

Project Title	Blue-Cloud 2026: A federated European FAIR and Open Research Ecosystem for oceans, seas, coastal and inland waters
Project Acronym	Blue-Cloud 2026
Project Number	101094227
Type of project	RIA – Research and Innovation Action
Topics	HORIZON-INFRA-2022-EOSC-01
Starting date of Project	01 January 2023
Duration of the project	42 months
Website	www.blue-cloud.org

D2.6 – Tuning between Blue-Cloud Data Lakes and DTO development, 1st report

Work Package	WP2 WP FAIR compliant Discovery and Access services
Task	T2.4 DTO Taskforce for tuning Blue-Cloud data lakes with DTO developments
Lead author	Marina Tonani (MOI)
Contributors	
Peer reviewers	Sara Pittonet (Trust-IT), Pasquale Pagano (CNR), Dick Schaap (MARIS)
Version	V1.0
Due Date	30/04/2024
Submission Date	15/05/2024

Dissemination Level

<input checked="" type="checkbox"/>	PU: Public
<input type="checkbox"/>	CO: Confidential, only for members of the consortium (including the Commission)
<input type="checkbox"/>	EU-RES. Classified Information: RESTREINT UE (Commission Decision 2005/444/EC)
<input type="checkbox"/>	EU-CON. Classified Information: CONFIDENTIEL UE (Commission Decision 2005/444/EC)
<input type="checkbox"/>	EU-SEC. Classified Information: SECRET UE (Commission Decision 2005/444/EC)

Version History

Revision	Date	Editors	Comments
0.1	30/04/2024	Marina Tonani (MOi)	First version
0.2	10/05/2024	Dick Schaap (MARIS)	Edits and suggestions
0.3	15/05/2024	Pasquale Pagano (CNR)	Review and approval
0.4			

Glossary of terms

Item	Description
EDITO	European Digital Twin of the Ocean
VRE	Virtual Research Environment
EMODnet	European Marine Observation and Data Network
DOF	Digital Ocean Forum

Keywords

DTO; EMODnet; VRE, Data Lake, Co-construction, collaboration, synergies, DOF

Disclaimer

The Blue-Cloud 2026 project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101094227. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

EXECUTIVE SUMMARY

The European Digital Twin Ocean (EU DTO) initiative, announced by President von der Leyen, aims to harness digital technology to provide comprehensive knowledge of the ocean. The European Digital Twin of the Ocean seeks to provide an innovative set of user-driven, interactive and decision-making tools backed by the best science and data. Its core development is underway with the European Union (EU) funding through the European Digital Twin Ocean (EDITO) R&D projects. Two sister projects, namely EDITO-Infra and EDITO-Model Lab, are building the operational backbone infrastructure of the European Digital Twin Ocean. Key components of the EDITO approach include a Data Lake, Processing Engine, and Virtual Simulation Environment. EDITO aims to integrate and expand existing European assets and capabilities, starting with the Copernicus Marine and EMODnet data and services, to develop the EU Public Infrastructure for the EU DTO. The vision is to widen the engagement to additional projects and infrastructures that can contribute with complementary data, services, algorithms, and computing resources. From that perspective, EDITO is considering engagement through projects such as ILIAD and Blue-Cloud 2026 and several other EU research projects and e-infrastructures, such as EuroHPC. The plan is also to integrate EDITO with Destination Earth (DestinE), the European Commission's flagship initiative for a sustainable future.

The DTO Task Force in Blue-Cloud 2026 has been set up to ensure the complementarity of Blue-Cloud 2026 with ongoing Digital Ocean initiatives. Therefore, the Blue-Cloud DTO Task Force seeks to identify complementary data resources and analytical services which Blue-Cloud could contribute to the EU DTO data lake and applications portfolio. To be able to deliver these on an operational basis, the Blue-Cloud DTO Task Force aims to promote and elaborate technical synergies and interoperability between EDITO and the Blue-Cloud architectures, data and services.

