computing resources as GitLab runners, providing control over the execution environment and scalability for CI/CD jobs.
Relation to other components	Configuration Management and Day2 Operations
Related sequence diagrams sections	4.2, 4.3, 4.4, 4.5

Component	**Auxiliary/support - Configuration Management and Day2 Operations**
Responsible partner	ENG, NKUA
Parent Component	Platform Controller (Auxiliary/support)
Description	**Terraform:** Terraform is an IaC tool that enables the provisioning and management of infrastructure resources using code. It's ideal for creating, updating, and destroying cloud resources to achieve the desired infrastructure state. **Ansible:** Ansible is an automation tool that uses YAML playbooks to configure, deploy, and manage servers and applications. It excels in configuration management and is well-suited for ongoing automation tasks, including Day 2 Operations such as maintenance and scaling.
Provided functionalities	**Terraform:** - **Infrastructure Provisioning:** Terraform enables the automated provisioning and management of infrastructure resources on cloud platforms.

	- **IaC:** It allows users to define infrastructure configurations in code, making infrastructure changes repeatable and version-controlled. **Ansible:** - **Configuration Management:** Ansible automates the configuration of servers, ensuring consistent and desired states across an infrastructure. - **Application Deployment:** It simplifies application deployment by orchestrating tasks and workflows across servers and environments. - **Task Automation:** Ansible streamlines automation tasks, making it easy to perform routine operations, updates, and scaling of infrastructure and applications.
Relation to other components	Auxiliary/support - GitLab Git Repository and Container Image Registry
Related sequence diagrams sections	4.2, 4.3, 4.4, 4.5

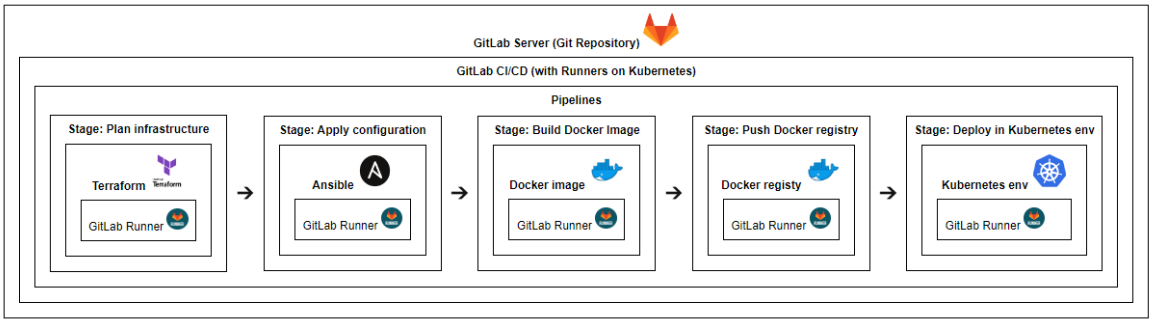

Figure 6 **Continuous integration general schema**

Component	**Platform Orchestrator**
Responsible partner	NKUA
Parent Component	Platform Controller
Description	Consisting of five (5) sub-components, the Platform Orchestrator will be the heart of the platform: (i) Monitoring Manager for event handling both for platform and workflow evolution; (ii) Communication Manager to handle messaging systems (Kafka, Nats, etc) and scratch (ephemeral) storage; (iii) Predictive Allocation to provide resource resilience in live workflows; (iv) Registry Handler interfacing the resource and service registry; (v) Provision Manager to handle the creation and the lifecycle of each user defined workflow.
Provided functionalities	- Leverage on monitoring component to retrieve relevant - Event categorization, statistical analysis, event/alarm push mechanism.

	- Create, Update, Delete messaging systems (pub/sub) entities. - Create, Update, Delete scratch storage (pvc, S3, etc) as per workflow data store. - Create, Read, Update, Delete entities in the Resource Registry according to the Resource Registry rules (rules of relations). - Pattern recognition in the Resource Registry and resource and service optimization actions. - Validate, Create, Read, Update, Delete of a user-defined workflow.
Relation to other components	All components
Related sequence diagrams sections	4.2, 4.3, 4.4, 4.5

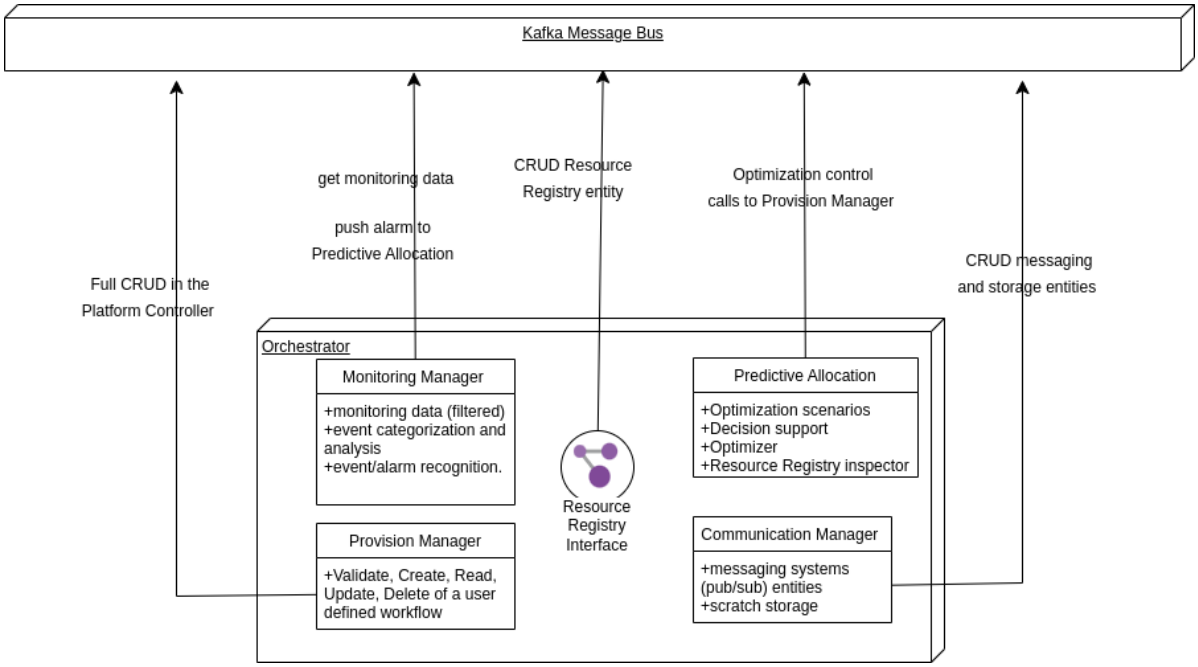

Figure 7 Platform Orchestrator Class Diagram

Component	**Authentication SSO**
Responsible partner	EBOS
Parent Component	Identity and Access Management (Key Cloak)
Description	The SSO (Single Sign On) functionality allows users to sign in once and get access to a set of different applications. Such functionality requires the use of cookies and therefore the access is granted per browser. The SSO is an identification and an authentication method that enables users to log into the EO4EU Software platform system that provides access onto multiple applications and services with one set of credentials. SSO streamlines the checking authentication process for users. The Authentication SSO mechanism is being integrated into the EO4EU Software Platform, which exploits a wide spectrum of AI/ML functionalities and EO Services to the end-user in a versatile cloud security framework. The advantage is to allow ubiquitous access to EO4EU services

	and data offerings. Furthermore, the Authentication SSO mechanism provides an easy and accelerated in time user data engagement with a checking and secure authentication. In particular, the user authentication and registration of the EO4EU will be based on ASP.NET Core, containing features for managing authentication, authorization, data protection, HTTPS enforcement, app secrets, XSRF/CSRF prevention, and CORS management. Furthermore, a session mechanism will be applied to assist users interact with the EO4EU framework. Currently, the assessment of the cloud security and authentication mechanism that is being designed, developed and integrated into the platform, is being program by the technology specialists in CINECA and ECMWF. In addition, from the technical side of development view, a robust security-authentication-encryption mechanism is being integrated into the EO4EU Software Platform System to allow only authorized entities to access system data and functionality subject to specific arrangements and approval previously granted.
Provided functionalities	An authentication SSO system provides several functionalities to enhance user authentication and streamline access to multiple applications. Here are listed some functionalities that may be provided by the EO4EU authentication SSO: • Centralized Authentication: The SSO Authenticator will allow users to authenticate themselves on the EO4EU platform once using a single set of credentials, such as username and password; gaining access to multiple applications or systems without re-entering credentials for each application. • User Identity Management: The SSO Authenticator may provide user identity management capabilities, allowing administrators to create, modify, and delete user accounts centrally. This may simplify user provisioning and de-provisioning processes. • Seamless Single Sign-On: Once a user logs in through the SSO system, they will be able to access different applications without re-authentication. This eliminates the need for multiple login prompts, saving time and improving user experience. • Password Management: The SSO authenticator may include password management features, such as password reset, password policy enforcement, and password synchronization across different applications. This helps improve security and simplifies password-related tasks for users. • Access Control and Authorization: The SSO Authenticator may enforce access control policies, allowing administrators to define who can access specific applications or resources based on user roles, groups, or other attributes. This helps ensure appropriate authorization and secure application access. • Audit and Logging: The SSO Authenticator may provide logging and auditing capabilities, allowing administrators to monitor user authentication events, track access attempts, and generate reports for compliance or security purposes. • Single Logout: The SSO Authenticator may offer a single logout functionality, allowing users to terminate their session across

	multiple applications simultaneously. This improves security by ensuring that users' sessions are fully terminated when they log out. • Session Management: The SSO Authenticator may include session management capabilities, allowing administrators to monitor and manage user sessions. This can include features such as session timeouts, session termination, and session revocation in case of suspicious activity. • Support for Multiple Authentication Protocols: The SSO Authenticator may support various authentication protocols such as SAML (Security Assertion Markup Language), OpenID Connect, OAuth, and LDAP (Lightweight Directory Access Protocol). This ensures compatibility with a wide range of applications and systems. • User Self-Service: The SSO Authenticator may offer self-service capabilities, allowing users to manage their own profiles, update personal information, reset passwords, and handle other account-related tasks without administrative intervention. • Integration with Multi-Factor Authentication Providers: The SSO Authenticator may integrate with external multi-factor authentication providers, allowing organizations to leverage additional authentication methods such as email verification. • Customization and Branding: The SSO Authenticator developed may offer customization options, allowing to customize the login page, branding, and user interface to align with the EO4EU identity and provide a consistent user experience. It's worth noting that the specific functionalities offered by the authentication SSO may vary depending on the use case or the organisation the user belongs to.
Relation to other components	
Related sequence diagrams sections	4.2, 4.8, 4.9, 4.10, 4.11

Component	**Communication Manager**
Responsible partner	NKUA
Parent Component	-
Description	Communication manager along with the Kafka layers will interact with the other components to provide the means of message dispatch and delivery as unified message bus. The underlying technology that powers and supports the Communication Manager within the Kubernetes ecosystem is the Strimzi operator. Strimzi will provide Apache Kafka instances that will act as the EO4EU message bus and will handle the topic creation and deletion. By interpreting the source and destination, it guarantees precision in message delivery. Specific interfaces will support CRUD operations on the Kafka topics that will be created and deleted, per the workflow demands and requirements. Furthermore, the EO4EU communication manager along with the means of reconfiguring the message queues will have a self-healing and

	rebalancing mechanism that will ensure peak operational efficiency. Based on gathered operational metrics, the message bus will identify and rectify performance bottlenecks ensuring continuous message passage even amid system disturbances while also empowering systems with structured communication avenues. Also, the communication manager will be the responsible component to overview the lifecycle of the needed storage requirements in each workflow. This way communication manager will be handle the CRUD processes of the dynamic storage.
Provided functionalities	• Deal with asynchronous notifications caused by specific platform events. • Handle Publisher/Consumer requests among different components. • Manage Kafka resources per other component requests • CRUD dynamic storage
Relation to other components	ALL components
Related sequence diagrams sections	4.2, 4.3, 4.4, 4.5

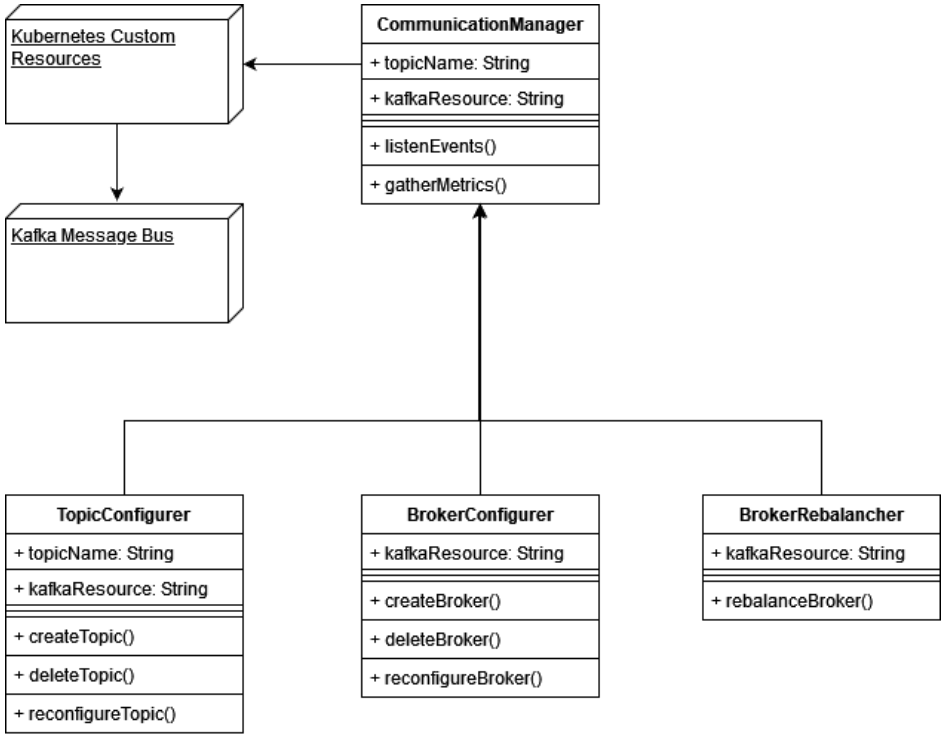

Figure 8 Communication Manager Class Diagram

Component	**DSL Engine**
Responsible partner	NKUA
Parent Component	None
Description	The Domain Specific Language (DSL) Engine's role is to act as a validation and control schema for the System Workflows. A DSL is developed for the needs and characteristics of the Workflow Editor (WFE) named Graph Description Language (GDL). This language is mainly comprised of

	structures containing the information of system nodes, their characteristics, metadata and relation other system nodes. The DSL Engine while being a standalone components is tightly integrated with the WFE through the WFE Auxiliary (AUX) Service. When a workflow is about to be deployed from the WFE ,it is first sent to the DSL Engine for compilation and validation, where in the case of it being valid will get compiled to YAML format and sent back to the AUX Service for deployments to the Systems ,or in case of it not being valid an error report will be sent to the AUX Service.
Provided functionalities	- Workflow compilation and validation.
Relation to other components	Consume and produce messages to the message bus. Read Marketplace Items Authentication SSO Platform Orchestrator
Related sequence diagrams sections	4.7, 4.8

