Destination Earth and Data Lake

powering Europe's high precision climate and environmental modelling

Michael Schick (EUMETSAT)
OGC iDAYS 2025 – Session #5
10th of December 2025

An intergovernmental organisation with 30 member states

AUSTRIA
BELGIUM
BULGARIA
CROATIA
CZECHIA
DENMARK
ESTONIA
FINLAND
FRANCE
GERMANY
GREECE
HUNGARY
ICELAND
IRELAND
ITALY
LATVIA
LITHUANIA
LUXEMBOURG
THE NETHERLANDS
NORWAY
POLAND
PORTUGAL
ROMANIA
SLOVAK REPUBLIC
SLOVENIA
SPAIN
SWEDEN
SWITZERLAND
TURKEY
UNITED KINGDOM

EUMETSAT mission

Primary objective:

Establish, maintain and exploit European systems of meteorological satellites.

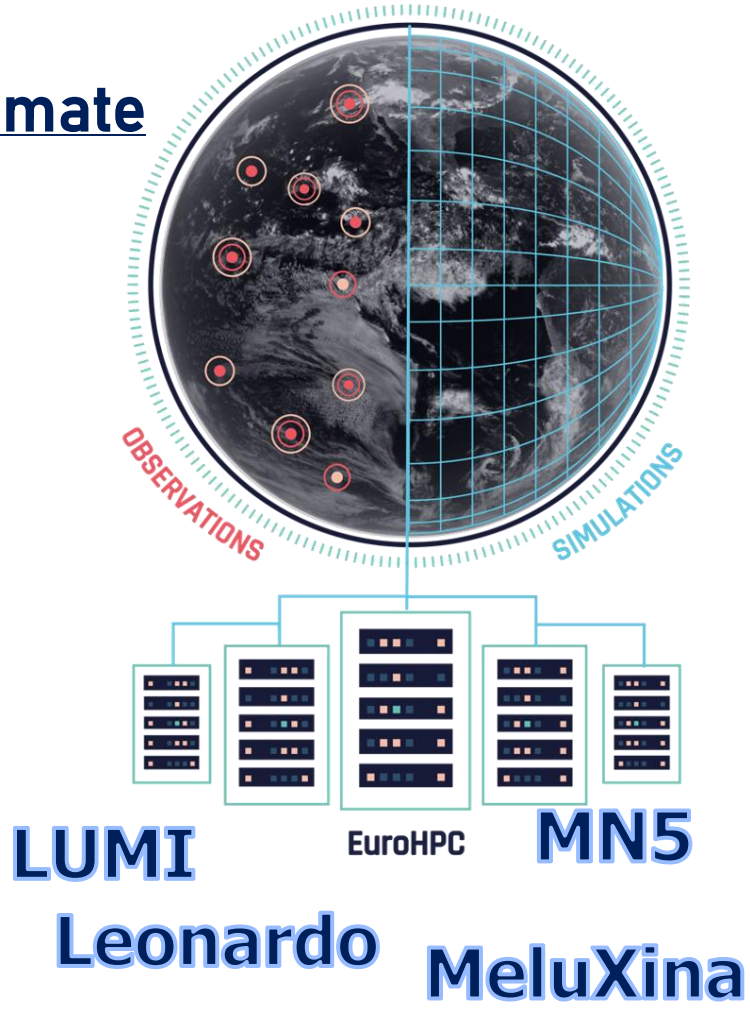
Further objective:

Contribute to the operational monitoring of the climate and the detection of global climatic changes.

