# Artificial Intelligence for the #EOSC

Valentin Kozlov for AI4EOSC, SCC-D3A

HPC-DIC meeting, 13-Jan-2023

European Commission (EC) initiative to **provide** researchers, companies and citizens with a **federated** and **open multi-disciplinary** environment for:

> publishing
>
> finding
>
> re-use

of **data**, **software**, and **services**
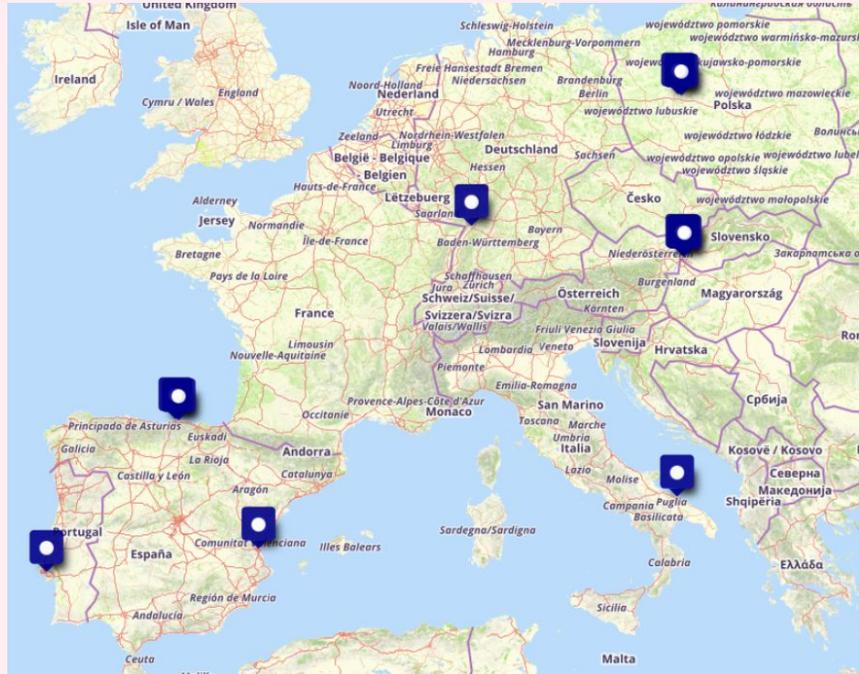
It is also about

> policies
>
> knowledge transfer
>
> funding opportunities
>
> …

# AI4EOSC

Artificial Intelligence for the #EOSC

- Evolution of the DEEP Hybrid DataCloud platform
- HORIZON-INFRA-2021-EOSC-01-04 call
- Runs September 1st 2022 – August 2025 (36 months)
- 7 academic partners

  + 2 SME

  + 1 non-profit organization

Advanced features for distributed, federated, composite learning, metadata provenance, MLOps, event-driven data processing, and provision of AI/ML/DL services

# Objectives

### Objective 1

Provide **feature rich services** and **platform** to **build** and **deploy** custom AI applications in the EOSC

### Objective 2

Enhance existing cloud services to support **AI on distributed datasets**, with a particular focus on **federated learning**

### Objective 3

Deliver methods to **compose AI tools**, enabling the development of complex data-driven composite applications