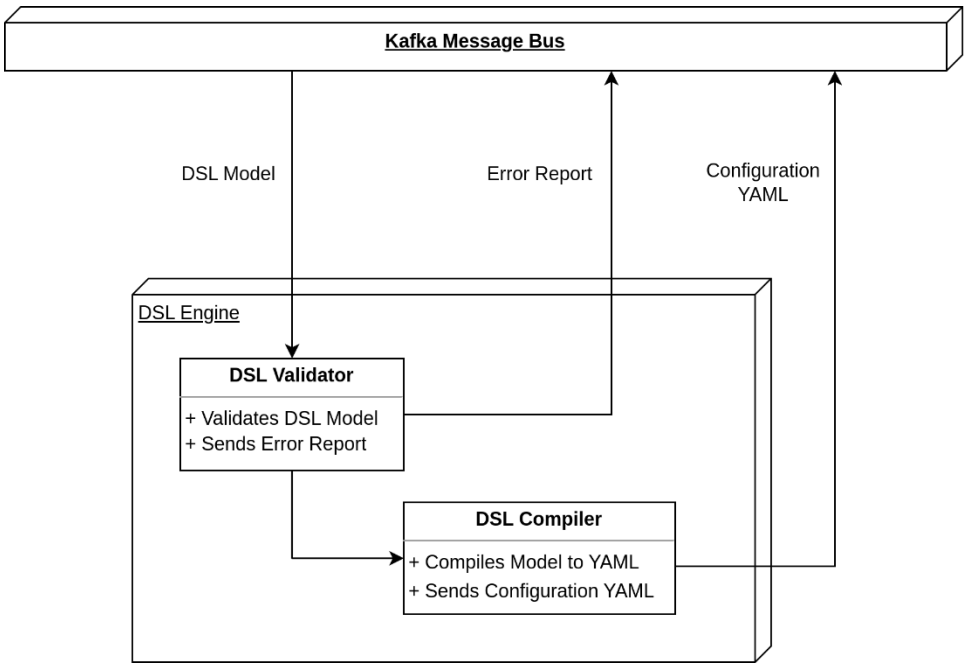

Figure 9 DSL Engine Class Diagram

Component	**Fusion engine**
Responsible partner	NKUA
Parent Component	None
Description	Fusion engine (FE) enables context awareness by combining the data readings and leading to situation awareness. FE has two main functionalities: i) create fusion models and pipelines to provide spatiotemporal aggregation of the data, and ii) execute multiple workflows on parallel coming as requests from the user in a dynamic way. Fusion pipelines are published in AI/Marketplace in order to be accessible to the users in WFE. When a new workflow is created, a request arrives to

	Fusion Proxy in order to create a specific workflow in Kubeflow environment, i.e. create a consumer triggering the start of the execution, initialize the namespace, run the pipeline and then publish the results in data repository and inform next component via Kafka. All computations will be investigated to be performed by using HPC or GPU environments and always in regards to high productivity rated. Each pipeline will be executed in the Kubeflow environment dockerized and serve the next components by producing messages to the EO4EU message bus.
Provided functionalities	- Execute User Workflow - Create pipelines for spatiotemporal processing - Publish pipelines in AI/ML Marketplace - Execute pipelines
Relation to other components	Consume and produce messages to the message bus. Read/ Write data to data buckets Publish models to AI/ML Marketplace
Related sequence diagrams sections	4.6

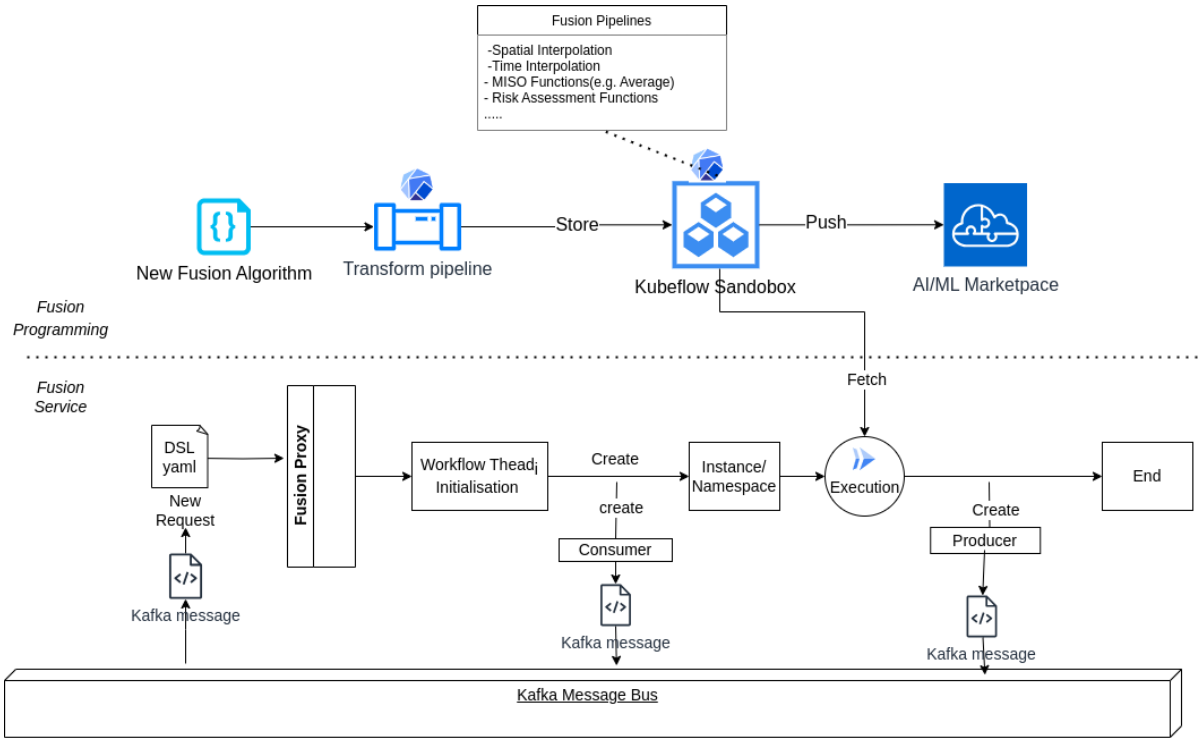

Figure 10 Fusion Engine Class Diagram

Component	**AI/ML Marketplace**
Responsible partner	EBOS
Parent Component	Platform Orchestrator
Description	The AI/ML Marketplace will be defined as an open-source library or database of the available AI/ML processing algorithms, metadata, techniques, data model structures for communicating from one block to the other functional block, that will be developed by each corresponding

	partner within the project. The AI/ML Market Place is not a software component or a software utility, however, it will be a typical library and will be shown as an assistive library list inside the Dashboard that will contain all the AI/ML algorithms and models as these will be developed by each assigned partner in the project. In addition, this library or repository called the AI/ML Marketplace will also contain configuration files, programming code for data transformation or processing, and documentation. All corresponding models will be programmed, developed, tested, prototyped, and evaluated by the corresponding Partner throughout the life cycle of the project, and will be available for automated reusability, but to also enhance the acceleration of the data processing. The AI/ML Marketplace will be defined as a European Open Science Cloud (EOSC)[20] portal, that will give direct access to the users or partners within the project to share and re-use the AI/ML Models for delivering promising innovations, advancements, and services to the end-users. Zenodo platform will be used as a wide-open source for publishing the results and other allowed documents for providing additional data to the research and science community. Furthermore, the usage of relevant platforms such the Acumos[21] and OpenML[22] will advance the full interoperability and reusability of the newly generated data computational models, and will significantly simplify the cross-platform re-usability by leveraging the new AI/ML models that will be developed by each partner. The AI/ML Market Place will be communicating with the Function Execution Engine based on the User Demands and Services, but also based on the System communication infrastructure and needs.
Provided functionalities	AI/ML Models-Algorithms-Techniques, Metadata, Data Models for Processing and Communication from Block to Block, Programming Code, Configuration Files, Documentation.
Relation to other components	
Related sequence diagrams sections	4.9

[20] https://eosc-portal.eu/
[21] https://www.acumos.org/
[22] https://www.openml.org/

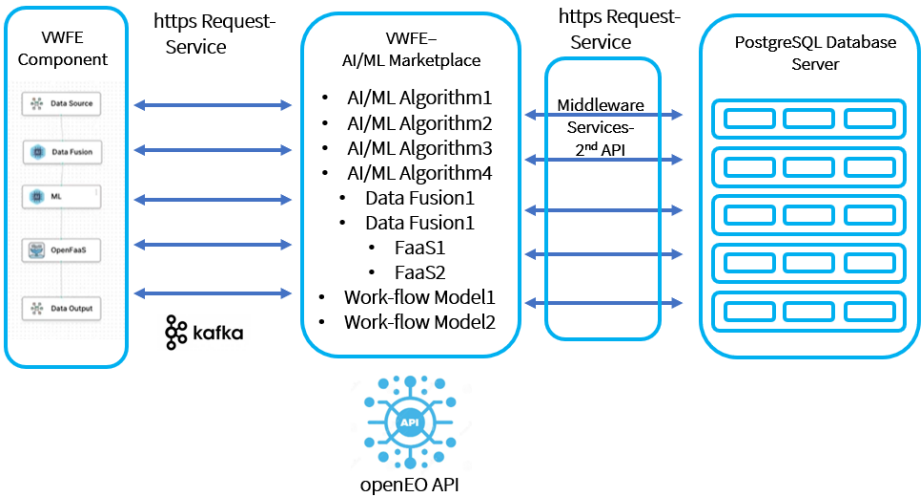

Figure 11 - AI/ML Marketplace - Block Diagram

3.4 ML Tier

Component	**Machine Learning Model**
Responsible partner	HESSO
Parent Component	None
Description	A machine learning model is a mathematical representation or algorithm that is trained on historical data to make predictions or decisions without being explicitly programmed. It is designed to learn patterns and relationships within the data and generalize that knowledge to new, unseen data. The model's key components include feature engineering, where relevant features are selected or created to enhance model performance; data loading, where appropriately pre-processed data are fed to the model; model training, where the model learns from the data to optimize its parameters; and model evaluation, where the model's performance is assessed using various metrics. Through the MLOps lens calibrated onto the project's goal, trained machine learning models will be integrated within the platform and usable by the user notably through inference tasks executed via the inference server embedded within the platform. Once developed within the course of the project, a developed and trained machine learning model is deployed to the inference server (described below) and is then ready for use.
Provided functionalities	- Ability to access trained models for inference - Documentation associated with every provided model, including a detailed description of the input types that each model can accept if the user desires to use the model for inference in a downstream task. Importantly, their shape is also given. - Ability to download models locally for local execution
Relation to other components	- Data Fusion Component - Machine Learning Inference Server
Related sequence diagrams sections	4.9

Component	Machine Learning Inference Server
Responsible partner	HESSO
Parent Component	None
Description	The inference server component consists of two main sub-components. The first sub-component is the core inference server, which includes a deployment of KServe. KServe is a standard Model Inference Platform on Kubernetes, designed for facilitating highly scalable use cases. It provides a performant, standardized inference protocol across multiple machine learning frameworks and supports modern serverless inference workloads with Autoscaling, including the ability to Scale to Zero with or without GPUs. Models are uploaded to storage, currently S3, in a format compatible with a supported inference server backend. Currently, we use the Open Neural Network Exchange (ONNX) format, an open-source project that defines an interoperable format for machine learning models. We use NVIDIA Triton as the backend server, which supports multiple model formats, including ONNX. KServe provides a RESTful API for inference, using JSON as the query language. The second sub-component interfaces with the EO4EU platform. This sub-component receives messages from the platform, reads data in various formats, and converts them into the format required by the internal Inference server. It also converts the results into a format expected by the platform. This component includes a converter from Kafka to HTTP and a webserver code which interacts with the inference server.
Provided functionalities	- Inference - Data Loading/Saving
Relation to other components	- Deploys ML Models - Listens to DSL - Data Type dependency with Fusion for reading - Data Type dependency with users components for saving.
Related sequence diagrams sections	4.9

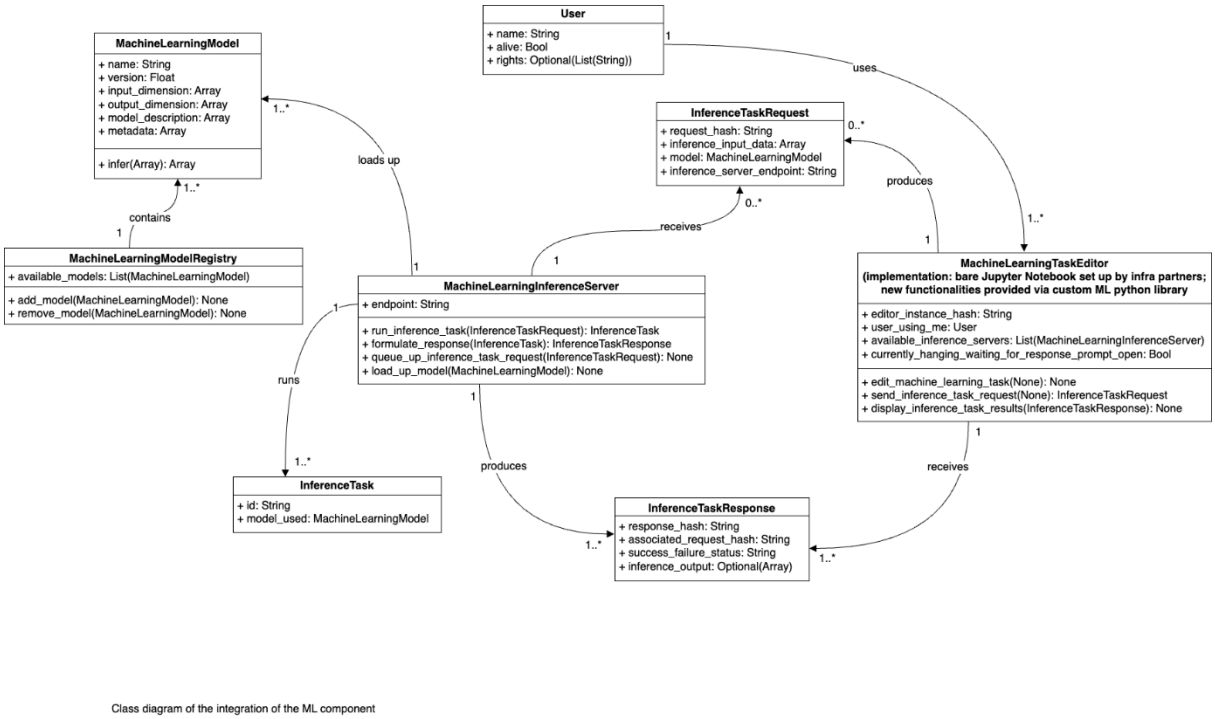

Figure 12 ML class diagram

3.5 Customer facing services

Component	**Dashboard - Data Analytics Visualization (DDAV)**
Responsible partner	EBOS
Parent Component	None
Description	User-friendly Data Analytics Visualisation tools will be designed-developed-integrated for decision-making and policy-maker use**.** Advanced data analytics visualizations will be developed for improved learning and evidence-based interpretations of environmental observations. Visualization and exploration of high-dimensional vector data will be interacting with AI/ML modules developed in this project, in order to generate the intelligent data desired by the user/users. Furthermore, a novel situated analytics approach will be developed, that will use augmented reality as an immersive analytical tool for exploring and analysing information in the physical world and adjust it to handle EO data. The visualization tool will be designed and developed for both scientific and EU civilian users who would like to get real time data analytics. The tool will provide an analytical scientific visualization to assist computational scientists and to better understand their research data. New scientific findings will be possible by visually studying the output of their computational models. The EO4EU will introduce a hybrid visualization mechanism, based on the convergence between exploratory and explanatory visualization. Smart triggering events of weather-climate-environment changes will be developed based on intelligent ML algorithms and visualization methods. These smart visualization methods and AI-ML algorithms that will be developed, will generate an intelligent awareness among citizens about urgent environmental changes, while the methods of explanatory visualization can enrich the exploration of data.