Destination Earth

Flagship initiative of the European Commission to develop

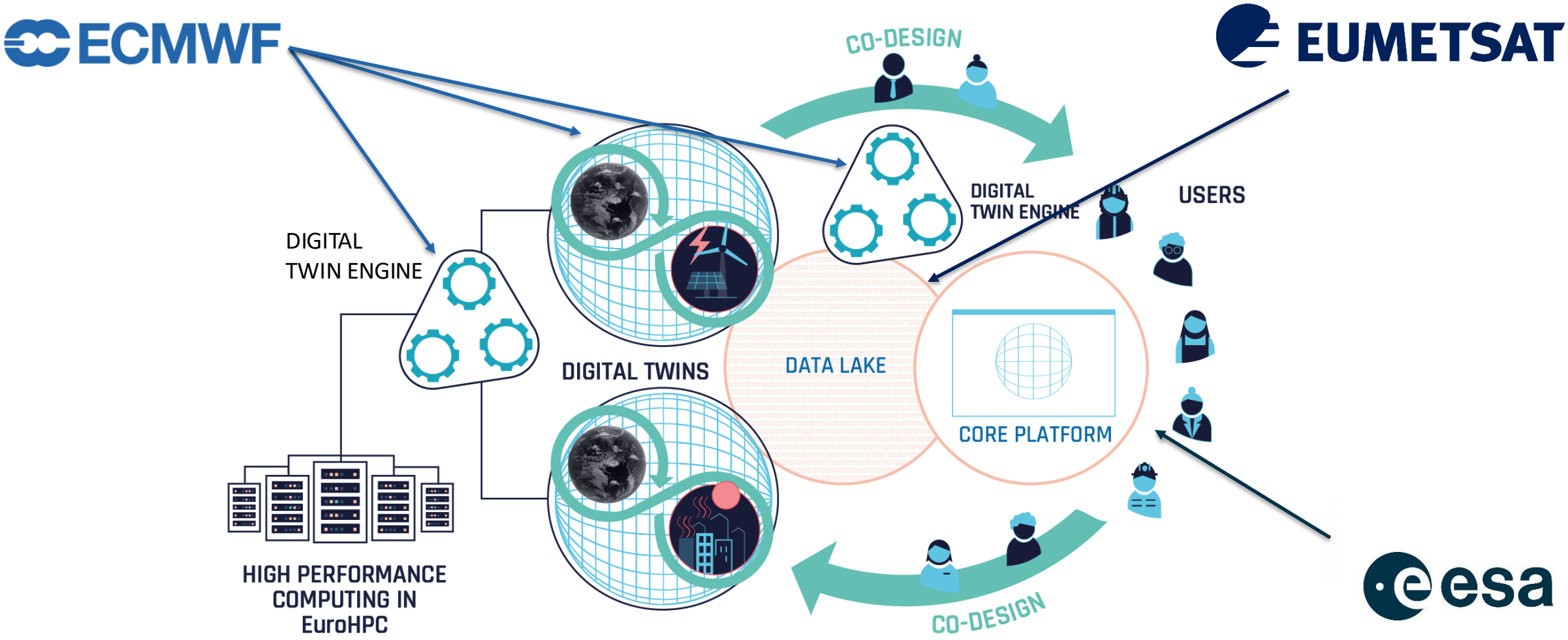
a digital replica of our planet to respond and adapt to climate change and extreme events



DestinE is a joint undertaking in strategic partnership with **EuroHPC**

- Establishes bespoke cutting-edge simulation capabilities
- Provides Earth-system information at scales where the impacts of **extreme** events and **climate change** are felt
- Fosters an innovative and thriving AI-enabled digital ecosystem
- DestinE Data is part of European Green Deal Data Spaces