### Objective 4

Deliver an **AI exchange** in the context of the EOSC, enhancing and increasing the application offer currently available
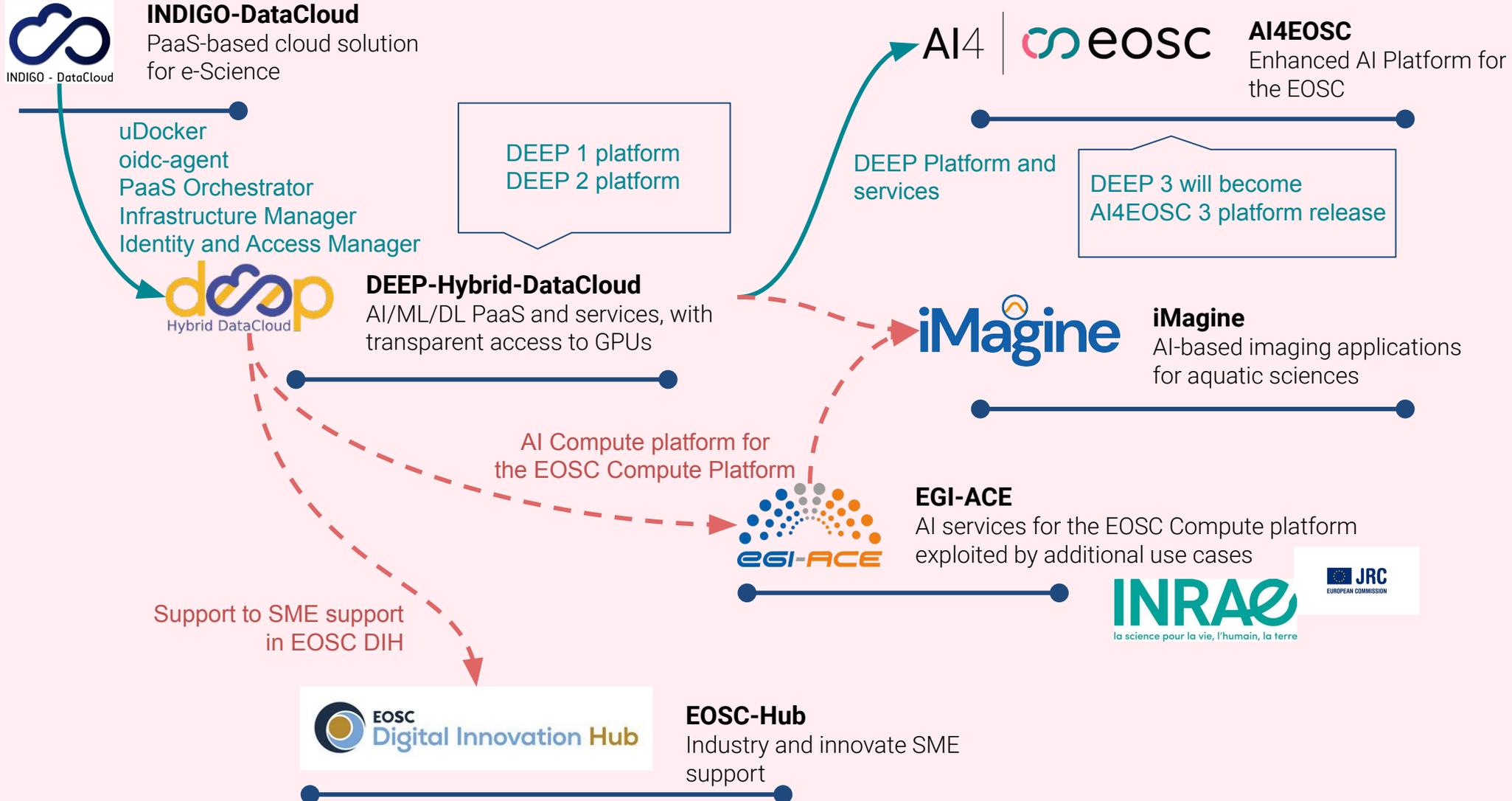
### Objective 5

**Extend the service offer** and the capabilities being offered through the EOSC portal, **with focus on AI**

### Goal

Foster an AI exchange in the EOSC context, with added value, innovative and easily customizable services

# Background and ecosystem



| 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 | 2026 | ... |

**INDIGO-DataCloud**
PaaS-based cloud solution for e-Science

uDocker
oidc-agent
PaaS Orchestrator
Infrastructure Manager
Identity and Access Manager

DEEP 1 platform
DEEP 2 platform

**AI4EOSC**
Enhanced AI Platform for the EOSC

DEEP Platform and services

DEEP 3 will become AI4EOSC 3 platform release

**DEEP-Hybrid-DataCloud**
AI/ML/DL PaaS and services, with transparent access to GPUs

**iMagine**
AI-based imaging applications for aquatic sciences

AI Compute platform for the EOSC Compute Platform

**EGI-ACE**
AI services for the EOSC Compute platform exploited by additional use cases

Support to SME support in EOSC DIH

**EOSC-Hub**
Industry and innovate SME support

# DEEP evolves in...

# AI4EOSC

Training on single site, centralized dataset expected

Federated learning, split learning, gossip learning, making possible training on decentralized datasets