Provided functionalities	1. Real-time data analytics and interpretation of environment observations 2. Real-time mapping and interactions 3. Smart Search Engine based on Text or Annotation – Select Data - Smart Category-Type-Parameters Search Engine of the searched item 4. Dashboard personalization 5. Web XR/VR Visualization methods 6. Data Analytics based on statistical metrics
Relation to other components	1. Machine Learning Model 2. AI/ML Marketplace 3. API
Related sequence diagrams sections	4.10

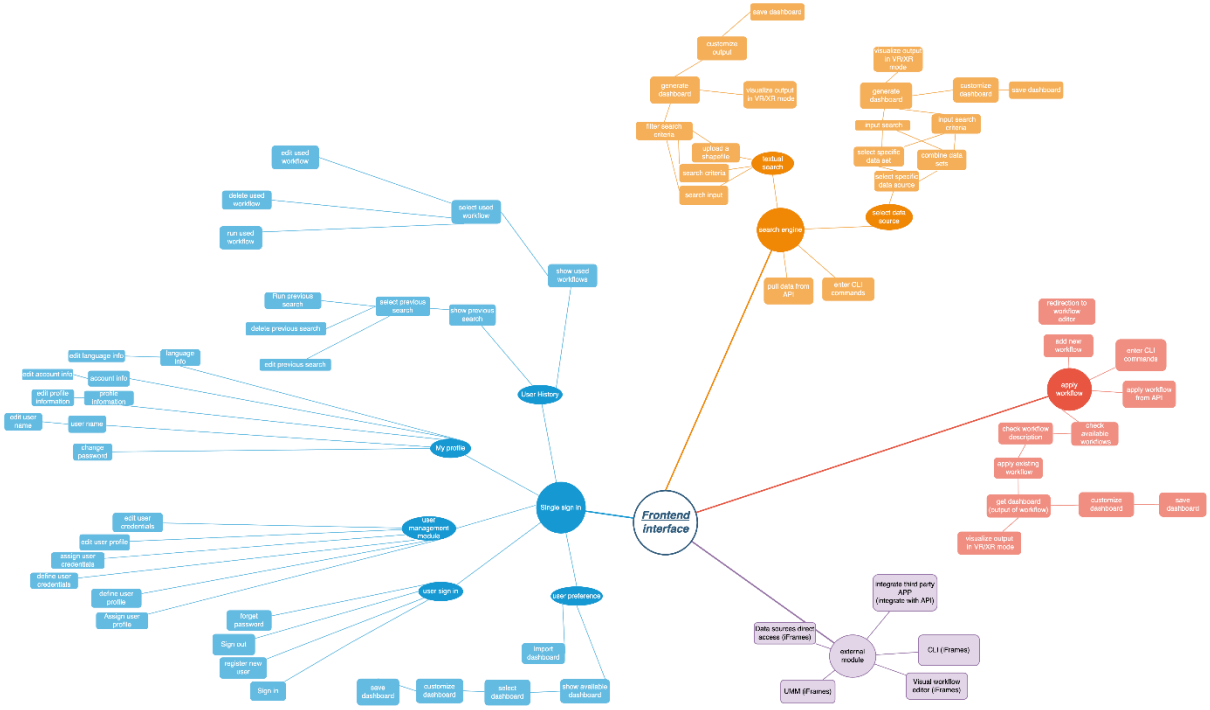

Figure 13 Dashboard - Data Analytics Visualization Class Diagram

Component	**EO4EU API**
Responsible partner	EBOS
Parent Component	None
Description	Application Programming Interface (API) will be designed and developed as a control interface management software tool for establishing communication with the user and the various EO4EU software components, but to also generate smart data interfaces and communications among the corresponding software modules/components. The developed API software module will define the user requirements and request services. In parallel the API will be designed based on the various system's components specs and

	requirements needed to serve and generate the requested data of the user. The API here will enable smart communication between EO4EU applications or software functions such as Web XR/VR, Dashboard, Visualization Data Analytics, CLI, etc. The API will play a significant role in a remote communication server that will be receiving requests from the user or users and sending EO Data in the form or type of the request. In addition, the API will be communicating with the Kubernetes Platform Software Management and other Control Software modules that will be connecting to retrieve real-time and continuous time from the requested Cluster or available servers or software resources to obtain processing EO sources' data.
Provided functionalities	• User (Internal/External) System Access • Communication Interface with Clusters • Communication Interface with Kubernetes Platform Software Management System • Communication Interface with other Software Control Modules • Communication Interface with 3rd Party UI • Communication Interface with UMM • Communication Interface with Work-Flow Editor
Relation to other components	1. Data Analytics Visualization 2. XR System 3. EO4EU Platform System 4. User Management Model UMM 5. Third-Party/User Interface UI 6. Visual Work Flow Editor
Related sequence diagrams sections	4.10

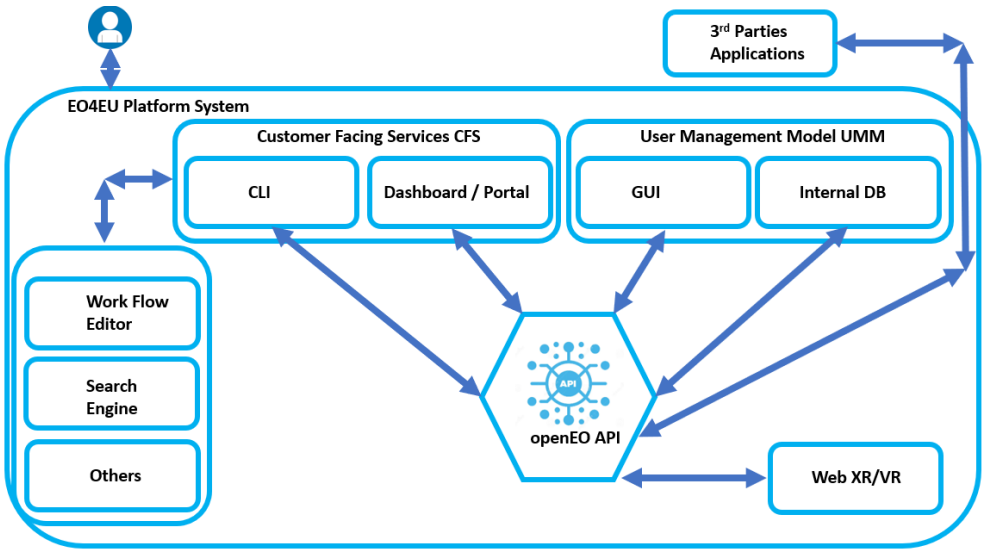

Figure 14 Customer Facing Services – Abstract System Architecture

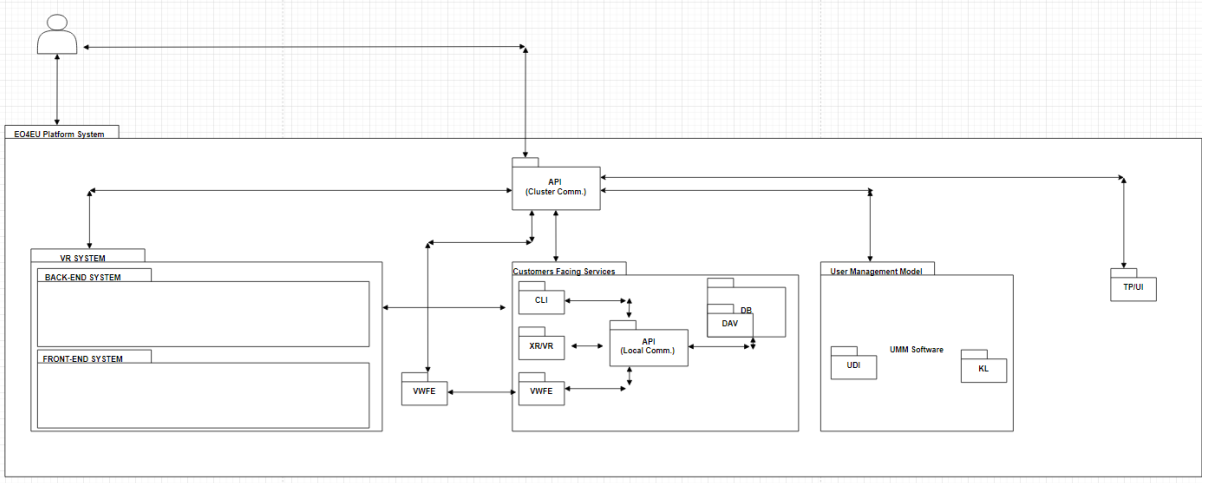

Figure 15 Customer facing services – Analytical System Architecture

Component	**XR System**
Responsible partner	CINECA
Parent Component	None
Description	The XR system will provide a web-based XR interface for visualizing and exploring the EO data in a more immersive way. The XR system will include two different applications. An augmented reality (AR) interface will enable the visualization and the possibility to analyse the EO data in the physical world taking advantage of the geolocation system of the device where the user will run the application. A virtual reality (VR) interface will enable the visualization of the EO data on 3D virtual scenes on hardware designed for supporting VR applications, such as cardboard and VR head-mounted devices. In the case of the VR head-mounted devices the user will have the opportunity to move within the 3D scene and interact with the environment using the device hand controllers. Both the applications will be optimized to handle EO data and will make available to the final user tools to interact with the 3D models and the EO data themselves to offer an improved way to explore and understand the environmental observations. The XR interactions and visualizations will be realized through the Web XR device API, a group of standards which are used together to support 3D scenes to hardware designed for presenting virtual worlds. The XR system will communicate with the other EO4EU applications to obtain processing EO sources' data through the provided API.
Provided functionalities	- VR interface for EO data visualization - AR interface for EO data visualization - Movement within the 3D scene - Interactions with both the EO data and the 3D models in the XR environment
Relation to other components	1. API 2. User Management Model UMM
Related sequence diagrams sections	4.10

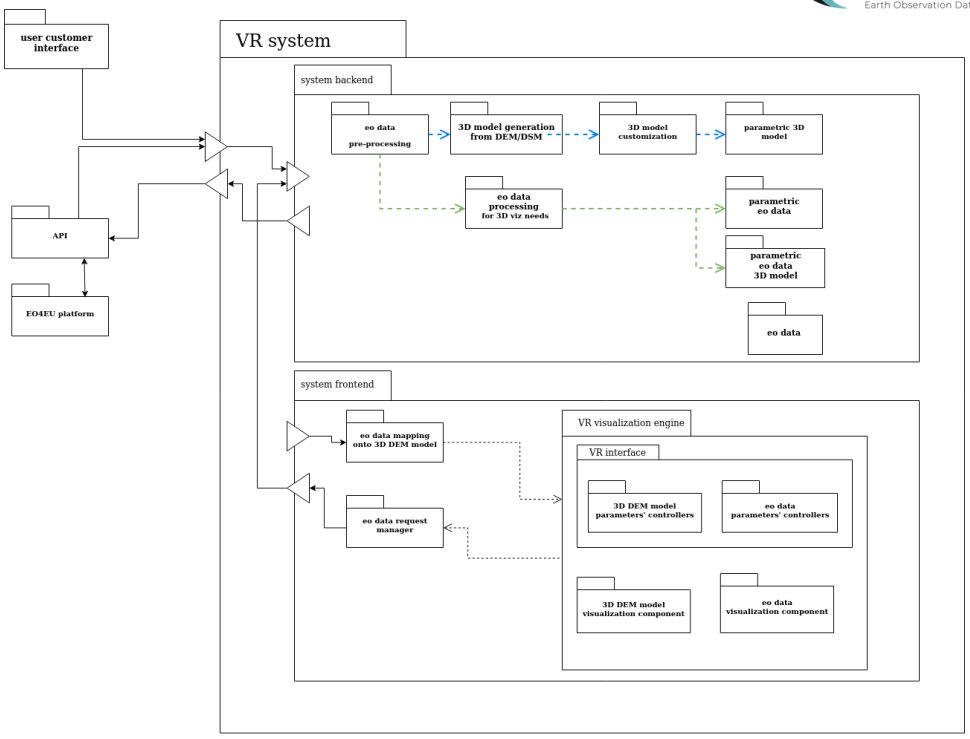

Figure 16 VR system - Analytical Block Diagram

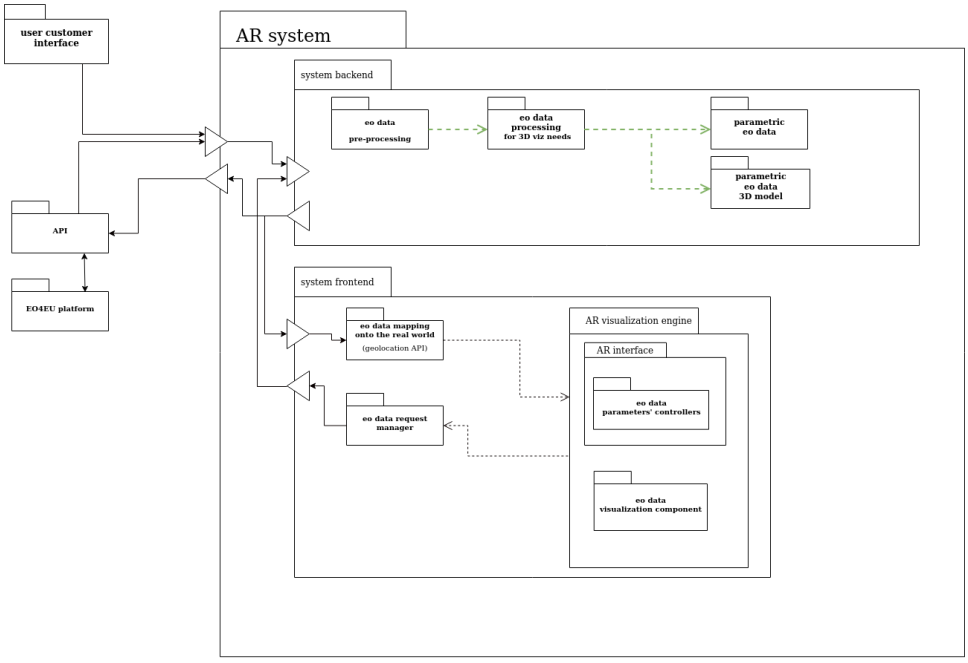

Figure 17 AR system (on continuous update) - Analytical Block Diagram

Component	User Management Model - UMM
Responsible partner	IVI

Parent Component	
Description	Management of login, SSO, user preferences and user history data.
Provided functionalities	User related functionalities, see sub-components
Relation to other components	Consists of - Key Cloak (Identity and Access Management) - User Data Management
Related sequence diagrams sections	Available for any Use Case (General Profile)