The Task Force's activities began with a presentation and short discussion at the Blue-Cloud 2026 Kick-Off meeting, where the EDITO initiative was introduced. This was followed by a dedicated workshop between managers and developers of the Blue-Cloud 2026, EDITO-Infra and EDITO-Modellab projects. The presentations and discussions created a level playing field and gave the early impression that the technologies in use in Blue-Cloud and EDITO are already quite aligned. Areas of potential collaboration were identified, and for the short term, a plan was made to demonstrate synergy between Blue-Cloud 2026 and EDITO at the upcoming Digital Ocean Forum 2024 (DOF 2024) in June 2024 by deploying a copy of a successful Blue-Cloud VLab at the EDITO infrastructure. In the meantime, this plan has been accepted by the DOF 2024 organisers, including EU DG RTD and DG MARE, and technical activities have started to roll out this plan. For the longer term, more structural cooperation and connectivity will be desired as Blue-Cloud is developing and delivering an increasing portfolio of Virtual Labs with potential for use in the

EU DTO context while also building an interesting data provision through a federation of multiple leading data repositories and deploying fast performing data lakes with high-quality data and data products, which might complement the core data offer for EDITO from Copernicus Marine and EMODnet. Therefore, ongoing dialogue and technical exchanges are required to identify in more detail the added value that Blue-Cloud 2026 could offer to the EU DTO data and services portfolio and to elaborate on how smart connectivity could be established between the Blue-Cloud digital ecosystem and the EDITO infrastructure to achieve a dynamic exchange. These discussions and analyses should become part of a roadmap for co-construction between Blue-Cloud 2026 and EDITO, which could then be proposed to the EU DTO management team.

TABLE OF CONTENTS

1. Introduction.....	6
2. DTO task force in BlueCloud2026.....	7
3. DTO task force activities.....	8
3.1. Next steps.....	12
APPENDIX on EDITO-Infra Architecture.....	13
1.1. EDITO-Infra project.....	13
1.1.1. EDITO-Infra architecture.....	13

TABLE OF FIGURES

Figure 1: List of participants attending the DTO Task force Workshop.....	9
Figure 2: Agenda first day of DTO task force Workshop	10
Figure 3: Agenda second day of DTO Task Force Workshop	10
Figure 4: Blue Cloud project (dotted line) and EU Digital Ocean Forum (DOF) events.	12
Figure 5 EDITO-Infra overall architecture overview	15

TABLE OF TABLES

Table 1 EDITO-Infra feature set.....	14
--------------------------------------	----

1. Introduction

The European Digital Twin Ocean (EU DTO) is a transformative initiative aimed at harnessing the power of digital technology to provide comprehensive knowledge and understanding of the ocean. It was announced by President von der Leyen at the One Ocean Summit in Brest in February 2022 as part of the EU Mission "Restore our Ocean and Waters by 2030." The EU DTO seeks to make ocean knowledge readily accessible to citizens, entrepreneurs, scientists, and policymakers through innovative user-driven tools and visualisation platforms.

At its core, the EU DTO is envisioned as a public service and a public good, leveraging existing European science and assets to provide consistent, high-resolution descriptions of the ocean. This includes capturing the physical, chemical, biological, socio-ecological, and economic dimensions of the ocean, with forecasting capabilities ranging from seasonal to multi-decadal scales.

The European Digital Twin Ocean will consist of a core DTO providing as a baseline a huge bulk of data, generic ocean models and AI processors as toolboxes, on top of which a multitude of tailor-made applications, or 'local twins', can be plugged in. These applications will answer individual priorities by and for public users and industry, as well as national, local, or thematic policies or sector areas. The core European DTO will be a public good, connecting the physical, biological and socio-economic dimensions of the ocean.

To achieve its goals, a collaborative effort involving various stakeholders, including EU research projects, national initiatives, policymakers, and industry representatives, is needed.

Its core development is underway with the European Union (EU) funding through the European Digital Twin Ocean (EDITO) R&D projects. Two sister projects, namely EDITO-Infra and EDITO-Model Lab, are building the operational backbone infrastructure of the European Digital Twin Ocean.

