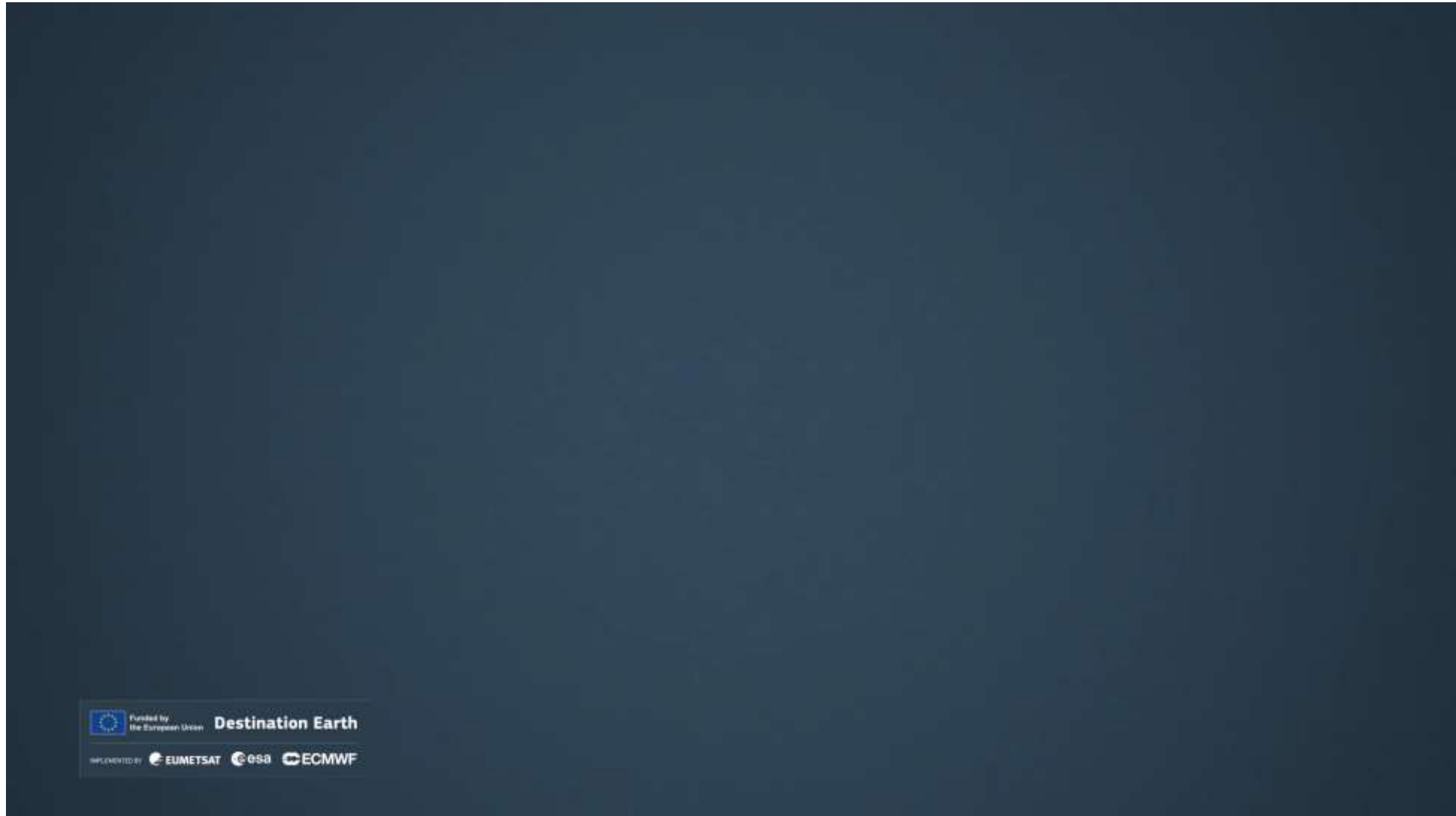


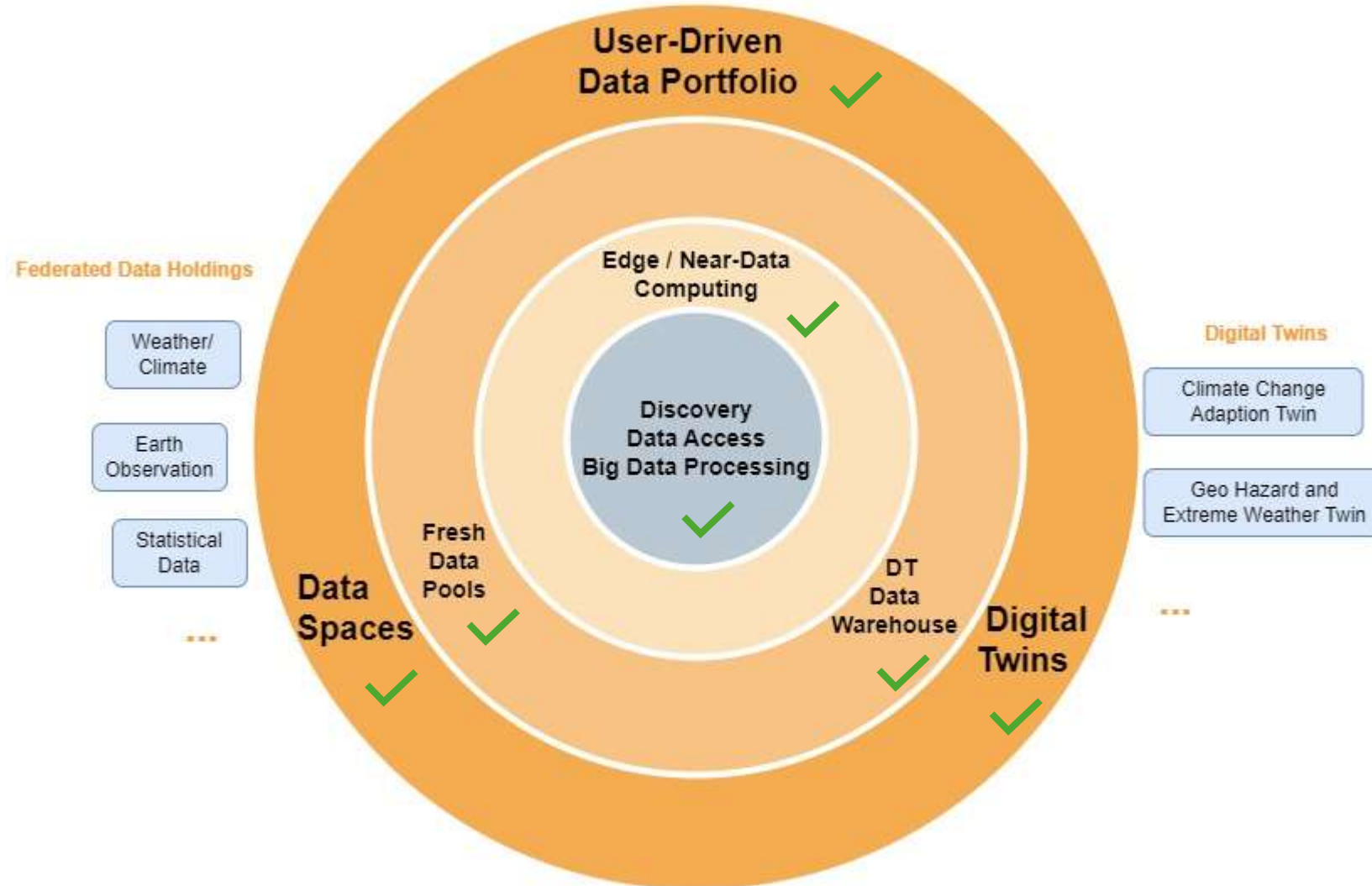
DestinE Data Lake characteristics and services