Component	**Identity and Access Management (Key cloak)**
Responsible partner	IVI
Parent Component	UMM
Description	Usage of the open-source tool "Key Cloak[23]". It is an identity and access management tool that provides user federation, strong authentication, user management, fine-grained authorization, single-sign-on and connection to external identity providers
Provided functionalities	- user and group management - management of roles of users and groups - authorization for resources based on user roles using scops, policies and permissions definitions - single-sign-on - integrate external identity providers (e.g., login via a Google account)
Relation to other components	
Related sequence diagrams sections	Available for any Use Case (General Profile)

Component	**User Data Service**
Responsible partner	IVI
Parent Component	UMM
Description	Stores additional data per user, like preferences or a history of actions
Provided functionalities	- Store preferences per topic and user - Store list of actions (history) per user
Relation to other components	- Any other component that needs to store preferences or user actions.
Related sequence diagrams sections	Available for any Use Case (General Profile)

[23] https://www.keycloak.org/

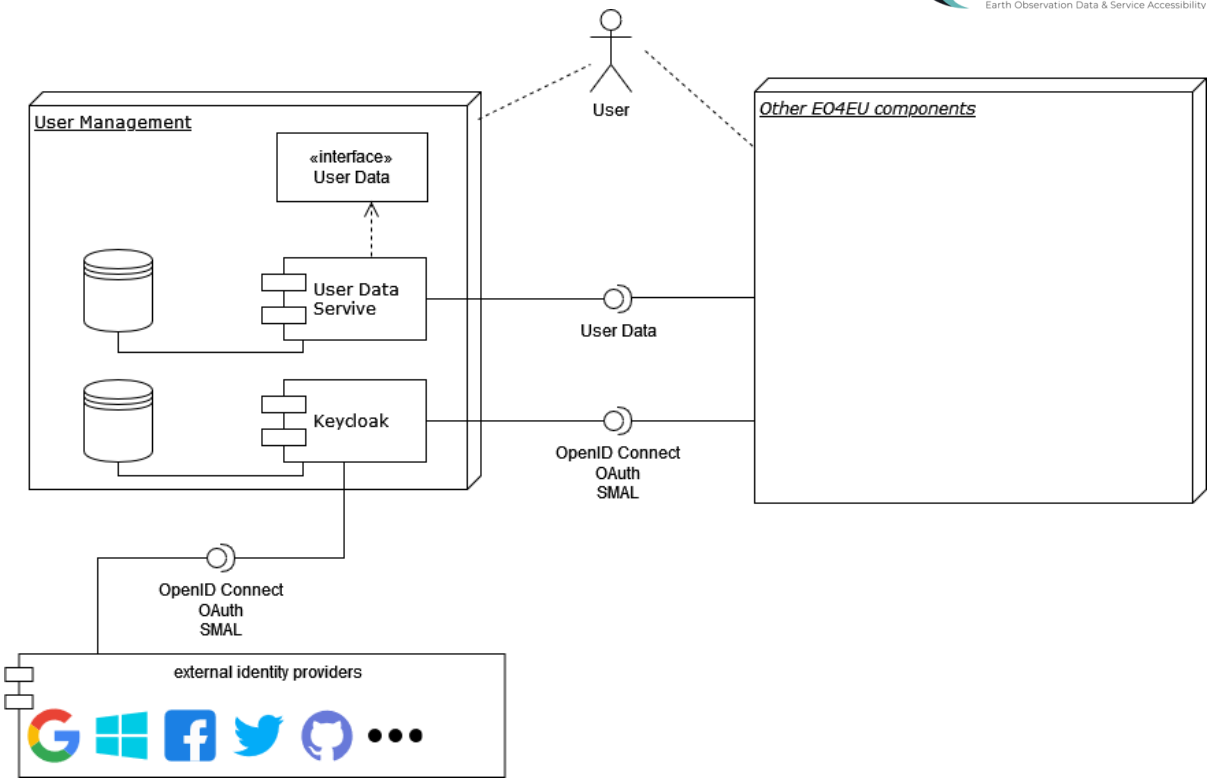

Figure 18 Overview of the User Management Component

Component	**Workflow Editor (WFE)**
Responsible partner	NKUA
Parent Component	None
Description	The Workflow Editor, along with its graphical environment, serves as the central point for describing workflows. The purpose of the tool is to parameterize and create personalized data processing workflows both in terms of the components that the user wishes to include in the process and in the individual parameters that need to be entered at the component level. The design of a workflow takes place in a user-friendly environment and follows the design logic of a single-path graph. In the graph, the components are presented as nodes and the edges represent the flow of information from the output to the input of the next node. The workflow editor as a user-friendly graphical environment, empowers users to effortlessly establish connections between components of the experiment using graphical elements. This component provides users with the capability to parameterize the nodes within the workflow. It also enables the description of functions that will be utilized to feed the FaaS subsystem. Moreover, users have the flexibility to incorporate custom implementations in the form of Docker images. The GUI translates the user's created opaque graph into the DGL (Definable Graph Language) through collaboration with the Core service of the DSL Engine. This transformation process results in the generation of the necessary configuration. Subsequently, the configuration is transmitted through the central Message Bus to the platform orchestrator

	subsystem for validation. Upon successful validation, the platform orchestrator subsystem commits the requisite computing resources and initiates the execution process. Lastly, the component provides users with information about the state of the workflow. The WFE also offers essential actions such as pausing, editing, and deleting the workflow, granting users the ability to interact with and control the workflow according to their needs.
Provided functionalities	Publish/Edit/Stop/Delete EO4EU Workflows Interact with Marketplace and gives the ability to add components to workflow Configuration Data Analytics Visualization Configuration for AR/AX Configuration for ML Component Create/Edit and Validate FaaS for Python3, C#, nodeJS18 and PHP8.2 Languages
Relation to other components	Consume and produce messages to the message bus. Read Marketplace Items Authentication SSO Platform Orchestrator DSL Engine
Related sequence diagrams sections	4.7, 4.8

3.6 Technical Requirements mapping

The following table shows an overview of all the requirements, i.e. functional or non-functional, that were gathered, analysed and applied during the design and the development project lifecycle. The requirements coming from DoW were thoroughly analysed in deliverable D2.2. In D2.2 a questionnaire to the end users was presented and the output of the requirements is consolidated in the following table.
The concatenation of requirements is a combination Tier-Component-Numbering. The mapping of the tiers is the following

- DATA- Data tier
- SS – Systems tier combing IaaS and PaaS
- ML – Machine Learning tier
- CF – Customer Facing Tier
- GEN – Generic requirements related to infrastructure or to data

For example, CF-GUI-06 is a requirement applied in Customer Facing tier and more specifically in Dashboard/Graphical User Interface component.
The main columns of the table have the following columns

- ID: unique numbering
- Component: in which component are applied
- Title: Title and Description
- Type: Functional or Non-Functional requirement
- Priority: priority in the project Low, Medium and High
- Source: the source generating the requirement, i.e. External users, DoW and Consortium partners

The priorities of the requirements as HIGH, MEDIUM, LOW were defined in D2.2.

Table 1 – Functional and Non-functional Requirements

ID	Component	Title & Description	Type	Priority	Source	Comments
CF-R-01	Dashboard - Data Analytics Visualization, EO4EU API	The EO4EU should allow users to download final and intermediate results of their workflows	FUNC	**HIGH**	External	

ID	Component	Title & Description	Type	Priority	Source	Comments
CF-DDDAV-01	Dashboard - Data Analytics Visualization	Usability and operability: The EO4EU platform shall provide a user-friendly interface which can be operated in an easy and intuitive way.	FUNC	**HIGH**	DoW	
CF-DDAV-02	Dashboard - Data Analytics Visualization, WFE	Adaptability: The EO4EU shall provide a fully configurable dashboard that will be able to personalize the viewable components depending on the user's needs.	FUNC	**MEDIUM**	DoW	
CF-DDAV-03	Dashboard - Data Analytics Visualization	The EO4EU platform's dashboard shall be developed to offer robustness, high performance, cross-platform support, improved quality and easy maintenance.	FUNC	**MEDIUM**	DoW	
CF-DDAV-04	EO4EU API	The EO4EU platform's dashboard shall rely on open standards and shall enable an easy and secure communication and data transfer between the interfacing and the backend system.	FUNC	**HIGH**	DoW	
CF-DDAV-05	Dashboard - Data Analytics Visualization	The EO4EU platform shall provide a multi-layered GUI for visual analytics.	FUNC	**HIGH**	DoW	
CF-DDAV-06	Dashboard - Data Analytics Visualization, EO4EU API	The EO4EU platform shall provide a Command Line Interface (CLI).	FUNC	**HIGH**	DoW	
CF-DDAV-07	XR System	The EO4EU platform shall provide an XR web-based application interface.	FUNC	**HIGH**	DoW	

ID	Component	Title & Description	Type	Priority	Source	Comments
CF-DDAV-08	Dashboard - Data Analytics Visualization, EO4EU API	~~The EO4EU platform shall be secure providing protection of the authorization information from unauthorized access, and the maintenance of user confidentiality.~~	FUNC	**MEDIUM**	External	Covered by CF-DDAV-04
CF-DDAV-09	Dashboard - Data Analytics Visualization, EO4EU API	~~The EO4EU platform shall provide an administrator console offering the capability to manage users / applications / services, authentication, and authorization.~~	FUNC	**MEDIUM**	External	Covered by CF-DDAV-01, CF-DDAV-02, CF-DDAV-03, CF-DDAV-04
CF-DDAV-10	Dashboard - Data Analytics Visualization, AI/ML marketplace	Alerting: sending information to either specific people registering to the service, or to the mass of people based on event triggering	FUNC	**MEDIUM**	External	The alerting functionality and event registering is related to error handling and events created by user workflow
CF-DDAV-11	WFE	Aggregator: to set a specific aggregator, listing some parameters, thresholds, that once reached will aggregate information from any data source	FUNC	**MEDIUM**	External	WFE will cover parameters of the workflow
~~CF-DDAV-12~~	~~Graphical user interface~~	~~Manage/evaluate specific event or situation, based on specific criteria, users can suggest solution, or propose decisions that could be shared with other people~~	~~FUNC~~	~~**MEDIUM**~~	~~External~~	N/A
CF-DDAV-13	AI/ML marketplace	Share of specific event, or decision that could be useful for other people/organization	FUNC	**MEDIUM**	External	Events are coming from a workflow added in AI/ML and a user can re-use it

ID	Component	Title & Description	Type	Priority	Source	Comments
CF-DDAV-14	Dashboard - Data Analytics Visualization, WFE, KG	GUI to include a search engine which should be the frontend interface for the semantic annotation being done in the background	FUNC	**HIGH**	External	
CF-DDAV-15	Dashboard - Data Analytics Visualization, WFE, KG	Search engine functionality should have 2 levels of granularity, 1st is the dataset which is semantically annotated, 2nd level is the product, feature/product falling on a specific dataset (not semantically, which entails to give all the products and features under each dataset)	FUNC	**MEDIUM**	External	
CF-DDAV-16	WFE	Workflow editor module, which is used by the user to run a specific workflow	FUNC	**HIGH**	External	
CF- DDAV -17	Dashboard - Data Analytics Visualization, WFE	GUI should have an access to the workflow editor UI, in which the user can create new workflow	FUNC	**HIGH**	External	
CF- DDAV -18	Dashboard - Data Analytics, KG	The user interface of the EO4EU has to allow the user to search for the available datasets	FUNC	**HIGH**	External	
CF- DDAV -19	Dashboard - Data Analytics Visualization, WFE, FaaS	The EO4EU platform should provide a ~~adequate~~ code editor ~~(e.g., jupyter lab)~~ to allow users to run their custom code on the platform	FUNC	**HIGH**	External	
CF-MP-01	AI/ML MarketPlace	Developed, tested and evaluated models will be publicized in a public repository.	FUNC	**HIGH**	DoW	
CF-MP-02	AI/ML MarketPlace, OpenEO API	Other FAIR, open repositories will also be potential places of publishing	FUNC	**MEDIUM**	DoW	

ID	Component	Title & Description	Type	Priority	Source	Comments
CF-UM-01	UMM, identity and Access Management	Create user account	FUNC	HIGH	Consortium	
CF-UM-02	UMM, identity and Access Management	User registration service in Keyclock	FUNC	HIGH	Consortium	
CF-UM-03	UMM, identity and Access Management, User Data Service	Integration of external identity providers	FUNC	MEDIUM	Consortium	
CF-UM-04	Identity and Access Management	Single-Sign-On	FUNC	HIGH	Consortium	
CF-UM-05	Identity and Access Management	Mapping of roles to users and groups	FUNC	HIGH	Consortium	
CF-UM-06	UMM, identity and Access Management, User Data Service	Define authorization for resources	FUNC	MEDIUM	Consortium	
CF-UM-07	UMM, identity and Access Management, User Data Service	Storing of user settings/preferences	FUNC	HIGH	Consortium	
CF-UM-08	UMM, identity and Access Management, User Data Service	History of user actions	FUNC	MEDIUM	Consortium	
CF-XR-01	Dashboard - Data Analytics Visualization, XR/VR	The EO4EU to allow data search using CLI and XR/VR	FUNC	LOW	External	
CF-XR-02	Dashboard - Data Analytics Visualization	The EO4EU should allow visualization of analysis results using XR/VR	FUNC	LOW	External	
GEN-COMP-01	IaaS,PaaS	The EO4EU platform should provide adequate storage needs for the use	OTHER	MEDIUM	External	