Key components of the EDITO digital core infrastructure approach include a Data Lake, a Processing Engine, and a Virtual Simulation Environment. These components will facilitate the aggregation, processing, and visualisation of ocean data, enabling users to explore, analyse, and derive insights from diverse datasets.

The development of the EU DTO follows an inclusive and open process, with stakeholders actively engaged in co-designing and co-developing the initiative. Platforms like the Digital Ocean Forum allow stakeholders to contribute their expertise, insights, and feedback to shape the EU DTO's functionalities, assets, and specifications. Additionally, the EU DTO aims to align with other international initiatives, such as the UN

Decade of Ocean Science for Sustainable Development and Destination Earth, to promote interoperability and collaboration in advancing ocean knowledge worldwide.

Overall, this initiative represents a pioneering effort to harness digital innovation for the sustainable management and conservation of ocean resources. By providing actionable knowledge and supporting informed decision-making, the EU DTO aims to contribute to the restoration and preservation of marine ecosystems, promoting a sustainable blue economy, and mitigating climate change impacts on the ocean.

2. DTO task force in BlueCloud2026

The DTO (Digital Twin of the Ocean) Task Force within the Blue-Cloud 2026 project serves several key objectives:

- **Alignment with EU Digital Twin Ocean Initiative:** The task force aims to align the Blue-Cloud 2026 project with the broader EU Digital Twin Ocean initiative, which aims to establish a digital infrastructure for oceans and freshwater within the Mission restore our Ocean.
- **Integration of Blue-Cloud Data Lakes with DTO Developments:** The task force focuses on tuning Blue-Cloud data lakes with DTO developments, ensuring interoperability and facilitating dynamic data exchanges between Blue-Cloud and the evolving DTO infrastructure. This involves regular dialogue and cooperation between the two initiatives.
- **Mobilisation of Additional Data Resources:** Blue-Cloud 2026 contributes to the DTO by mobilising and making available major additional data resources beyond those managed by existing initiatives such as Copernicus Marine and EMODnet.
- **Provision of Virtual Labs:** Blue-Cloud 2026 develops and operates a series of Virtual Labs with analytical workflows which could be complementary and of added value for the EU DTO collection of analytical applications.
- **Facilitation of Technical Synergies:** The task force seeks to identify and leverage technical synergies between Blue-Cloud 2026 and DTO, including topics like semantic brokering, harmonisation of data lakes, optimisation of data access services,
- **Establishing dynamic exchange:** the taskforce will analyse and evaluate whether a smart connectivity could be established between the e-infrastructures of Blue-Cloud and EDITO as 'porting' of applications and data packages in operational practice will hinder the co-creation and co-production processes.

- Promotion of Sustainable Investments: By aligning with the DTO initiative, Blue-Cloud 2026 aims to make sustainable investments in marine data infrastructure, ensuring the long-term viability and usability of the developed services and resources.
- Community Engagement and Collaboration: The task force fosters a collaborative community of practice by establishing a sustained dialogue with stakeholders across the EU ocean and marine data management landscape. This includes engaging with relevant EU initiatives, policymakers, and key target audiences to ensure alignment with broader policy objectives and uptake of key exploitable results.

Overall, the DTO Task Force plays a critical role in ensuring that the Blue-Cloud 2026 project contributes effectively to developing the EU Digital Twin Ocean initiative, promotes sustainable investments in marine data infrastructure, and fosters collaboration and engagement within the marine science and technology community.

3. DTO task force activities

The DTO Task Force within the Blue-Cloud 2026 project has been active since the start of the project, beginning with a presentation on the EDITO project at the Blue-Cloud 2026 Kick-Off meeting. This initial step aimed to establish an ongoing open dialogue between the two projects, focusing on interoperability and collaboration to maximise the EU's contribution to the development of digital twins of the ocean.

Subsequently, a dedicated workshop was organised in February 2024 to facilitate technical and strategic exchanges between the Blue-Cloud 2026 and EDITO projects.

The workshop saw significant participation, with 26 attendees representing different WPs and tasks involved in the DTO Task Force in BlueCloud2026, as well as representatives from EDITO-Infra, EDITO ModelLab, and the EU Commission's RTD and MARE directorates (see the full list in **Error! Reference source not found.**).

