# PHIDIAS
Prototype of HPC/Data Infrastructure for On-demand Services

# PHIDIAS BOOSTING THE USE OF CLOUD SERVICES
## To benefit marine data management, services and processing

Through the development of **3 Use Cases**, Phidias HPC will develop data post-processing methods coupled with HPC capabilities, which will be deployed as a service for several end-users (including scientific communities, Public authorities, private players and citizen scientists).

Intelligent screening of satellite data

Big data earth observations (EO)
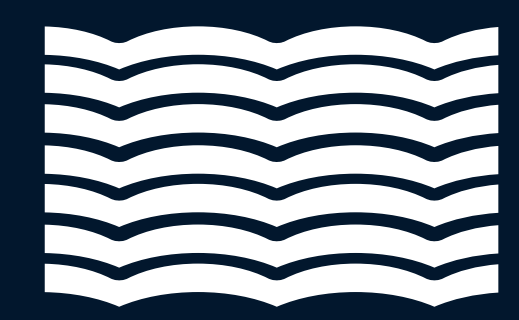
Ocean

## Observing the ocean is challenging:
missions at sea are costly, different scales of processes interact, and the conditions are constantly changing.

This is why scientists say that "*a measurement not made today is lost forever*".

That is why it is fundamental to properly store both the data and metadata, to:

- Guarantee Data access for the widest community
- Align with the FAIR principles, to be:

**Findable**     **Accessible**     **Interoperable**     **Reusable**

Long-term data archiving procedures will be specified in the PHIDIAS use case in Ocean Data testing, relying on the HPC and HPDA (high-performance data analytics) expertise of PHIDIAS partners, aimed at improving the usage of cloud services for marine data management, data services to the user in a FAIR perspective, and data processing on demand.

### TASK 1 — Improving Long-term stewardship of marine in-situ data
Ensuring marine data are preserved within procedures that can be certified by the Research Data Alliance (Core Trust Seal Approval).

### TASK 2 — Improvement of data storage for services to users
Provide users with:

(1) Fast and interoperable access to data from multiple sources,

(2) Parallel processing capabilities within dedicated high-performance computing, using, for example, Jupyter notebooks or the PANGEO software ecosystem.

### TASK 3 — Marine data processing workflows for on-demand processing
Allow users to access data, software tools and computing resources in a seamless way to create added-value products, for example quality-controlled, merged datasets or gridded fields.

## Working Plan topics

**1** Service scalability: Limited data upload size

Solution: Other protocols e.g. Virtual File Systems, shared storage between infrastructures.

**2** Back-office exchanges: Increased exchanges between Data Centres.

Solution: iRODS (Integrated Rule-Oriented Data System)

**3** Securing long-term archive: Unsuitable data archiving and lack of dedicated staff

Solution: Professional long-term repositories and geographically distributed repositories.

**4** Fast access to datasets: Assembly of multidisciplinary datasets are difficult

Solution: Data Lake/ technical cache

**5** On-demand processing: Using specialised tools can be time-demanding

Solution: DIVAnd interpolation software tool (Deliverable 6.3.1)
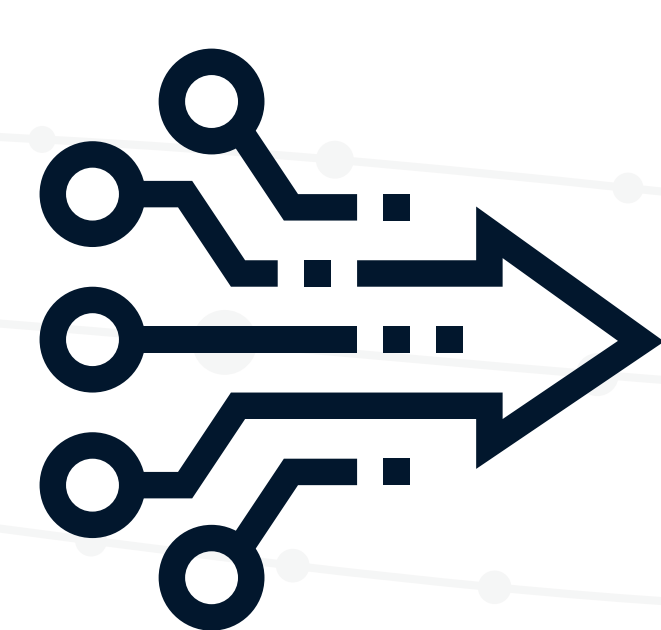
## Key result