ID	Component	Title & Description	Type	Priority	Source	Comments
		cases to process/store some data (average of 500 GB)				
GEN-COMP-02	IaaS,PaaS	The EO4EU platform should allow storing data of various formats such as GeoTiff, NetCDF, GRIB, Shapefiles, CSVs, GeoJSON, etc.	OTHER	**HIGH**	External	
GEN-COMP-03	IaaS,PaaS	The EO4EU should include GPU processing capabilities	OTHER	**HIGH**	External	
GEN-COMP-04	IaaS,PaaS	The EO4EU should provide an average of 30-50GB RAM for processing purposes of the users	OTHER	**HIGH**	External	
GEN-DATA-01	Dataset metadata	Add geographical extent to the dataset metadata	OTHER	**HIGH**	External	Different service providers provides the geographical extent of their datasets in different types
GEN-DATA-02	Dataset metadata	Add temporal coverage to the dataset metadata	OTHER	**HIGH**	External	Different service providers, provides the temporal extent of each in different types
GEN-DATA-03	Dataset metadata	Add the dataset URL to the metadata	OTHER	**MEDIUM**	External	The metadata of a dataset is extracted through dedicated parsers from the URL of the dataset itself. In the cases where dedicated APIs provide access to the metadata of a dataset, the URL of the metadata collection can be provided

ID	Component	Title & Description	Type	Priority	Source	Comments
GEN-DATA-04	Dataset metadata	Add data format to the dataset metadata	OTHER	MEDIUM	External	All the metadata is being provided by the service providers
GEN-DATA-05	Dataset metadata	Add processing level to the dataset metadata	OTHER	MEDIUM	External	Processing level is being provided by the service provider. Not all the service providers provide the processing level
GEN-DATA-06	Dataset metadata	Add the acquisition info to the dataset metadata (e.g., instrument/sensor)	OTHER	MEDIUM	External	Not all the service providers provide this information. Specifically, this information is not supported to all datasets by ADS/CDS service provider
GEN-DATA-07	Dataset metadata	Add keywords to the dataset metadata (e.g., application domain)	OTHER	MEDIUM	External	Not all the service providers provide this information. Specifically, this information is not supported to all datasets by ADS/CDS service provider
GEN-DATA-08	Dataset metadata	Add DOI to the dataset metadata	OTHER	LOW	External	Same with 08
GEN-DATA-09	Dataset metadata	Dataset metadata should include information about FAIR complaint, licence type	OTHER	LOW	External	License type in ADS/ CDS is included, otherwise according to the openEO api we should add “proprietary”
GEN-DATA-10	Dataset metadata	Dataset metadata should include information about the data quality	OTHER	LOW	External	Not all the service providers provide this information.

ID	Component	Title & Description	Type	Priority	Source	Comments
GEN-DATA-11	Dataset metadata	The EO4EU platform should provide two-level metadata system, a collection (dataset) and granule (file) levels to improve discoverability and accessibility of the data	OTHER	**HIGH**	External	Datasets which are accessed through the KG are based on the OPEN APIS and the capacity supported by each one, while dedicated custom parsers have been developed to extract extra metadata provided by the service providers through open access license
GEN-DATA-12	Data source	The datasets provided by the platform should have global coverage if applicable such as the free access satellite images	DATA	**HIGH**	External	The KG provides access to already established and open source EO datasets. The geographical coverage is dependent to the service providers.
GEN-DATA-13	Data source	The EO4EU platform should provide access to Sentinel data products	DATA	**MEDIUM**	External	KG will provide access to Sentinel through ADAM platform
~~GEN-DATA-14~~	~~Data source~~	~~The EO4EU platform should provide access to Digital Elevation Model datasets~~	~~DATA~~	**MEDIUM**	~~External~~	No access
~~GEN-DATA-15~~	~~Data source~~	~~The EO4EU platform should provicde access to Landsat collections~~	~~DATA~~	~~**MEDIUM**~~	~~External~~	N/A ~~KG provides access to the extracted datasets as per D2.1. Additional datasets included in the ADS, CDS, ADAM services can be included.n/~~

ID	Component	Title & Description	Type	Priority	Source	Comments
GEN-DATA-16	Data source	The EO4EU platform should provide access to ERA-5 land data	DATA	**MEDIUM**	External	KG provides access to the collections of CDS service.
GEN-DATA-17	Data source	The EO4EU platform should provide access to ERA-5 data	DATA	**MEDIUM**	External	KG provides access to the collections of CDS service.
GEN-DATA-18	Data source	The EO4EU platform should provide access to Copernicus Services datasets	DATA	**MEDIUM**	External	KG provides access to the ADS, CDS, LAND, MARINE copernicus services
~~GEN-DATA-19~~	~~Data source~~	~~The EO4EU platform should provide access to MODIS collections~~	~~DATA~~	~~**MEDIUM**~~	~~External~~	KG provides access to the extracted datasets as per D2.1.
~~GEN-DATA-20~~	~~Data source~~	~~The EO4EU platform should provide access to commercial VHR data (Copernicus)~~	~~DATA~~	~~**MEDIUM**~~	~~External~~	KG provides access to the extracted datasets as per D2.1.
GEN-DATA-21	Data source	The EO4EU platform should provide access of archive data for long time periods	DATA	**HIGH**	External	Supported as per the provisions of each data provider
GEN-DATA-22	Data source	The EO4EU platform should allow downloading specific features from a dataset if the entire product is not needed by the user (e.g., specific bands in S2 or polarization is S1)	FUNC	**HIGH**	External	KG is regenerating API calls to download data from the service provider or provide the download link supported by the service provider. The access t to specific portions of data is dependent to the service provider.
~~GEN-DATA-23~~	~~Data search~~	~~The EO4EU should allow searching data using classic criteria such as time range, sensor, processing level, coverage, cloud cover, etc~~	~~FUNC~~	~~**HIGH**~~	~~External~~	N/A

ID	Component	Title & Description	Type	Priority	Source	Comments
GEN-DATA-24	Data search	The EO4EU should allow searching data using natural language	FUNC	**MEDIUM**	External	
~~GEN-DATA-25~~	~~Data search~~	~~The EO4EU platform should allow searching using image similarity~~	~~FUNC~~	~~**LOW**~~	~~External~~	N/A
KG-01	KG	To augment the FAIRness of EO data knowledge	FUNC	**HIGH**	DoW	
KG-02	KG	To support a sophisticated representation of EO data entities and their dynamic	FUNC	**MEDIUM**	DoW	
KG-03	KG	To semantically relate different resources and descriptions	FUNC	**MEDIUM**	DoW	
KG-04	KG	To investigate and map data and services transparently and efficiently	FUNC	**MEDIUM**	DoW	
KG-05	KG	To merge descriptions/ontologies by expressing semantic information	FUNC	**MEDIUM**	DoW	
KG-06	KG	To conduct queries to discover EO resources that match users' requirements	FUNC	**HIGH**	DoW	
ML-01	ML	A self-supervised learning model based on contrastive learning, utilizing the vast volume of unlabelled data, by learning robust representation that will facilitate downstream tasks minimizing their labeled data budget requirement will be developed	FUNC	**HIGH**	DoW	
ML-02	ML	A generic ML pipeline will be designed and developed.The Generic ML pipeline of EO4EU will enable the learning of a	FUNC	**HIGH**	DoW	

ID	Component	Title & Description	Type	Priority	Source	Comments
		robust and transferable representation of the input data in a latent space, in an unsupervised way.				
ML-03	ML	Machine learning models will facilitate the volume reductions of the data that will be transferred.	FUNC	**HIGH**	DoW	
ML-04	ML	An encoder-decoder pair will serve the compression process in our platform.	FUNC	**HIGH**	DoW	
ML-05	ML	The final version of all developed models will be shaped in the form that will allow their publication in the EOSC	FUNC	**HIGH**	DoW	
ML-06	ML	Deep learning models following a self-supervised learning approach, and in particular contrastive learning will be developed, tuned, evaluated and optimized.	FUNC	**HIGH**	DoW	
ML-07	ML	Deep learning models will exploit HPC infrastructure for the computational aspects.	FUNC	**MEDIUM**	DoW	
ML-08	ML	The models developed and trained will be available for inference use through a inference server service deployed on the platform.	FUNC	**HIGH**	Consortium	
ML-09	ML	The inference service will be set so that it interfaces (receives and sends data) through the agreed upon, unified, message bus	FUNC	**HIGH**	Consortium	
ML-10	ML	The inference service will handle load balancing.	FUNC	**HIGH**	Consortium	

ID	Component	Title & Description	Type	Priority	Source	Comments
SS-API-01	EO4EU API	The EO4EU platform has to expose an API RESTful with JWT authentication	FUNC	**MEDIUM**	Consortium	
SS-API-02	EO4EU API	The EO4EU platform should be compliant with the normative as CAP V1.2 Message Producer&Consumer	FUNC	**MEDIUM**	Consortium	
SS-API-03	EO4EU API	The EO4EU platform should be able to reply with a JSON or CAP message	FUNC	**MEDIUM**	Consortium	
SS-API-04	EO4EU API	The API authentication should be compatible with the current trends (e.g., access token, HTTPs basic auth, Oauth with openid)	FUNC	**HIGH**	External	
~~SS-API-04~~	~~EO4EU API~~	~~API should be able to authenticate user through the integration between UMM, Frontend and API~~	~~FUNC~~	~~**HIGH**~~	~~External~~	CF-UM-01, 02,03,04,05,06,07,08
~~SS-API-05~~	~~EO4EU API~~	~~API should be able to upload a workflow once created within the workflow editor~~	~~FUNC~~	~~**HIGH**~~	~~External~~	Covered by CF-MP-01
~~SS-API-06~~	~~API~~	~~API should be able to execute any available workflow~~	~~FUNC~~	~~**HIGH**~~	~~External~~	Covered by CF-GUI-16
SS-API-07	API	API should be able to get data from the EO4EU system and send it to the visual data analytics	FUNC	**HIGH**	External	
SS-API-08	API	The EO4EU platform has to expose a REST API to search the available data	FUNC	**HIGH**	External	

ID	Component	Title & Description	Type	Priority	Source	Comments
SS-API-09	API	The EO4EU platform capabilties should be exposed via API/CLI to ease the integration with user's existing applications and workflows	FUNC	HIGH	External	
SS-DDAV-01	Dashboard - Data Analytics Visualization	The user select a use case from the preconfigured list of use cases	FUNC	HIGH	Consortium	
SS-DDAV-02	Dashboard - Data Analytics Visualization	Data Analitics Visualization analyze data related to the use case selected	FUNC	MEDIUM	Consortium	
SS-DDAV-03	Dashboard - Data Analytics Visualization	The user can create a dashboard starting from a list of preconfigured charts	FUNC	MEDIUM	Consortium	
SS-DDAV-04	Dashboard - Data Analytics Visualization	The user save the configured dashboards for future visualization	FUNC	MEDIUM	Consortium	
SS-DDAV-05	Dashboard - Data Analytics Visualization	Data analytics Visualization show the list of dashborads previously configured and saved	FUNC	MEDIUM	Consortium	
SS-DDAV-06	Dashboard - Data Analytics Visualization	The user visualize a pervious saved dashboard and edit it for a new save	FUNC	MEDIUM	Consortium	
SS-DDAV-07	Dashboard - Data Analytics Visualization	The user visualize the list of saved dashboard after the choise of the use case	FUNC	MEDIUM	Consortium	
SS-DDAV-08	Dashboard - Data Analytics Visualization	Thw User customize dashboard applying filters	FUNC	MEDIUM	Consortium	
SS-DDAV-09	Dashboard - Data Analytics Visualization	The Data Analytics Visualization visualize data previously processed by the Machine learning algorithms of the EO4EU platform (resultset)	FUNC	HIGH	Consortium	

ID	Component	Title & Description	Type	Priority	Source	Comments
SS-DDAV-10	Dashboard - Data Analytics Visualization	The resultset has to be previously stored on Elastic search	FUNC	HIGH	Consortium	
SS-DDAV-11	Dashboard - Data Analytics Visualization	The result set data format is in JSON	FUNC	HIGH	Consortium	
SS-DDAV-12	Dashboard - Data Analytics Visualization	The user select time range (future and past)	FUNC	MEDIUM	Consortium	
SS-DDAV-13	Dashboard - Data Analytics Visualization	The EO4EU should provide dashboards to visualize GIS data (rasters,vectors)	FUNC	MEDIUM	External	
SS-DDAV-14	Dashboard - Data Analytics Visualization	The EO4EU should provide dashboards for charts of numerical analysis results	FUNC	MEDIUM	External	
SS-FE-001	FE	The fusion engine shall integrate heterogeneous data sources.	FUNC	HIGH	DoW	
SS-FE-002	FE	Fusion engine shall provide preprocessing assessment and refinement of data. Preprocessing of data includes spatiotemporal intrapolation.	FUNC	HIGH	DoW	
SS-FE-003	FE	Fusion engine comprises the guidelines from Domain Specific Language.	FUNC	HIGH	DoW	
SS-FE-004	FE	The fusion engine provides risk assessment of the raw data.	FUNC	MEDIUM	Users	
SS-FE-005	FE	Dynamic Intgration of new algorithms in fusion engine	FUNC	MEDIUM	Users	
SS-FE-006	FE	Fusion engine should allow aggregating heterogenous sources of data	FUNC	HIGH	External	
SS-FE-007	FE	Fusion engine should allow aggregating S1 and S2 data	FUNC	HIGH	External	

ID	Component	Title & Description	Type	Priority	Source	Comments
SS-FE-008	FE	Provide mechanism for automatic selection of data sources to be fused for a specific use case	FUNC	HIGH	External	
SS-FSO-001	FaaS	FaaS Orchestrator shall provide computational resources and services as a per service instance of a containerized environment.	FUNC	HIGH	DoW	
SS-FSO-002	FaaS, IaaS, PaaS	Provide GPU capabilities supporting accelerated AI and High Performance Computing (HPC)	FUNC	MEDIUM	DoW	
SS-FSO-003	FaaS , IaaS, PaaS	Virtualization layer based on Linux/KVM	FUNC	HIGH	DoW	
SS-FSO-004	FaaS ,IaaS	Infrastructure as a Service (IaaS) layer using Openstack	FUNC	HIGH	DoW	
SS-FSO-005	FaaS , IaaS, PaaS	Kubernetes for running the operational workloads	FUNC	HIGH	DoW	
SS-FSO-006	Faas	Deployment isolation for every service/workflow with minimum footprint to the system.	FUNC	MEDIUM	DoW	
SS-FSO-007	FaaS	(FaaS) orchestrator with resource reservation capability	FUNC	MEDIUM	DoW	

4 Sequence diagrams

In the following section functionalities workflow end-to-end are presented. Some of the components especially those included in PaaS Tier, cannot be presented in a user story as long as their main functionality takes place in the backend. Each use case consists of a description/analysis followed by a sequence diagram that visualizes the interactions between the components.