Lothar Wolf, EUMETSAT

A possible journey using the data lake



DestinE Data Lake – in a nutshell



DEDL is self-standing component

- Built from geographically distributed physical elements
- Distributed services – seamless access

Discovery & Data Access

- Harmonisation of data access (HDA) to simplify data discovery & access
- External federated data spaces
- Digital Twin data (ECMWF):
 - Extreme Weather and Climate Change Adaptation
- DestinE User generated data

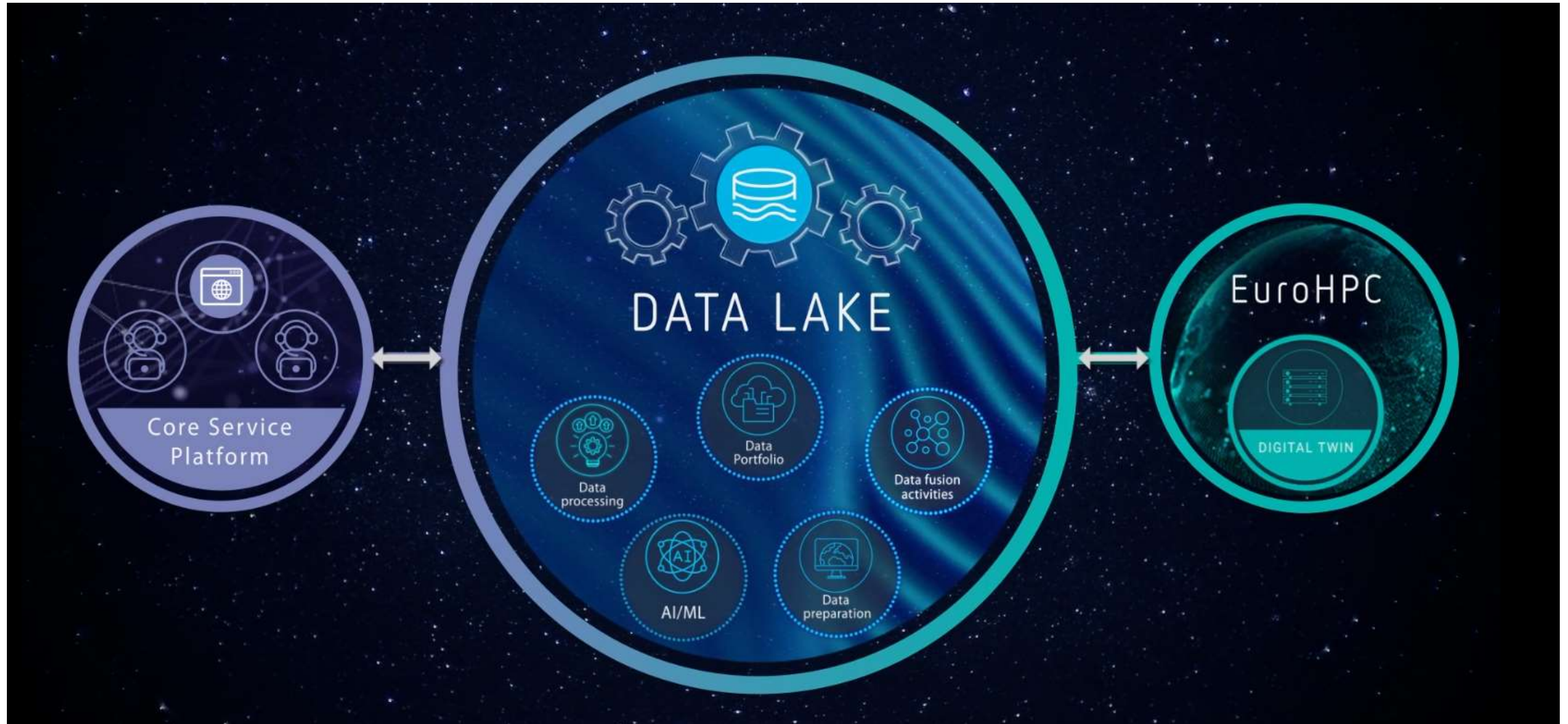
Big Data Processing

- Processing near data including distributed computing & workflows
- Supports & enables AI/ML applications