	Name Surname	Affiliation	Presence (in presence – remote)
1	Nicolas Segebarth	EC DG RTD	remote
2	Zoi Konstantinou	EC DG MARE	remote
3	Sara Pittonet Gaiarin	Trust-IT & Blue-Cloud2026	Remote
4	Alessandra Giorgetti	OGS	Remote
5	Simona Simoncelli	INGV	remote
6	Conor Delay	SSBE	remote
7	Tjess Hernandez	VLIZ	remote
8	Paul Weerheim	MARIS	In presence
9	Dominique Obaton	IFREMER	remote
10	Robin Kooyman		In presence
11	Quentin Gaudel	MOI	In presence
12	Simon Lyobard		remote
13	Mary Malicet	MOI, EDITO-Model Lab project	remote
14	Sebastian Luna-Valero	EGI Foundation, BlueCloud-2026	Remote
15	Marina Tonani	MOI	In presence
16	Dick Schaap	MARIS	In presence
17	Pasquale pagano	CNR-ISTI	In presence
18	Massimiliano Assante	CNR-ISTI	In presence
19	Alain Arnaud	MOI	In presence
20	Julia Vera	SSBE	In presence
21	Patricia Cabrera	VLIZ	In presence
22	Samuel Fooks	VLIZ	In presence
23	Erwan Bodere	IFREMER	remote
24	Simon Van Gennip	MOI	In presence
25	Mary Malicet	MOI	In presence
26	Rita Giuffrida	Trust-IT & Blue-Cloud2026	Remote

Figure 1: List of participants attending the DTO Task force Workshop

Figure 2 gives the agenda for the first day of the Workshop.

9:30	Welcome at Mercator (Alain) and objective of the meeting (Marina)
9:40	EU -DTO (Nicolas Segebarth, DG RTD and Zoi Konstantinou, DG MARE). The EU commission perspective - minutes
10:10- 11:00	EDITO Infra short presentation + platform (MarinaTonani and Quentin Gaudel) + questions/discussion - minutes
11:00 -11:20	Coffee break
11:20-12:20	Data provision. Blue-Cloud presentations + questions/discussion
	Blue-Cloud federated Data discovery and access- MARIS
	Blue-Cloud development for EOVS Workbenches- Ifremer
12:20-14:00	Lunch break
14:00-15:00	Data Lake
20 min	EDITO Data Lake - MOI (Quentin) and VLIZ
20 min	Blue-Cloud BEACON approach - MARIS
20 min	Questions/discussion
15:00-15:30	VRE Environment (CNR + VLIZ) and Virtual Labs
15:30-16:00	Question/discussion
16:00-16:20	Coffee Break
16:20- 16:50	EDITO – Model Lab models and tools + questions /discussions - MOI Yann
16:50-17:20	EDITO- Model Lab ex of one demonstrator + questions/ discussion–MOI Simon
17:20-17:30	Wrap-up

Figure 2: Agenda first day of DTO task force Workshop

On the first day, managers and developers of the projects presented and demonstrated their developments for Data Lakes, the Virtual Research Environment, and Analytical Workflows (VLabs) and explained the underpinning architectures for the e-infrastructures. Questions and answers gave more technical and content insights and created a level playing field, whereby the impression arose that the technical basis of both projects was already quite aligned. Also, it appeared that both projects could learn from each other and, by regular dialogue and cooperation, could progress on various topics, like data interoperability, technical standards, user interfaces, visualisations, AAI, and others.

Technical details on EDITO architecture are described in the ANNEX at the end of this document.