4.1 KG Generation and Requests

The sequence diagram depicted in Figure 19 provides a visual representation of the KG generation as well as the consumption of the service by the end users. We note here that the user can either be a physical person requesting EO resources through a given user interface and/or an automated internal component that requests EO resources (driven by a user) through a dedicated API mechanism.

KG Generation

In Figure 19 the EO resources are being consumed through the ADAM platform. The KG imports EO datasets (processed datasets) from the various Services Providers with two different approaches.

1. A given API from the respective services and third-party providers (ADS, CDS, ADAM)
2. Through custom parsers for the services that do not expose an API to consume datasets.

The following steps describe the KG generation process:

Step 1: EO resources of interest are imported by the ADAM platform.

Step 2: Datasets are parsed through the ADAM parser

Step 3: The description of the datasets is parsed (from services that do not provide an API), by using custom parsers.

Step 4: The datasets from step 2 and step 3 are passthrough into the integrator. The integrator is consuming and handles the textual descriptions and the metadata extracted from the datasets

Step 5: The integrator packages the information and structures the KG.

Step 6: The KG pushes the data to the language processor for further analysis

Step 7: The language processor updates the indexes

User Request

Step 1: The user sends a request in free text Ito the graph.

Step 2: The search engine transforms the request into an embedding and sends it to the language processor.

Step 3: The language processor searches for the internal embedded database for the right indices, after receiving the embedding.

Step 4: The results are pushed to the KG.

Step 5: The results are pushed to the user (see step 1).

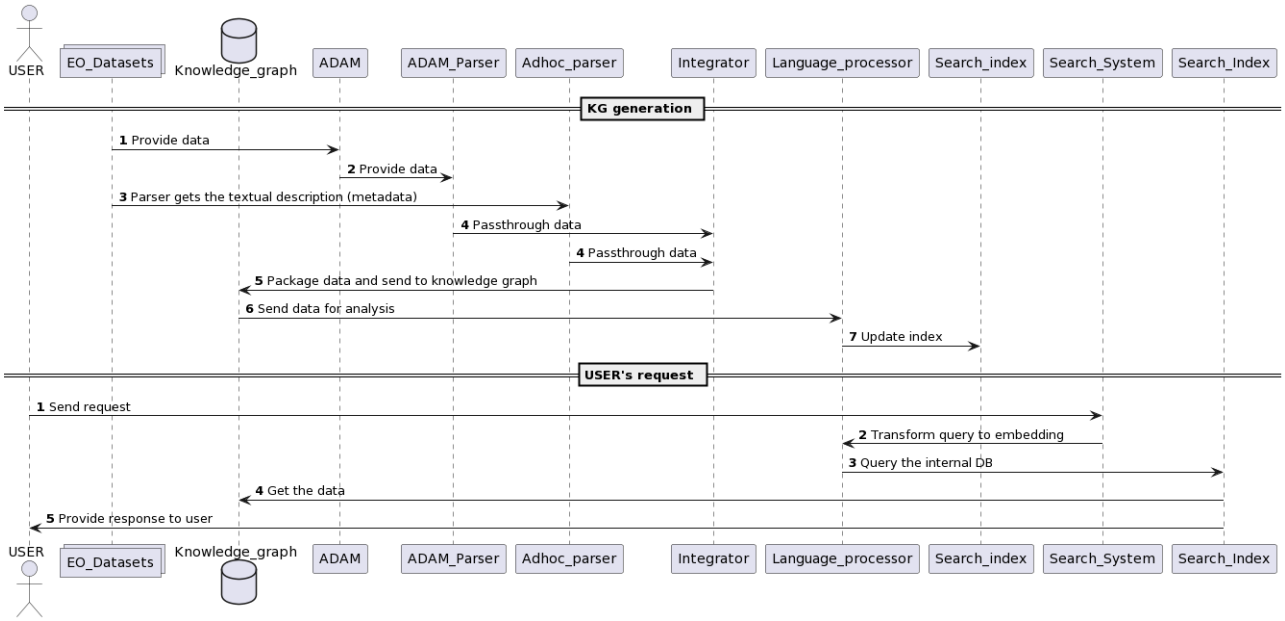

Figure 19 - KG Sequence Diagram

4.2 Service Provision

This sequence diagram depicts a holistic preview of the provisioning process of a workflow.

- Provide the needed authorization to access the platform
- Interact with the workflow editor to create a valid design.
- The workflow editor executes the appropriate language validation checks and proceeds with provision validation.
- As the design is correct and the expected resources are available for deployment the platform orchestrator executes the design.
    - In each deployment, the resource registry is advised or updated for the action of the platform orchestrator
    - Will (or not) provision virtual machines, deployments and services in Kubernetes and FaaS functions.
- A response message will be provided to the User Interface to inform the actor about the state of the workflow.
- Finally, the monitoring mechanism will continuously attend to the lifecycle of the workflow.

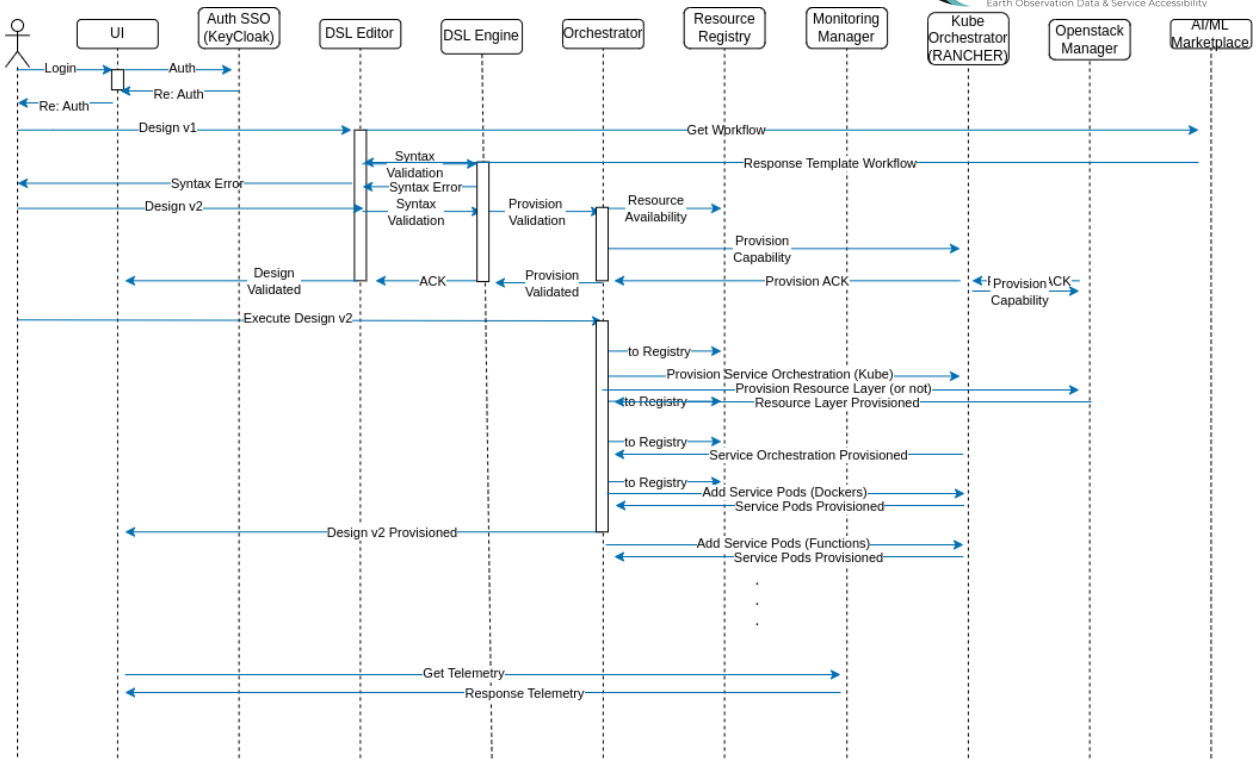

Figure 20 Service Provision SD

4.3 Service Overloading

This diagram shows how the platform scales the workload based on the processing load when the user initiates a transaction on the platform. The platform orchestrator analyses the health information of the application based on the telemetry data and when it detects that the application is overloaded, it deploys new service pods for the container-based or function-based workload.

In cases where the virtual server resources under the platform orchestrator management are not sufficient due to the increase in workloads, the platform orchestrator will request new virtual servers via Openstack Manager and create a new resource to run the workloads.

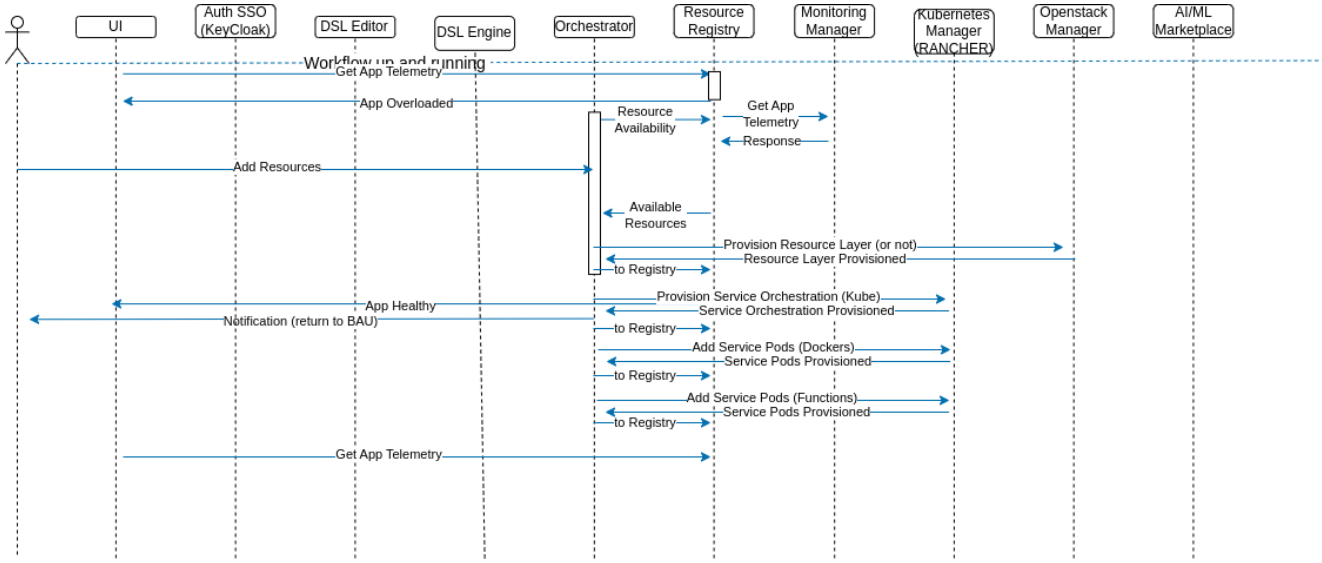

Figure 21 Service Overloading SD

4.4 Monitor Manager

In the following monitoring alarm system diagram, the Platform Controller interfaces with the Predictive Allocation component, which in turn communicates with the platform Orchestrator. The platform Orchestrator interacts with the Resource Registry to manage and allocate resources. The Monitoring Manager is responsible for monitoring the components and generating alarms when necessary. It receives information from the platform Orchestrator, Resource Registry, and other components for monitoring purposes.

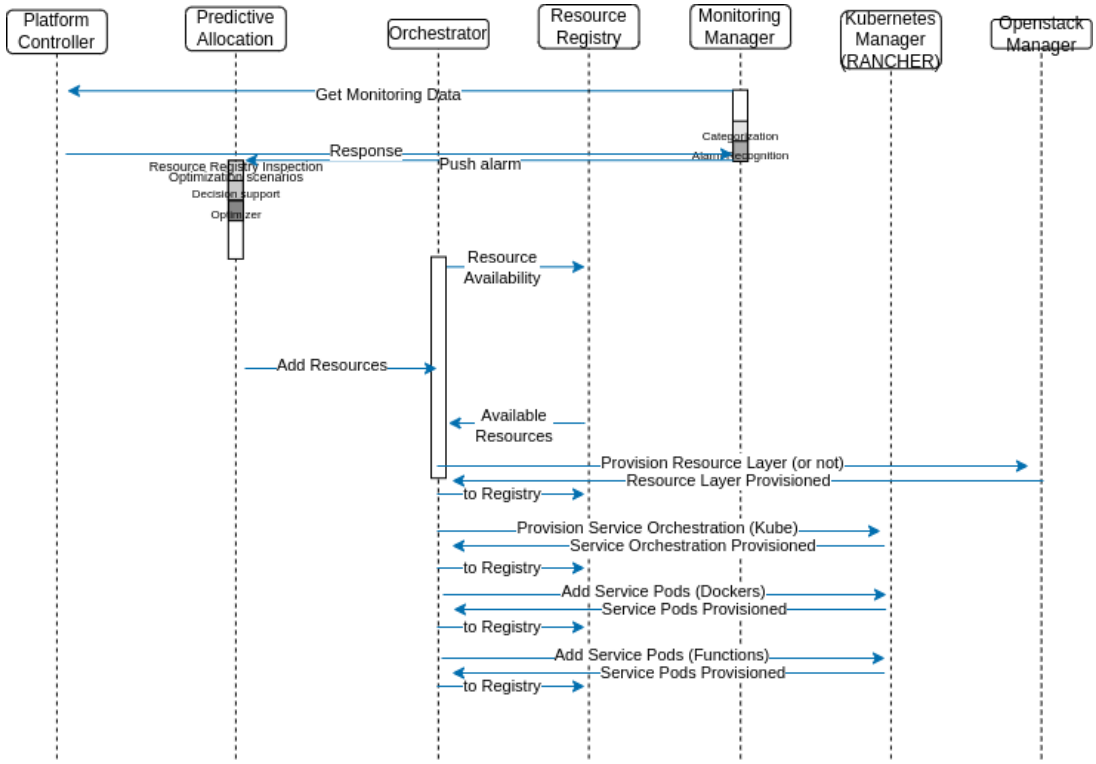

Figure 22 Monitor Manager SD

4.5 Service Scale Out/Up Manager

The Kubernetes Manager/Rancher component is connected to the Monitoring Manager and OpenStack. It handles the management and orchestration of Kubernetes clusters, ensuring their proper functioning. Openstack is responsible for providing infrastructure services such as virtual machines, storage and networking.