**PHIDIAS** is supporting the EU's policy of open science and its goals will be pursued in line with the development of the European Open Science Cloud (EOSC) and the Copernicus Data and Information Access Services (DIAS).

Copernicus — Europe's eyes on Earth

EUROPEAN OPEN SCIENCE CLOUD

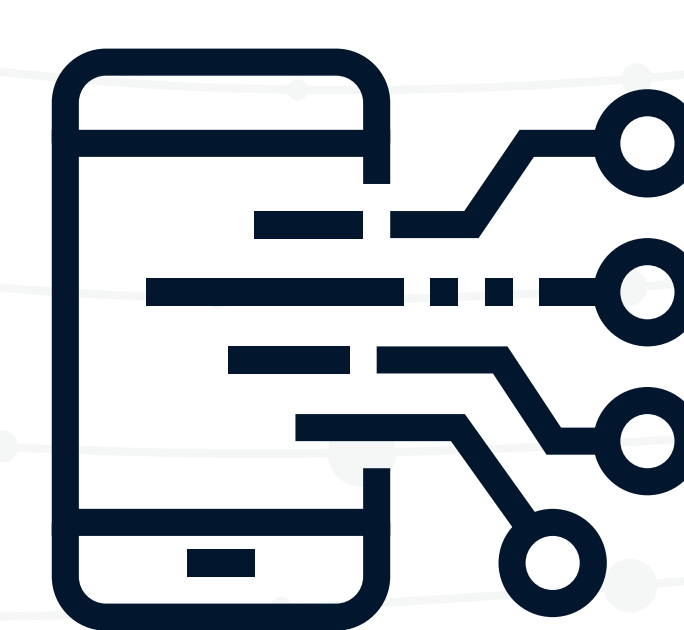**Improving the activities of the researchers and specialists.**

### Data publication
Seamless upload of large datasets, ensuring their long-term archiving and publishing them following standards, best practices and recommendations from Data Management groups.

Enhancing the ingestion of long-tail data, available to a larger community.

### Data access
Performing operations such as sub-setting (based on regions, parameters), quality-control, visualisation or spatial interpolation
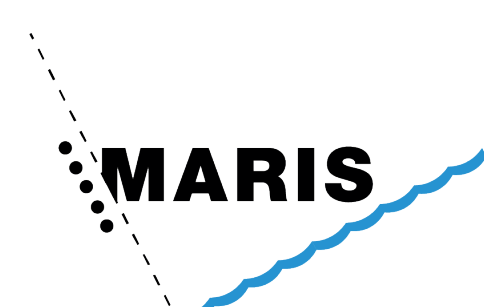
### Data processing
Performance of Spatio-temporal interpolation of large datasets.

North Atlantic Ocean and the Baltic Sea, 10 million observations for a total of approx. 250 GBytes.

The final product will consist of an inter-comparison of satellite data and in-situ data of sea surface salinity, including Inspire-compliant online services for data visualisation and access.

CINES · CERFACS · cnrs · CSC · GEOMATYS · Ifremer · IRD Institut de Recherche pour le Développement · MARIS · NEOVIA INNOVATION · SPASCIA · SYKE · Trust-IT Services Communicating ICT to markets · LIÈGE université

More information

phidias-hpc.eu     @PhidiasHpc     /company/phidias-hpc