Single AI application, self deployed or on serverless computing

Composite AI for complex AI tools and applications through function composition and serverless computing

Central management of onboarded sites, complex on-premises deployment

Enhanced onboarding of resources, easier deployment on-premises

# AI4EOSC conceptual diagram
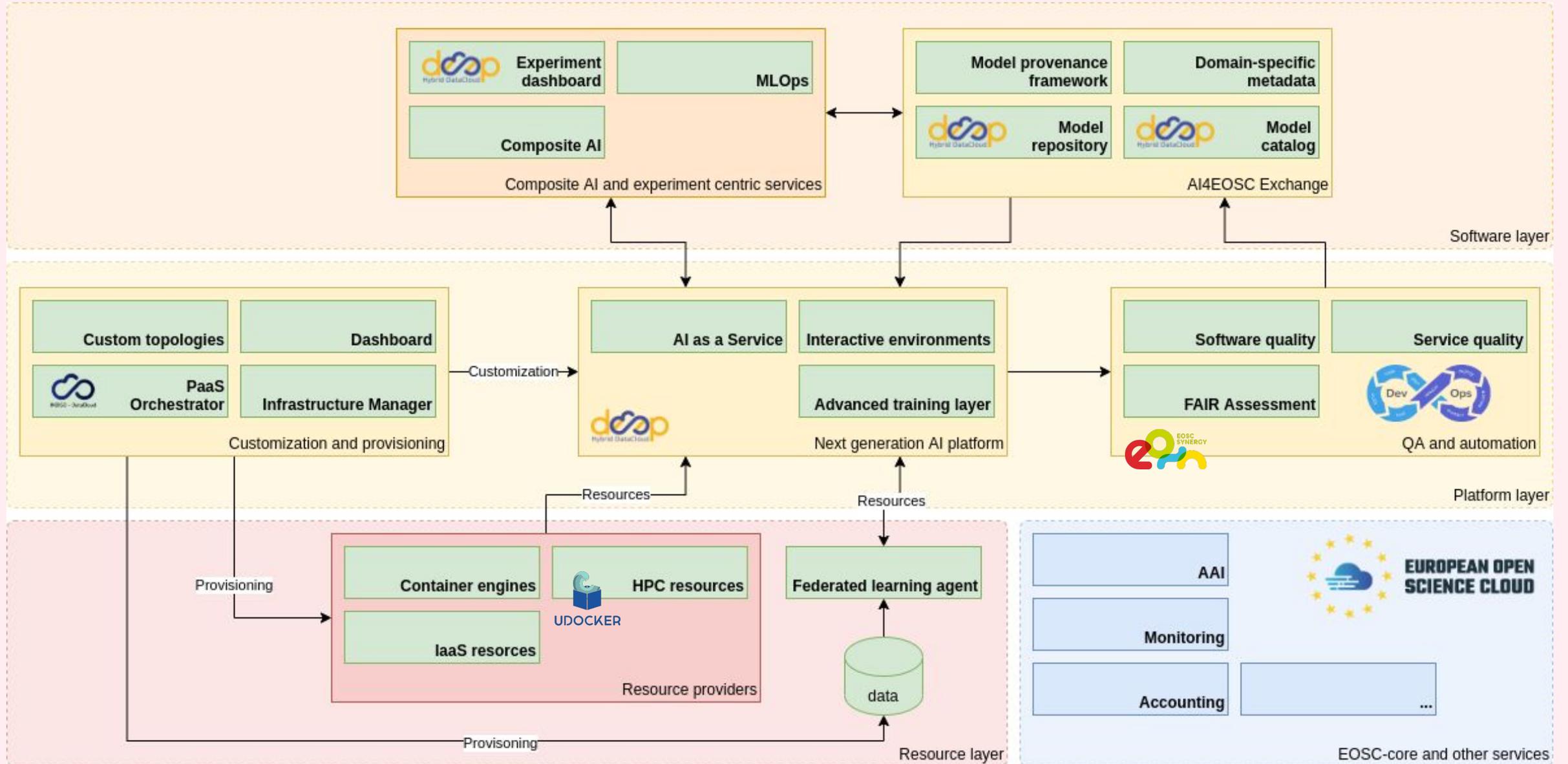
# AI4EOSC:
# use cases



Agrometeorology
_____

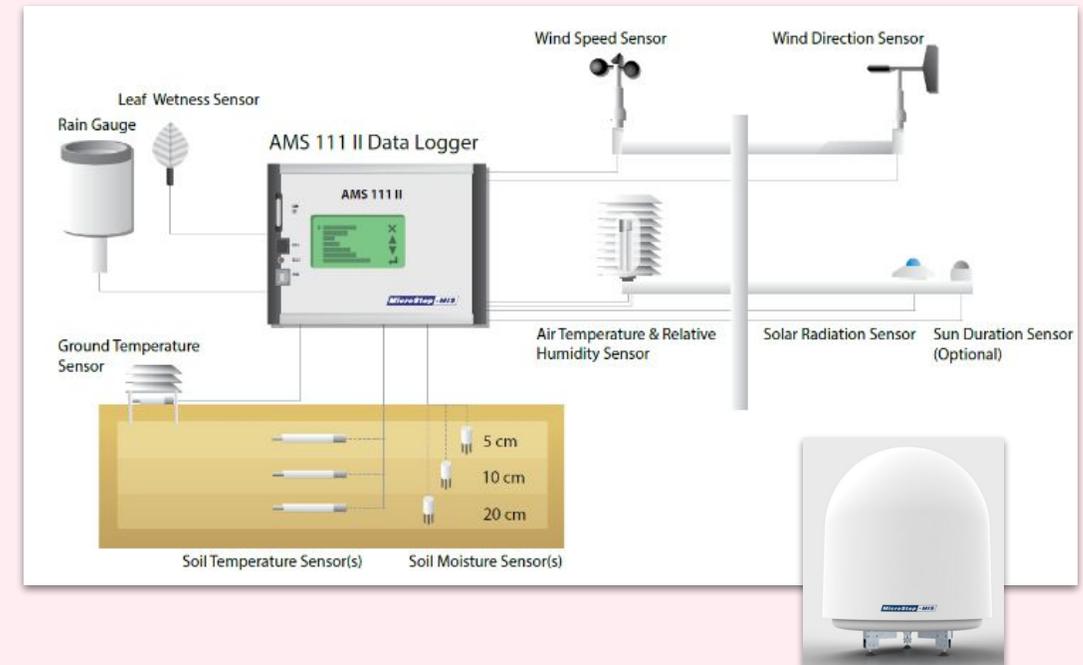Integrated plant protection
_____

Automated thermography
_____

# Agrometeorology

**Aim**: Usage of satellite imagery, in-site measurements, and weather forecasts to generate added-value products for improving farmers activity: **Agro nowcasting system for farmers** to better plan their actions in short range horizon.

**Currently**: Measurement system - TRL9, prediction system - TRL3

**Within AI4EOSC**: Enhancement of the prediction subsystem following a Composite AI approach to combine the different machine learning models used for the different data sources

**Partners**: Microstep, IISAS, Predictia

# Integrated plant protection

**Aim:** To **determine the risk of disease and pests in agricultural crops** and determine the phases of plant growth and the condition of crops. The developed AI models are going to be integrated into existing national advisory platforms, operated by WODR and PSNC.

**Currentl**y: WODR and PSNC operate a national advisory platform for farmers (eDWIN), which includes a network of meteorological ground stations, the Farm Management System, and ground observations of the occurrence of diseases and pests. The current solutions are based on predictive mathematical models.

**Within AI4EOSC**: The plan is to add to current mathematical prediction models the ML/DL-based models used for recognition of the plant diseases. Adding the new sources of the data. Initial focus on wheat and sugar beats and detection of the fungal diseases.

**Partners**: WODR, PSNC

# Automated Thermography

**Aim:** To **identify heat losses and thermal bridges in buildings and infrastructures using drone-based images and ML/DL approach** in order to provide a corresponding automated AI-based service.