9:00	DOF 2 outcome – co-construction process (Zoi or Nicolas TBC)
9:30-10:40	Blue-Cloud contribution to EDITO (chaired by Marina)
10:40- 11:00	Coffee break
11:00 -12:00	Roadmap for the Blue-Cloud 2026 towards EU DTO (chaired by Marina)
12:00-12:30	Wrap up and meeting closure

Figure 3: Agenda second day of DTO Task Force Workshop

The last half-day was dedicated to discussing Blue-Cloud 2026 possible contributions and topics for synergies with EDITO and initiating a roadmap for future collaboration (see **Error! Reference source not found.**). Further discussions took place on items from the previous day, clarifying the remaining questions and increasing understanding. Encouraged by the EU representatives, it was also discussed how Blue-

Cloud 2026 could demonstrate its cooperation and potential added value for the EU DTO at the coming Digital Ocean Forum 2024 (DOF 2024), which will be held on 12 and 13 June 2024 in Brussels, Belgium. This is an important event, where the progress with EDITO will be presented to many high-level EU attendants and representatives of tens of EU research projects, who are seen as potential stakeholders. It is planned that three examples of synergy will be presented in order to highlight how also other projects and initiatives could engage with EDITO for the co-creation and co-production of the EU DTO. The discussion resulted in the choice to deploy a copy of a successful Blue-Cloud Virtual Lab, namely for Zoo & Phytoplankton EOVS products, at the EDITO infrastructure. The choice of 'porting' was made because of the short lead time to the DOF2024 event, and it seemed feasible as VLIZ is the core developer of the Blue-Cloud VLab and is also a core partner in EDITO-Infra. For the longer term, porting is not seen as an effective method as it might give various complications in practice, such as different versions of used software libraries and missing connectivity to specific data storage spaces. Also, considering that Blue-Cloud 2024 might offer an increasing collection of analytical workflows (VLabs) and dynamic Data Lakes, establishing smart connectivity between the Blue-Cloud and EDITO e-infrastructures should be a better choice for the near future.

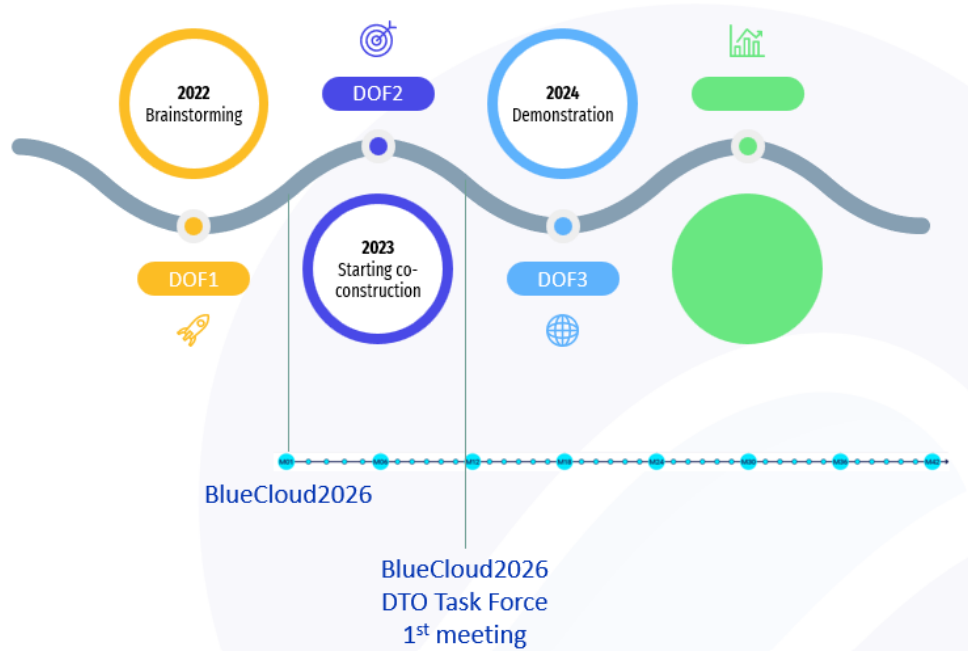


Figure 4: Blue Cloud project (dotted line) and EU Digital Ocean Forum (DOF) events.

It was concluded that the ongoing dialogue between Blue-Cloud 2026 and EDITO should continue to identify and elaborate synergies and collaboration opportunities. EDITO Infra will keep Blue-Cloud 2026 regularly informed about platform evolution, and both sides will promote technical exchanges to test and understand interoperability levels, constantly increasing the synergies between these two important initiatives.