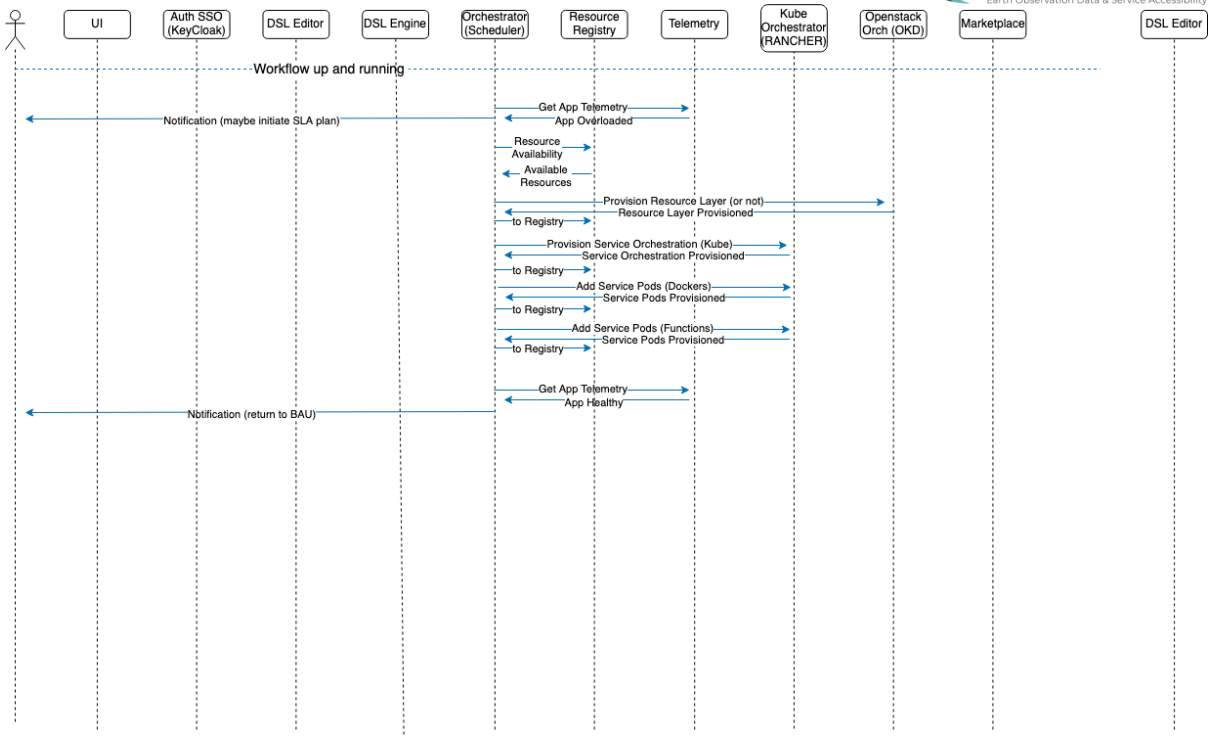

Figure 23 Service Scale Out/Up SD

4.6 Fusion Engine

Fusion engine (FE) is responsible to implement the spatiotemporal processing of several types of data and sources. A user in the EO4EU platform defines in the DSL a specific pipeline of fusion functionalities. These functionalities can be found in the WFE.

1. When a workflow is generated the DSL script is disseminated by Kafka message bus to the relevant components.
2. When Fusion Proxy receives a DSL message starts the initialization of the specific workflow, e.g. topics to be connected to, namespace, S3 buckets, pipelines requested etc.
3. A specific thread is created waiting for a message via Kafka to start the processing of the data.
4. A pipeline is called via Kubeflow Client and the pipeline is executed.
5. The results of the processing is stored back to the specific S3 bucket.
6. A message is generated for the next component to start the processing.

The workflows of users run in parallel. The output format of data produced is defined by the user in DSL yaml. If a specific computational chain creates a valuable product for the user, this chain can be stored as metadata in the KG in order to be re-executed and easily reproduce the results.

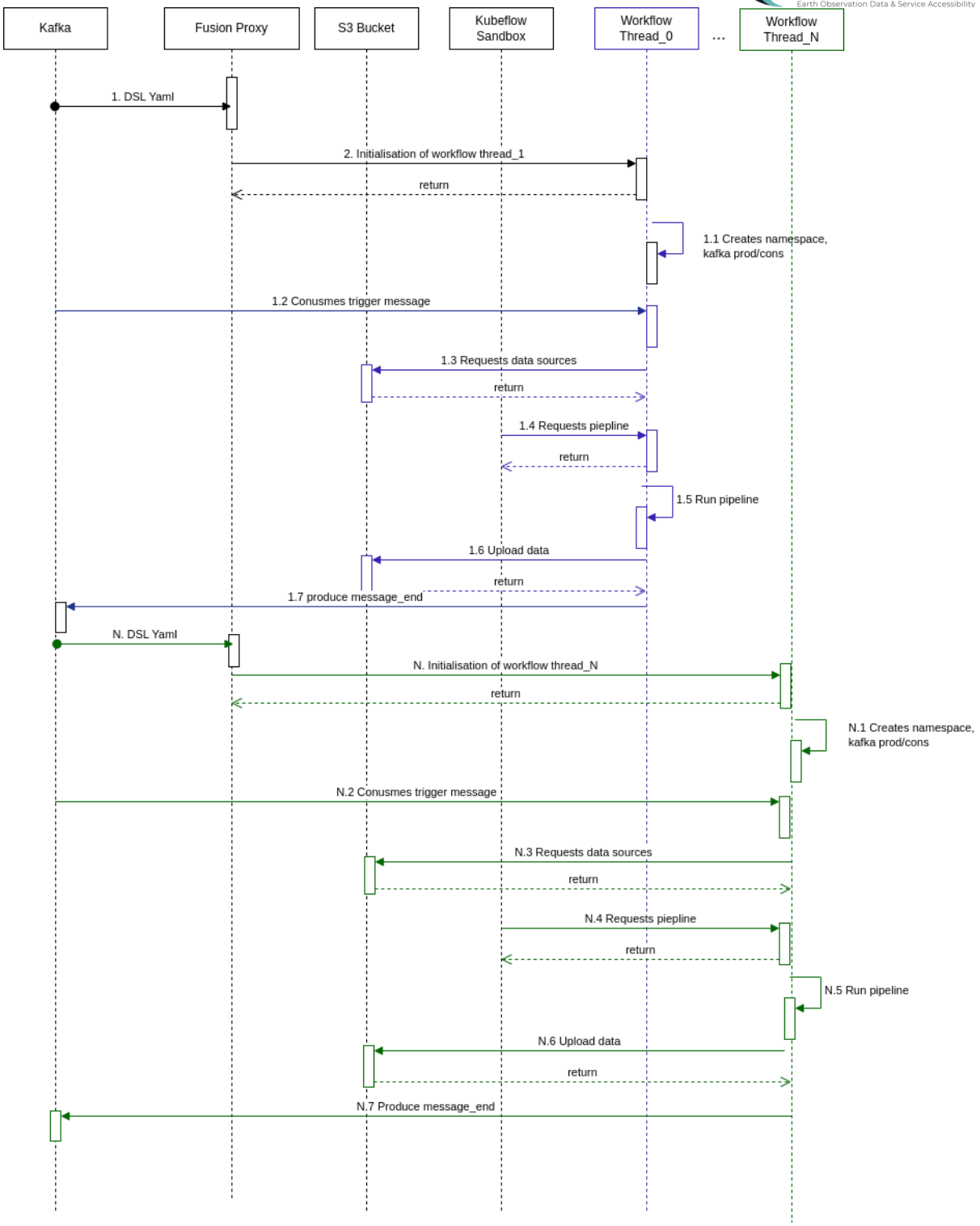

Figure 24 Fuson Engine SD

4.7 DSL Engine

The following sequence diagram describes the communication between DSL Engine sub-components during the main phases of operation. At the same time, it shows the dependency during the phase of

initialization and publishing with systems layer. Finally, it presents the internal transcode process from editor blocks to DGL.

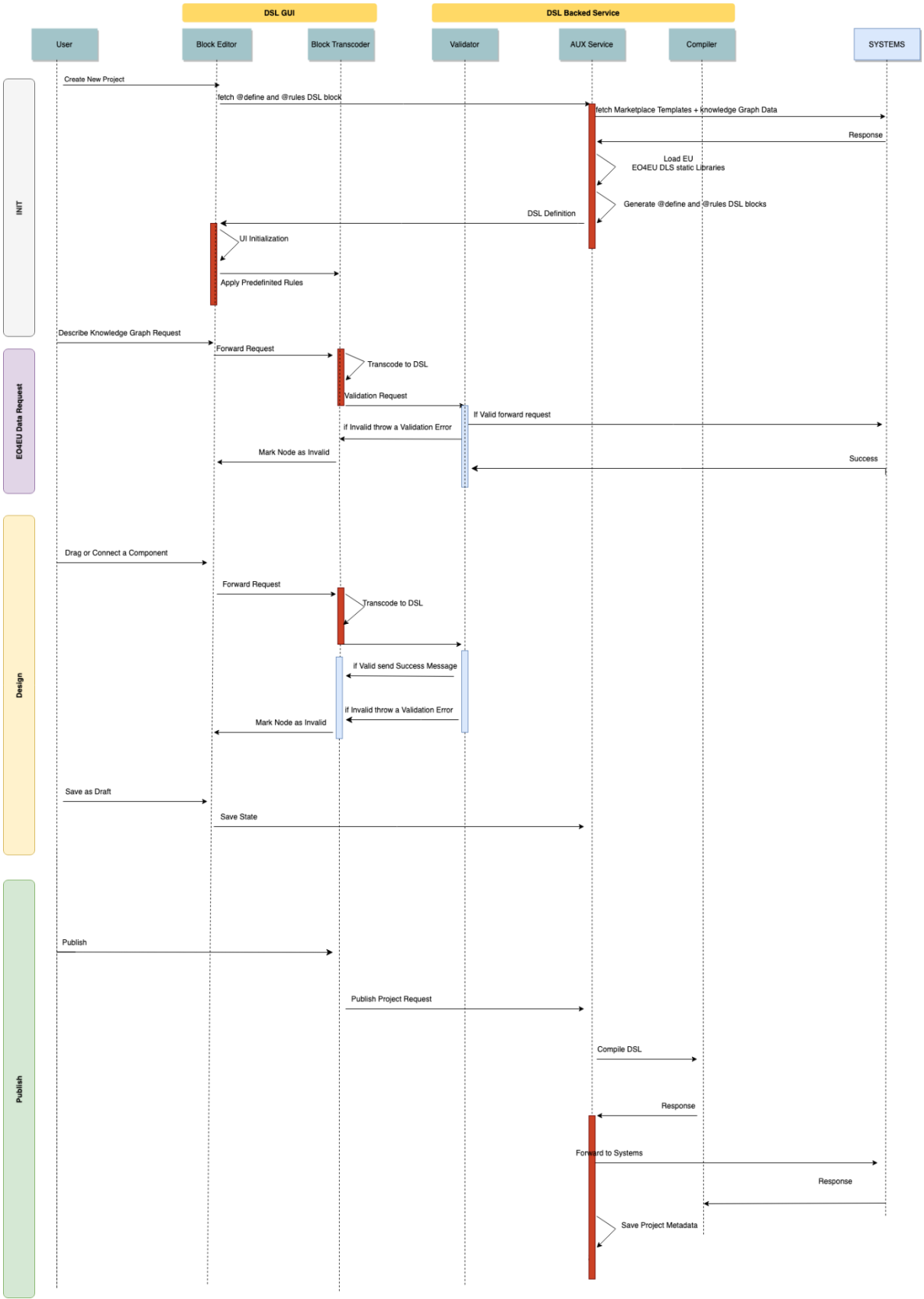

Figure 25 DSL Engine SD

4.8 DSL Engine Authentication

The sequence diagram below describes the Authentication-Authorization process of the DSL Engine component. The workflow editor as a user interaction platform implements the implicit flow as described in RFC7636 (Auth2.0) using the global SSO component of the platform as an identity provider.

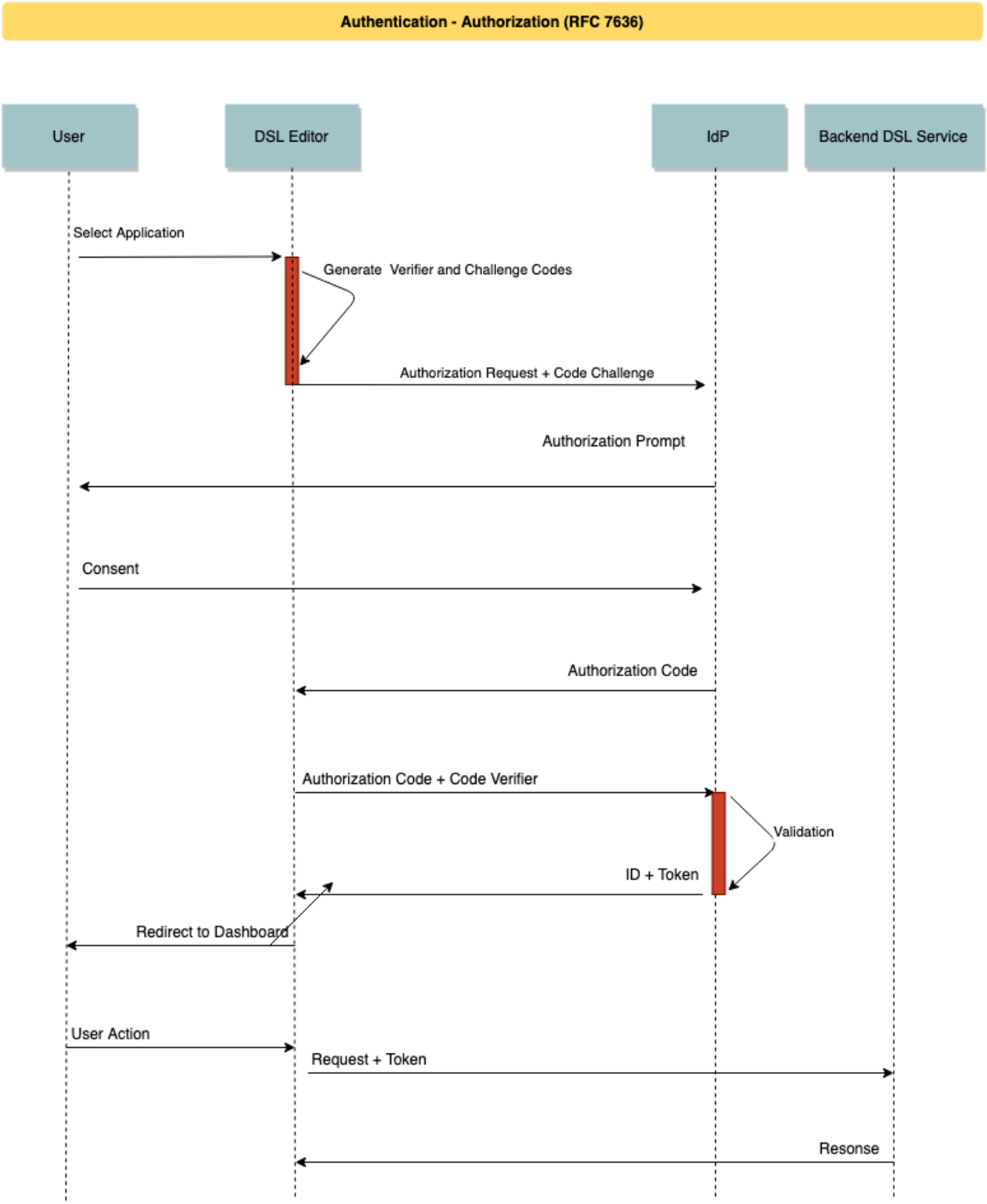

Figure 26 DSL Engine Authentication

4.9 ML user running an inference task

The sequence diagram illustrates the interactions between a user and a machine learning inference server. The user initiates the process by sending a request to the server, providing input data for inference. The server receives the request and forwards the input data directly to the machine learning model. The model performs the necessary computations and generates predictions based on the input. The server then receives the predictions from the model and sends them back to the user as a response. Finally, the user receives the predicted results and can utilize them for further analysis or decision-making purposes.