DestinE Data Lake Distributed Infrastructure



Data Lake services



DestinE Data lake services

HDA



Harmonised Data Access: seamless access to

- DestinE Data – DT Outputs & User Generated Data for DestinE
- Federated Data

as per defined & evolving “[DestinE Data Portfolio](#)”.

API => Spatio Temporal Asset Catalog (STAC)

Usage on request



STACK: SaaS suite which enables near data processing

- JupyterHub, Dask/ Dask Gateway and Open Data Cube



ISLET Compute: (IaaS/PaaS) enables near data processing by allowing users to manage and deploy virtualised workloads

ISLET Storage: S3 Object Storage to store user’s data and processing results



HOOK: allows to execute high level pre-defined or own workflows

- Data harvest: to harvest data from a federated data provider a priori of planned data processing or analysis task

How to access DEDL services

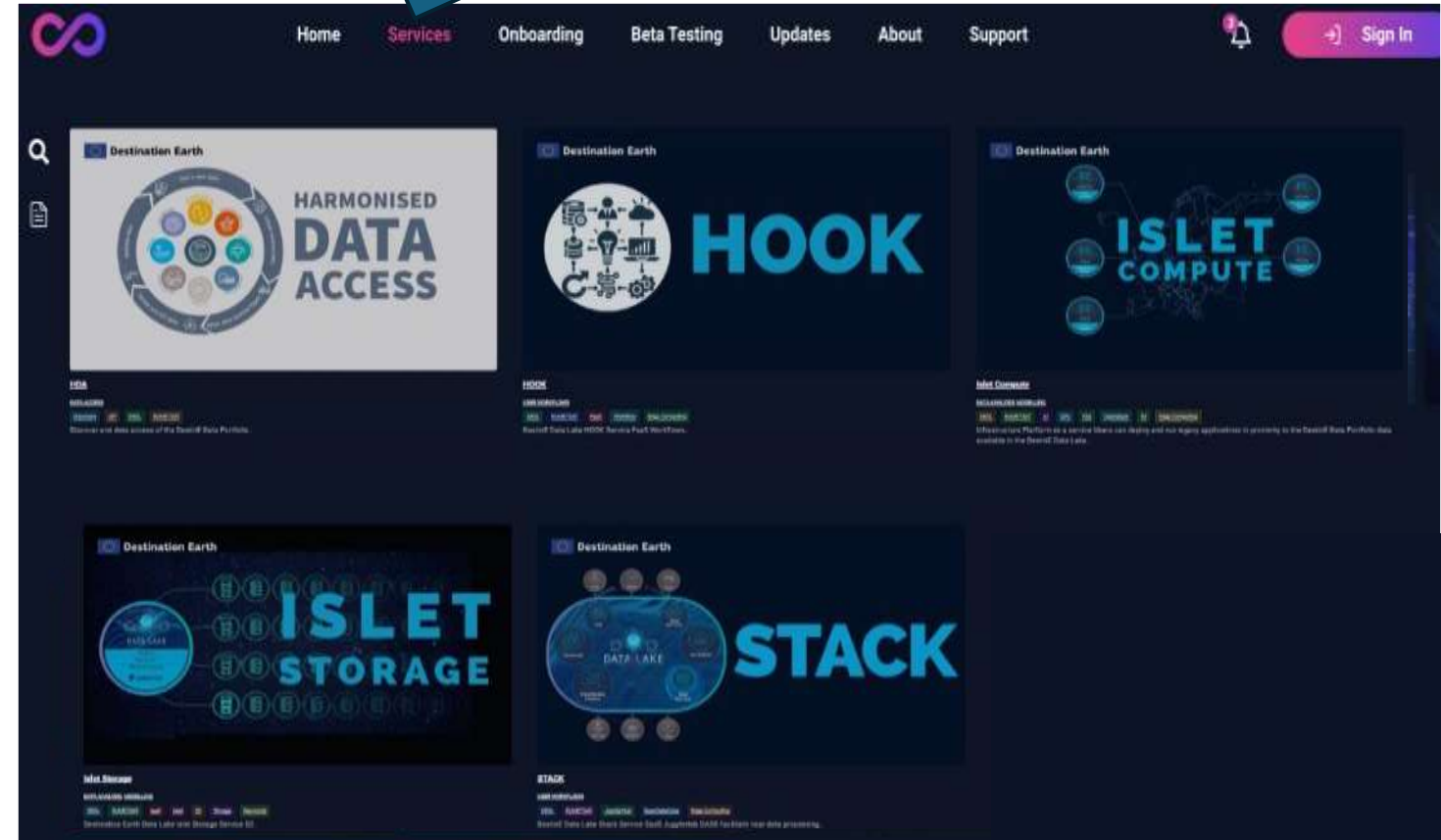
LANDING PAGE



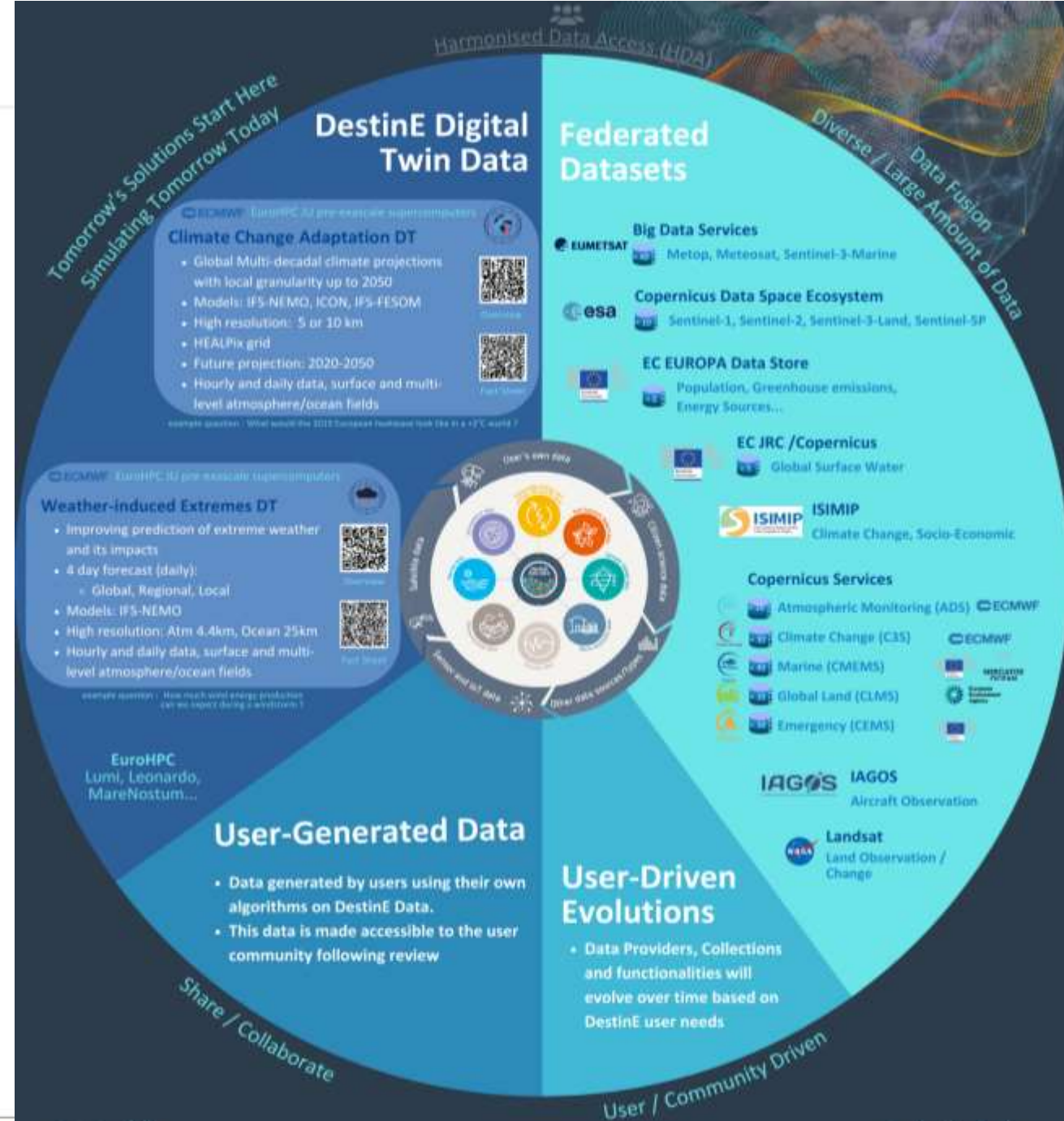
<https://platform.destine.eu>

- register/ sign in
- discover services
- work with HDA
- request access to edge services

SERVICE REGISTRY



DestinE Data Portfolio



DestinE use cases in DEDL

- Demonstrate usage of DestinE data and Data Lake services