3.1. Next steps

The direct next steps for Blue-Cloud 2026 involve active engagement in preparing its contribution to the Digital Ocean Forum 2024. This effort will strengthen synergies with the EU DTO and provide initial feedback on potential technical complexities related to integrating and interoperating VRE demonstrators into the EDITO_Infra platform.

Therefore, ongoing dialogue and technical exchanges are required to identify in more detail the added value that Blue-Cloud 2026 could offer to the EU DTO data and services portfolio and to elaborate on how smart connectivity could be established between the Blue-Cloud digital ecosystem and the EDITO infrastructure to achieve a dynamic exchange. These discussions and analyses should become part of a roadmap for co-construction between Blue-Cloud 2026 and EDITO, that then could be proposed to the EU DTO management team.

APPENDIX on EDITO-Infra Architecture

1.1. EDITO-Infra project

The main aim of the “EU Public Infrastructure for the European Digital Twin Ocean (EDITO-Infra)” project is to build the public infrastructure backbone for the European Digital Twin of the Ocean. This will be done by upgrading, combining, and integrating key service components of the existing EU ocean observing, monitoring and data programs, namely “Copernicus Marine Service” and the “European Marine Observation & Data Network (EMODnet)” into a single digital framework. In that context, EDITO-Infra is the both the name of the project and the name of the platform that will be deployed. As such, EDITO-Infra will provide the foundation for the further development of the European Digital Twin Ocean, hosting the deployment of multiple digital twin applications from ongoing and future digital twin projects, supporting the deployment of new generation of ocean models (i.e., EDITO-ModelLab) and other related initiatives, including Horizon Europe “Mission Lighthouses” projects.

1.1.1. EDITO-Infra architecture

EDITO is a European platform that serves as a centralised hub for searching, exploring, and utilising products and knowledge to support digital twins of the ocean. Specifically:

- It aggregates data from other European platforms, including Copernicus Marine Service and EMODnet.
- It comprehensively indexes all available ocean data products.
- It presents data through the STAC API, facilitating metadata-driven and semantic search queries.
- users can incorporate new data into the platform either directly (by uploading) or by deriving new data from existing sources through computational functions.
- All processes offered by the platform adhere to the OGC (Open Geospatial Consortium) API processes standard, ensuring smooth integration with third-party systems.

Following this definition, the consortium converged on the following features, which were presented at Digital Ocean Forum 2 in 2023.

The EDITO-Infra platform exists to fulfil the following feature sets for any end-user:

	EXPLORE USE THE DIGITAL TWIN OCEAN PLATFORM	CREATE INDUSTRIALIZE YOUR EXTERNAL THIRD- PARTY PROJECT SERVICES	CONTRIBUTE SHARE DATA, SERVICES AND PROCESSES TO THE DIGITAL TWIN OCEAN PLATFORM
DATA	Access the largest EU Marine Portfolio across many Topics including data from the Copernicus Marine Service and EMODnet services. Import external data and benefit from the platform functionalities.	Interface your project to our data lake to build your project fast data access.	Share your data so that it will be added and referenced in the data lake.
CLOUD AND HPC COMPUTING RESOURCES	Benefits from EU DTO own cloud capabilities and HPC computing provided through EuroHPC infrastructures. Optimize your performances thanks to near-data computing.		
OCEAN MODELS	Launch pre-defined ocean analysis and model forecasts on the platform. Derive all our ocean models, create and launch your own models on the platform.	Integrate your project into the platform ecosystem to validate your models.	Share your ocean models and results so that it will be added to the Digital Twin Ocean
SERVICES	The platform comes with numbers of out of the box services including, a smart viewer and tools for development, processing and analysis. Use any of the services for your own needs.	Automatize and test your workflow thanks to the shared out of the box tools. Integrate to your project other services available on the marketplace. Host and make your project services accessible through the platform.	Propose your services in the service marketplace.
ON-DEMAND PROCESSES WHAT-IF-SCENARIOS	Remotely call pre-defined on-demand processes and run what-if-scenarios. Create and run your own processes and what-if scenarios.	Integrate on-demand processes and what-if-scenarios results to your project. Schedule or program their runs.	Reference your on-demand processes and what-if-scenarios to the process registry.