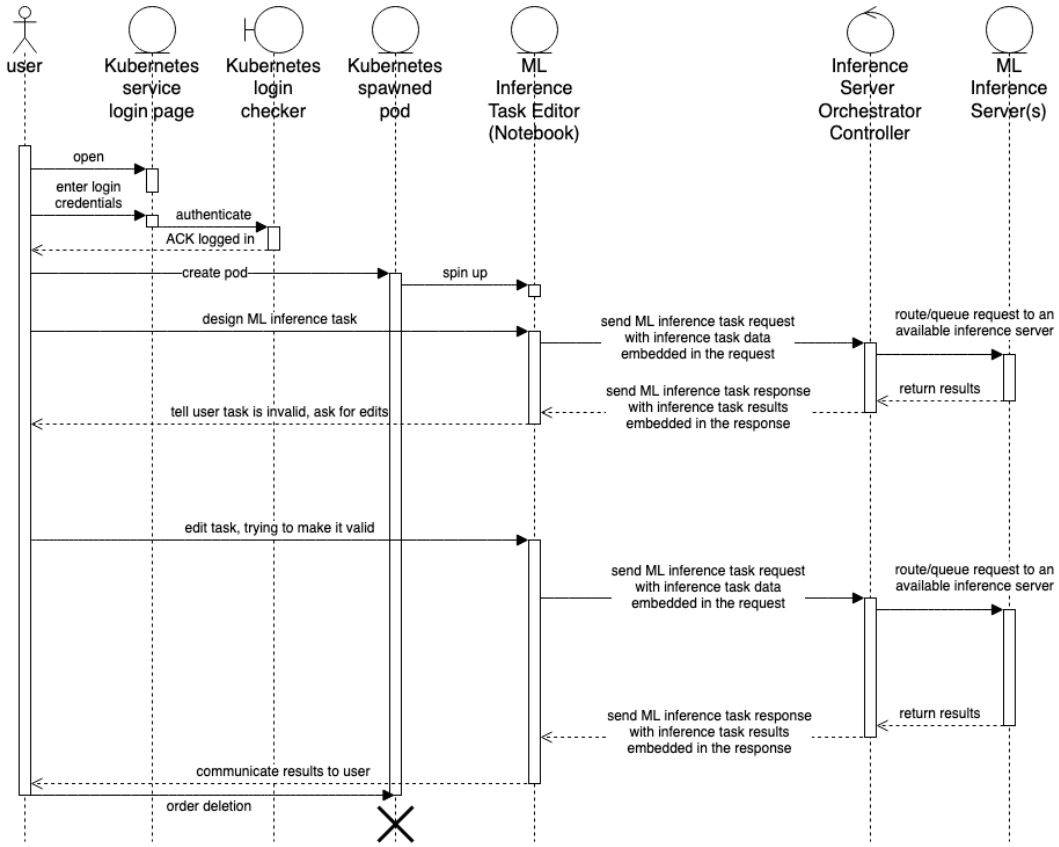

Figure 27 ML Inference Server SD

4.10 Customer Facing services

The following figure represents EBOS' CFS-Customers Facing Services-Sequence Signal Flow Diagram. The CFS Signal Flow Diagram consists of 5 basic communication access requests and authorized responses.

(1) Mainly, in the beginning the user must access the open source main EO4EU Platform System where the user has to login by registering his credentials, such as username and password to request access to the Customers Facing Services through the API Component. The API is then communicating with the UMM-User Management Model checking the Key-Cloak and verifying the personal data of the user. Once the personal username and password are verified, a granted permission-authority key is provided to the user to access the CFS components.

(2) In case the user desires to get access through the CLI, a signal request is sent to the API, which also communicates with the UMM for granting the authorized permission to the user to use command

lines-instructions for obtaining data analytics or visualization data on the dashboard, or XR/VR component.

(3) If the user requests access through the GUI of the EO4EU Platform, a corresponding GUI command is again sent through the API Component, which checks and verifies with the UMM and Key-Cloak component, in order to authorize the GUI access request.

(4) If the optional request of the user is to obtain a web XR/VR Service, the procedure is again the same where a signal request is sent to the API, which also sends a request to the UMM and Key-Cloak to get the authorized response. Once the access is verified, the communication returns to the XR/VR component in order to generate the 3D XR/VR Visualization Graphics and Data displayed to the user.

(5) The user may communicate and request any of the CFS components as a 3rd party user, by communicating through a 3rd part application and sending the requests to the API, UMM, and Key-Cloak. Once personal user account details are authorized, the user obtains the granted access to use the utilities-services of the CFS.

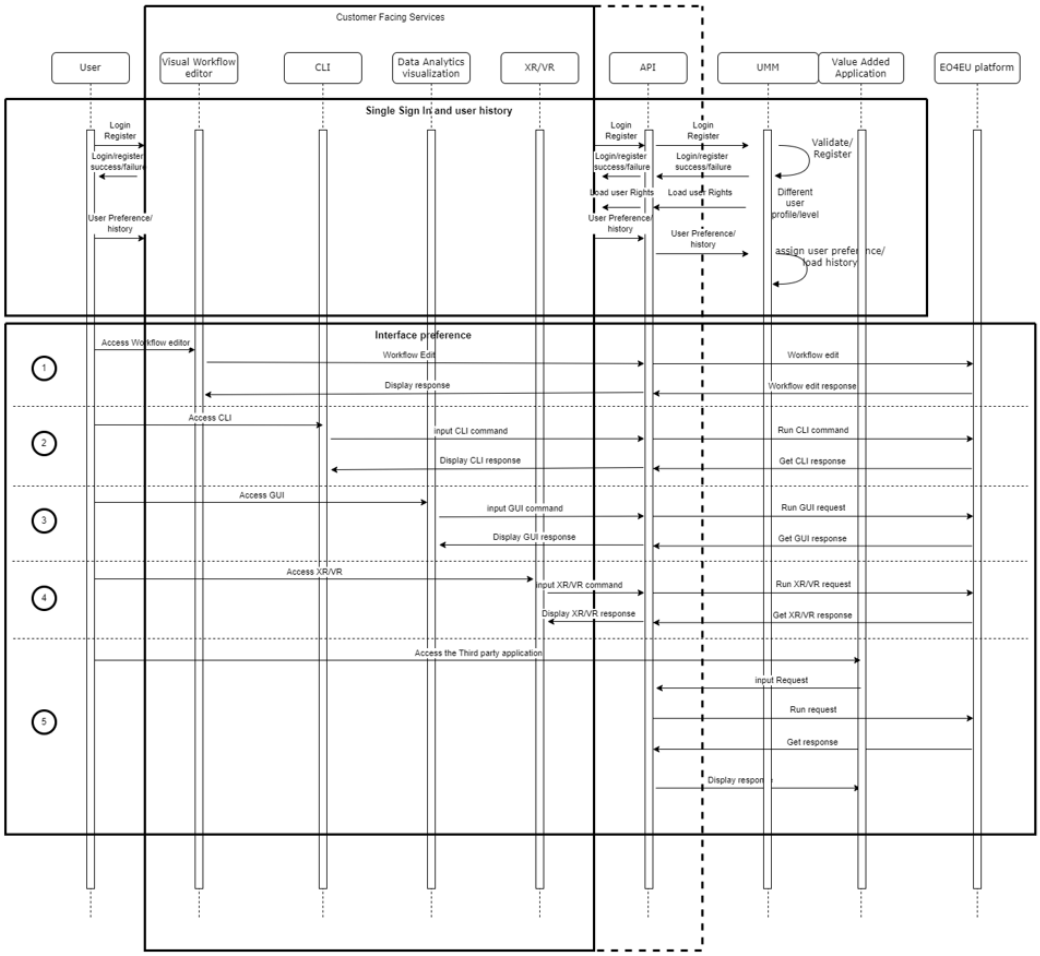

Figure 28 Customer Facing Services - Sequence Diagram

5 Related Projects

Table 2 - State of the Art and Related Projects and Architecture Systems

Name	Description	Start date
EOPEN [1]	EOPEN offers a user-friendly platform designed for non-expert Earth Observation (EO) data users, experts, and the SME community. It simplifies the utilization of Copernicus data and services for Big Data applications by delivering EO data analytics services, decision-making tools, and infrastructure support throughout the Big Data processing life cycle. This enables the seamless integration of value-added processes across various platforms.	November 2017
CANDELA [2]	The CANDELA project is dedicated to creating a platform that offers essential components and services to facilitate the rapid utilization, manipulation, exploration, and processing of Copernicus data. Its primary goal is to bridge the divide between big data technology and the Earth Observation data user community. By combining robust, established resources with innovative tools integration, the platform should empower existing Copernicus users to embrace big data technology and unlock maximum value. In addition to the toolkit that the consortium established within the platform, CANDELA empower users to seamlessly integrate existing building blocks into a unified, robust platform, thereby fostering collaboration and enabling the exploration of new approaches and services. This approach is exemplified by the development of two reference scenarios, which not only serve as validation cases but also represent genuine commercial, operational scenarios involving existing customers.	May 2018
EO4AGRI [8]	EO4AGRI is dedicated to establishing a European capability for enhancing agricultural monitoring operations, spanning from local to global scales. This is accomplished by harnessing information obtained from Copernicus satellite observations and leveraging associated geospatial and socio-economic data services.	November 2018
openEO [4]	openEO aims at creating an open application programming interface (API) that seamlessly links clients such as R, Python, and JavaScript to large Earth observation cloud infrastructures in a straightforward and cohesive manner. With this API in place: - Every client gains the ability to interact with any available back-end. - It enables the evaluation of back-ends based on factors like performance, cost, and outcomes (including validation and reproducibility)	October 2017

Name	Description	Start date
CALLISTO [6]	CALLISTO seeks to narrow the divide between Copernicus Data and Information Access Services (DIAS) providers and end-users of applications by employing specialized Artificial Intelligence (AI) solutions. This initiative aims at establishing an interoperable Big Data platform that harmoniously combines Earth Observation (EO) data with crowd-sourced, geo-referenced data, and observations from Unmanned Aerial Vehicles (UAVs). CALLISTO's capabilities put to the test in practical settings, offering geolocation-based services in fields such as agricultural policymaking, water management, journalism, and border security.	January 2021
BETTER [3]	BETTER implemented a Big Data intermediate service layer that focused on creating user-centric services and tools, while addressing the complete data lifecycle associated with EO data. This initiative aimed to bring more downstream users into the EO market and maximize the exploitation of Copernicus data and information services. These customized solutions, referred to as Data Pipelines, were driven by a large number of Data Challenges set by users deeply involved in addressing key Societal Challenges. The primary objective of BETTER was to facilitate the use of large, diverse datasets by downstream users, enabling them to focus on the analysis and extraction of potential knowledge within the data, rather than on the data processing itself.	Novemb er 2017
DEEPCUBE [5]	DeepCube, known as "Explainable AI pipelines for big Copernicus data," is a three-year project funded by the Horizon 2020 program of the European Union under the topic "Big data technologies and Artificial Intelligence for Copernicus." The project aims to unlock the potential of Copernicus data by leveraging advances in the fields of Artificial Intelligence and Semantic Web. The goal of DeepCube is to address new and ambitious problems that imply high environmental and societal impact, enhance our understanding of Earth's processes correlated with Climate Change, and feasibly generate high business value. To achieve this, the DeepCube Consortium combines mature and new ICT technologies, such as the Earth System Data Cube, the Semantic Cube, the Hopsworks platform for distributed Deep Learning, and a state-of-the-art visualization tool tailored for linked Copernicus data. They integrate these technologies to deliver an open and interoperable platform that can be deployed in several cloud infrastructures and	January 2021

Name	Description	Start date
	High-Performance Computing, including cloud-based platforms providing centralized access to Copernicus data, known as DIAS (Data and Information Access Services). These tools are then used to develop novel Deep Learning pipelines to extract value from big Copernicus data.	
CENTURION [7]	CENTURION combines and advances groundbreaking innovations in spatio-temporal datacubes and AI-as-a-Service, creating a platform that unleashes the use of Copernicus data for both EO and non-EO markets.	2021

6 Conclusions

This document contains an overview of the planned high-level architecture and an explained description of the technical specifications. D2.4 provides a clear outline of the planned components, define the capabilities of the component interfaces, shows possible interactions and dependencies between different components and describes the runtime environment of the components/ In addition this deliverable includes an analysis of the sequence diagrams of the components.

7 References

[1] EOPEN - https://eopen-project.eu/ , last accessed November 2021
[2] CANDELA - http://candela-h2020.eu/overview , last accessed November 2021
[3] BETTER - https://www.ec-better.eu/pages/better-project , last accessed November 2021
[4] OpenEO - https://openeo.org/ , last accessed November 2021
[5] DEEPCUBE - https://deepcube-h2020.eu/ , last accessed November 2021
[6] CALLISTO - https://callisto-h2020.eu/ , last accessed November 2021
[7] CENTURION - https://www.centurion-project.eu , last accessed November 2021
[8] EO4AGRI https://eo4agri.eu/ , last accessed November 2021

Annex 1

Table 3 – Mapping between Tiers and Components

Components/Tiers	Data T.	IaaS T.	PaaS T.	ML T.	Front-end T.
Knowledge Graph	+				
OpenStack		+			
Kubernetes Platform Manager			+		
Auxiliary/support - GitLab Git Repository and CI/CD			+		
Auxiliary/support - GitLab Container Image Registry			+		
Auxiliary/support - Configuration Management and Day2 Operations			+		
Platform Orchestrator			+		
Authentication SSO			+		
Message Bus			+		
DSL Engine			+		
Fusion Engine			+		
AI/ML Marketplace			+		
Machine Learning Model				+	
Machine Learning Inference Server				+	
Dashboard - Data Analytics Visualization					+
EO4EU API					+
XR System					+
UMM					+
WorkFlow Editor					+