- Demonstrators for DestinE future services

Use cases completed & delivered

Description/ Details available on DestinE community website

[Danube Delta Water Reservoir Monitoring](#) (Cloud Ferro, CS Group)

[Italy Drought 2022](#) (Cloud Ferro, EODC)

[Pakistan Flood 2022](#) (Cloud Ferro, EODC)

DestinE use cases in DEDL

Under development

Energy in Catalonia (CIMNE)

Development of machine-learning models for predicting electricity consumption in Catalonia at the postal code level based on extreme weather data and electricity distribution system operator datasets

Dust in Barcelona (CIMNE)

Implementation of a micro-scale pollutant (PM10) transport model using DestinE Data and DEDL services

Agriculture and Climate Change (Senckenberg, Oslo University)

Integration of Crop Wild Relatives model in DestinE using DestinE Data Portfolio and DEDL services

Algae Strom (RISE, SMHI)

Usage of DestinE Data from the portfolio in AI models with the aim to describe the timing and distribution of the spring bloom in the Baltic Proper Sea and to describe the development of cyanobacteria blooms

Forest Biodiversity (VTT, University of Jyväskylä)

Portage of Physically-based forest growth and productivity model and application of biodiversity model based on modelled forest structural variables and climate data in DestinE Data Lake

DATA LAKE: PHASE II KEY TASKS (1)

1

Build on Previous Accomplishments

- Operate, coordinate, maintain, and evolve DEDL baseline services
 - Enhance KPIs and SLAs
- Manage industrial service support contracts

2

Mature DEDL Services

- Achieve 24/7 operations level early in Phase II
 - 5x8 support for services
 - 24 x 7 NOC
 - high available
 - high redundant

3

Extend DestinE Data Portfolio

- Align with user needs from DestinE use cases and emerging end-user services
 - new federations, new datasets, harvest
- Evolve in line with partnerships and data governance policies (ensure data policy, terms & conditions, availability, SLAs)

4

Harmonise and Standardize Data Services

- Contribute to the overall harmonization and standardization of data services and APIs
 - alignment of HDA STAC implementation with other EO initiatives

DATA LAKE: PHASE II KEY TASKS (2)

**Enhance Near
Data Processing
Services**



**Evolve AI/ML
Preparedness**



**Support
Integration and
Interoperability**



**Support User
Engagement**





Destination Earth

**Questions
and
Answers**

Thank you!



Destination Earth

Funded by
the European Union



Implemented by

