EDITO

Copernicus Marine Service

EMODnet European Marine Observation and Data Network

European Digital Twin Ocean

EDITO for HORIZON-MISS-2025-03-OCEAN-08

EU Digital Twin Ocean: Contribution to the EU DTO core infrastructure through applications for sustainable ocean management

Funded by the European Union

This work was funded by the European Union under grant agreement no. 101227771. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Contents

1. Introduction 3
2. Overview of EDITO and its architecture 3
3. Key Enhancements in EDITO 2 3
3.1. Scalable Infrastructure for High-Performance Use 4
3.2. Enhanced DTO Engine for Developers and Scientists 4
3.3. Interoperability 4
3.4. Federated Identity and Role-Based Access Control 5
3.5. Data Lake and Metadata Management 5
3.6. AI Assistant for Smarter Search and Discovery 5
3.7. User Support 5
3.8. Scientific Validation of Data and Model Outputs 6

1. Introduction

The EU Digital Twin Ocean (DTO) platform, EDITO, is being developed within the framework of Horizon Europe projects and is intended to evolve into a long-term, sustained European infrastructure that extends beyond the lifetime of the EDITO 2 HE project.

- EDITO-Infra (Oct 2022-Feb 2025) and EDITO-ModelLab (Jan 2023 – Dec 2025);
- EDITO-2 (Mar 2025 – Aug 2028).

More information on the project website: **European Digital Twin Ocean - Powered by EDITO - EDITO**

The EDITO2 team will collaborate with the selected projects to identify and support their resource requirements, ensuring that their needs are met while maintaining a balanced allocation of resources across all strategic users of the EDITO2 infrastructure. These projects will be given priority in the co-construction process of EDITO2.

2. Overview of EDITO and its architecture

You can find documentation on EDITO platform at the following links:

EDITO documentation on EDITO platform **https://docs.lab.dive.edito.eu/** ;
EDITO project deliverables **https://zenodo.org/communities/eu-dto/** .

3. Key Enhancements in EDITO 2

EDITO 2 is the next phase of the European Digital Twin Ocean, significantly expanding its capabilities to support large-scale scientific collaboration, near-real-time processing, and seamless integration across cloud, HPC, and data ecosystems. The platform is built on the foundations of Phase 1 and introduces improved performance, simplified user interactions, robust data management, and expanded integration with scientific infrastructure across Europe.

3.1. Scalable Infrastructure for High-Performance Use

To support increasing demand and complex digital twin applications, EDITO 2 upgrades its core infrastructure.

- **Free computing resources** shared among users expand both in CPU and GPU nodes;
- **Free storage** shared among users **expands** to **1 Petabyte** of elastic, scalable storage;
- Integration of a credit system to manage user usage;
- Integration of a billing system to enable usage continuity after exceeding the free plan;
- Implementation of **redundancy and high availability** to reduce downtime;
- Continued integration with **HPC centers** (e.g. LUMI and BSC);
- Experimentation with **EDITO sidecar deployments** in HPC/data centers (e.g., Mercator private computing infrastructure).

3.2. Enhanced DTO Engine for Developers and Scientists

The DTO Engine has been evolving into a more powerful, user-friendly tool for service deployment and orchestration.

- Users can now **contribute to services and processes** without GitLab accounts;
- Simplified workflows for **data scientists and developers.** Replacement of the
- existing Helm Chart based deployment with a one click deployment of services from either a Dockerfile, an existing image or a git repository;
- Improve services and processes monitoring;
- Add data, services and processes usage analytics;
- Support for workflow creation, scheduling, and monitoring within EDITO;
- Seamless launching of services on HPCs, GPUs, or external data/computing centers;
- **Personal/group catalog management** for services, processes, and data;
- Updated IDEs (VSCode, JupyterLab, RStudio) with oceanographic tools;
- New services for **distributed computing** (Dask, Spark) and **AI training** ;
- Introduction of automated testing and expert validation for contributions;
- Only tested content is published in public catalogs;
- Option for human validation with visible quality control level.

3.3. Interoperability

EDITO 2 aligns with **open standards** (OpenAPI, STAC, OGC) and EU initiatives (Destination Earth, EOSC, Copernicus).

3.4. Federated Identity and Role-Based Access Control

EDITO 2 enhances security and collaboration through standardized, federated user management.

- Support for **federated authentication** via existing institutional credentials (e.g., Copernicus, EOSC);
- Use of **OAuth2** and EDITO Keycloak for authentication and authorization;
- Definition of a **common authorization vocabulary** for consistent access control;
- **Location-based access controls** to meet legal and strategic data restrictions.

3.5. Data Lake and Metadata Management

EDITO 2 introduces robust, automated workflows for handling increasing data complexity and volume.

- Development of technical guidelines for interoperable data and metadata
    - standards, including STAC integration;
- Implementation of **automated data ingestion, curation, and validation**. This includes automatic data and metadata content validation and built-in data transformation tool (e.g., netcdf to zarr conversion);
- Integration of an expert-led **"Stamp Validation Process"** for scientific quality assurance.

3.6. AI Assistant for Smarter Search and Discovery

A natural-language-driven AI assistant enhances the user's ability to find relevant data, services, and models.

- Creation of an **AI-powered assistant** to guide users in searching EMODnet and Copernicus Marine metadata catalogs;
- Support for **natural language queries** via a search bar on the landing page;
- Future expansion to include discovery of **Digital Twin models and services**
    - onboarded on EDITO;
- Integration with actionable AI to trigger computation on the fly.

3.7. User Support

EDITO 2 provides a two-tier support model to guide users at every stage—from discovery to advanced contribution.

- **Level 1 support**: AI chatbot, human agents, ticketing system, and a public knowledge base;
- **Level 2 support (Explore)**: Personalized help for data navigation, process use, and service integration;
- **Level 2 support (Create & Contribute)**: Expert assistance for publishing, processing, HPC usage, and best-practice contributions.

3.8. Scientific Validation of Data and Model Outputs

Ensuring trust and quality, EDITO 2 introduces a rigorous validation framework for contributed datasets and model results.

- Introduction of a **Stamp Validation Process** for models and datasets;
- **Expert review** ensures data meets accuracy and quality standards and assigns a "Validated" label after compliance checks;
- Validated datasets are flagged, and unvalidated ones remain user responsible;
- Users can clearly distinguish **trusted vs. unverified data**, improving confidence in operational use.