**Currently:** The group owns a dataset of drone-based images on urban districts and drone-based thermal images on a campus district (ca. 0.8TB). The identification of thermal bridges on roofs is already possible using DL (TRL 4). The identification of leakages in district heating networks is possible too (TRL 5/6).

**Within AI4EOSC:** Targets enlargement of the training dataset, AI model improvement, optimisation of the workflows, and creation of a cloud-based automated service

**Partners:** KIT (IIP, SCC)

# AI4EOSC Current state

D3.1 State of the art landscaping and technology scouting and venturing

   review of various Deep/Federated Learning, MLOps Frameworks

_____

D3.2 Requirements elicitation for co-design

   architecture diagramming: [C4 Model](#)

_____

D6.1 Analysis of user applications, collection of requirements

   methodology inspired by the [DSDM Agile Project Framework](#)

_____

# AI4EOSC Expected results

Cloud based **AI platform**, integrated into the **EOSC**, with distributed training capabilities

**Reusable AI/ML applications** offered through AI4EOSC exchange, with easy deployment paths

**Best practices** and recommendations for AI practitioners and data scientists

**MLOps** technological framework providing drift detection capabilities

**Model provenance** metadata framework, covering the whole AI/ML

# iMagine – Imaging data and services for aquatic science

iMagine

iMagine aims to deploy, operate, validate, and promote a dedicated AI framework and platform, connected to EOSC and AI4EU, to give researchers in aquatic sciences **open access to a diverse portfolio of AI-based image analysis services and image repositories** from multiple RIs, of relevance to the overarching theme of **Healthy oceans, seas, coastal and inland waters.**

## Use cases

1. Aquatic Litter Drones: Aquatic Litter monitoring system using drones (DFKI, MARIS, OGS)

2. Zooscan – EcoTaxa pipeline: Taxonomic identification of zooplankton using Zooscan (EMSO ERIC, UPC, Ifremer, MI)

3. Ecosystem monitoring at EMSO sites by video imagery (SU)

4. Oil Spill Detection: Oil spill detection from satellite images (CMCC, OrbitalEOS, UNITN)

5. Flowcam phytoplankton identification: Taxonomic identification of phytoplankton using Flowcam images (VLIZ)

## 3 Prototypes

### Prototype 1
Underwater Noise Identification: Underwater noise identification from acoustic recordings using spectrograms (VLIZ)

### Prototype 2
Beach Monitoring: Posidonia oceanica berms and rip-currents detection from beach monitoring systems (SOCIB)

### Prototype 3
Freshwater diatoms identification: Identification of freshwater diatoms using microscopic images (UL-LIEC – lead)

Use cases & Prototypes led by domain experts and supported by the iMagine Competence Centre

## Best practises and training materials
A competence centre will oversee the best practices developed by the use cases and collect them in ad-hoc documentation for wider distribution. Training materials on the infrastructures will also be available for interested users to explore all options and start experimenting themselves.

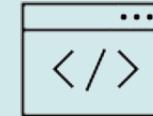POWERED BY egi POWERED BY

# AI4 | ∞ eosc

**Funded by the European Union**

---

AI4EOSC

ai4eosc-po@listas.csic.es

ai4eosc.eu

## ↑↑ Reach us! ↑↑

## Thank you for your attention

# Links used in the presentation

**AI4EOSC** project:  https://ai4eosc.eu

**iMagine** project:     https://www.imagine-ai.eu/

**DEEP-Hybrid-Datacloud**: https://deep-hybrid-datacloud.eu/

**DEEP** Platform in the EOSC Marketplace:

https://marketplace.eosc-portal.eu/services/deepaas-training-facility

**udocker** container tool:  GitHub:          https://github.com/indigo-dc/udocker

Documentation:  https://indigo-dc.github.io/udocker/

**oidc-agent**:                    GitHub: https://github.com/indigo-dc/oidc-agent

Documentation: https://indigo-dc.gitbook.io/oidc-agent/

**EOSC-Synergy** project: https://www.eosc-synergy.eu/

**SQAaaS** platform:          https://sqaaas.eosc-synergy.eu/

(To try the platform, please, "Sign In" via EGI Check-In using e.g. KIT account)