Table 1 EDITO-Infra feature set.

The EDITO platform comprises three primary elements: the EDITO Data Lake, the EDITO Engine, and the EDITO Service (Table 1 EDITO-Infra feature set.). These components interdependently contribute to satisfying the design criteria.

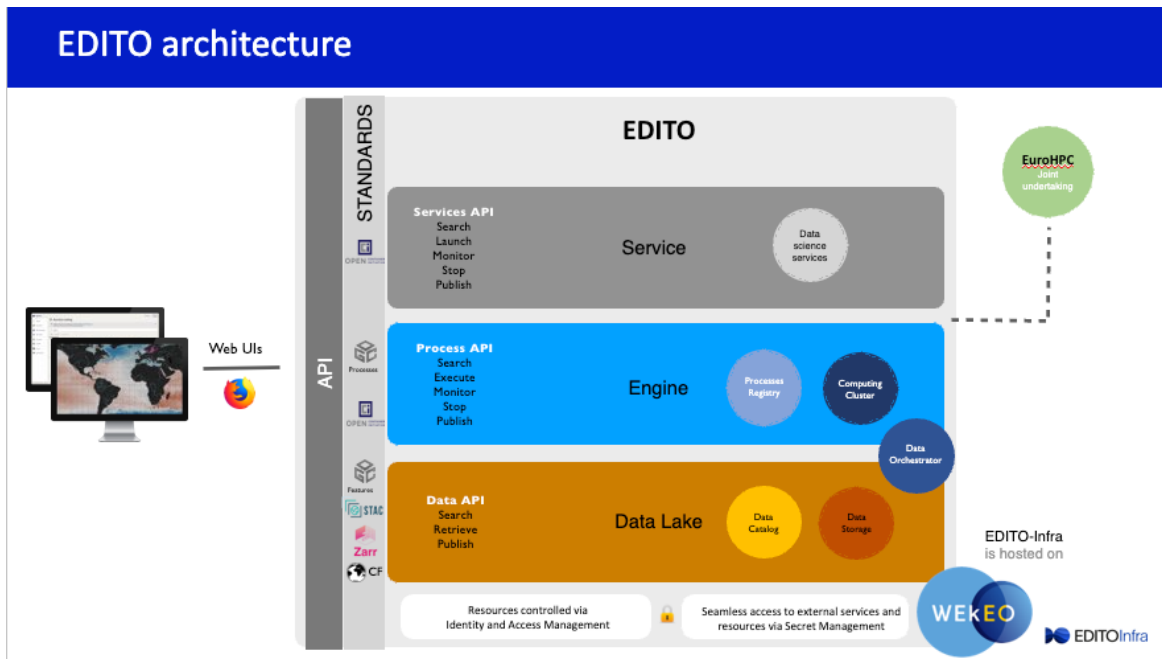


Figure 5 EDITO-Infra overall architecture overview

The "EDITO Home" page is the central hub of the EDITO platform, providing users with a cohesive structure for intuitive navigation. It begins by introducing users to the platform's overarching capabilities, ensuring a clear understanding of the diverse features available. Three main parts of the platform are highlighted:

- **The tutorials section**, offering users a structured and comprehensive learning path. Covering platform essentials, tools, and best practices, this section is designed to empower users with the resources needed for proficiency.
- **The datalab section**, dedicated to facilitating the design of digital twin ocean applications. It provides an interactive environment with powerful tools for seamless integration, simulation, and robust application development.

- **The data viewer**, that serves as a vital resource for users interested in exploring oceanographic data. Equipped with sophisticated visualization tools, this section allows for in-depth exploration of diverse datasets, including real-time and historical oceanographic information.

To improve user experience, the "EDITO Home" page is equipped with a responsive design to ensure accessibility across various devices. This focus on a user-centric design emphasises efficiency and engagement, positioning the "EDITO Home" page as an ideal starting point within the platform.