A novel information system – implemented by the 3EEs



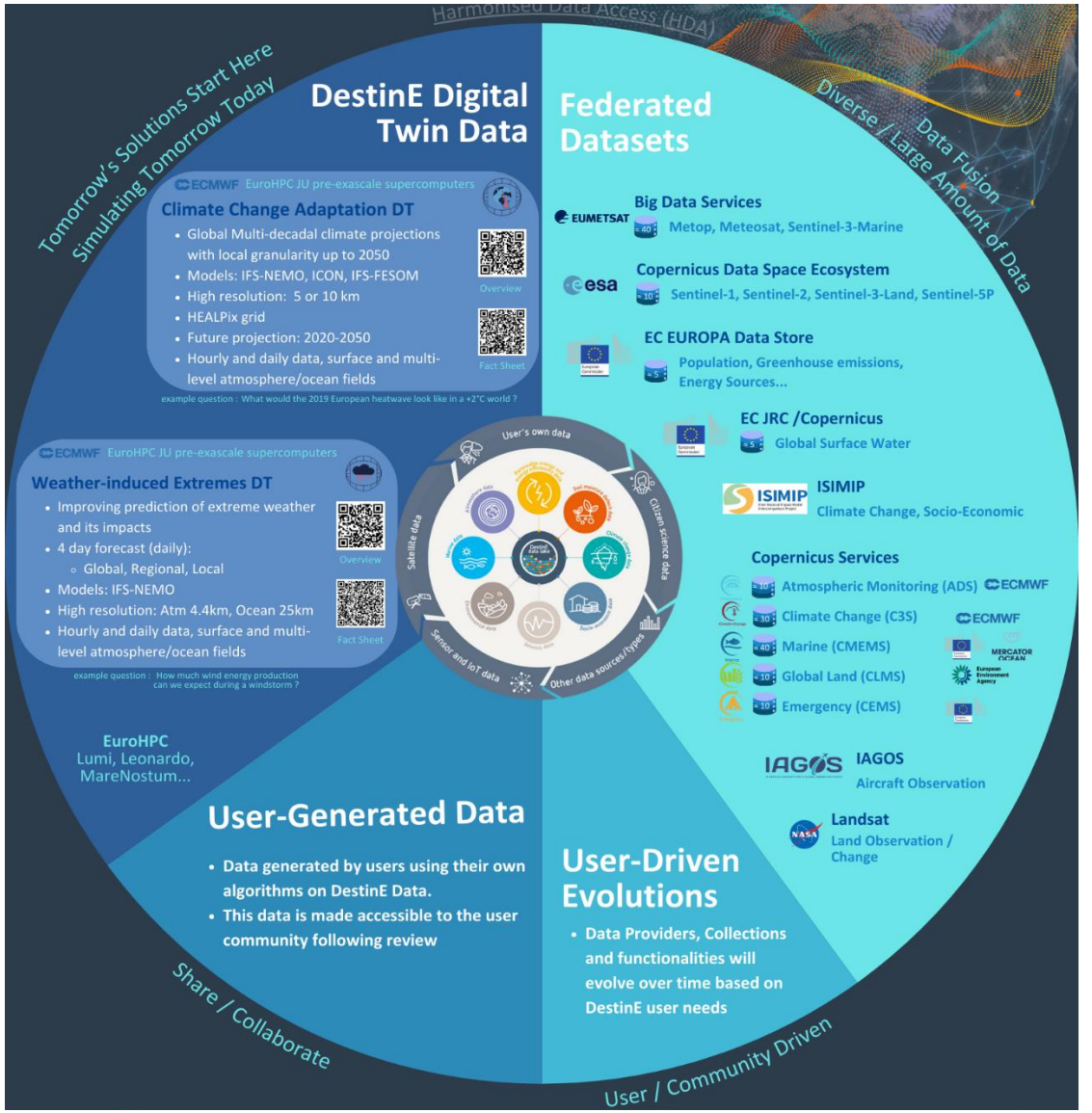
DestinE Data Portfolio

DestinE Data Lake offers a wide variety of data.

- Digital Twin Data
- Federated datasets from various data spaces, also beyond traditional Earth Observation
- User generated data
- User-driven datasets, based on DestinE user needs

This data portfolio is enabled by a harmonized data access solution (HDA) that abstracts away the heterogeneity and complexity of the underlying data sources

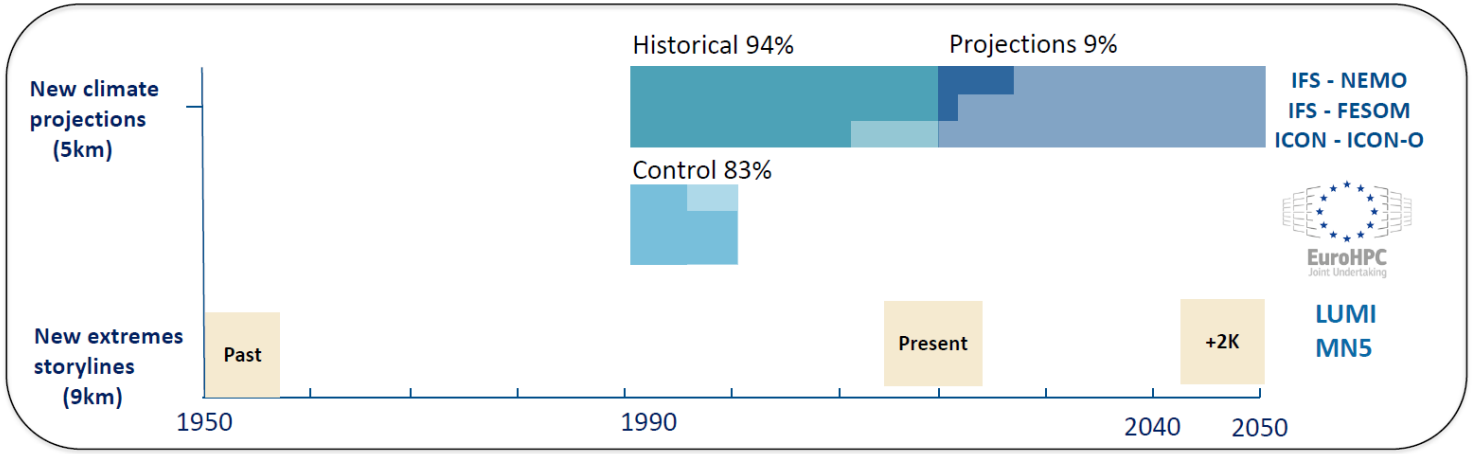
https://data.destination-earth.eu/data-portfolio



DestinE – Climate Change Adaption DT

CLIMATE DT

Defined development / evaluation / operational - suites: production started with o-suites



More info and full list of available simulations
https://destine.ecmwf.int/climate-change-adaptation-digital-twin-climate-dt/

ECMWF | DestinE Platform | CSC

KEY FEATURES

Km-scale Earth system models
Uses km-scale Earth system models, at 5 to 10 km resolution, to better represent critical processes like storms and ocean eddies.

Global multi-decadal projections with local granularity
Produces global multi-decadal climate projections with local granularity, up to 2050, by exploiting the EuroHPC pre-exascale supercomputers.

Tailored climate information
Tailors the climate information to match the needs of users from relevant impact sectors, through co-design and innovative data streaming and data handling techniques.

Routine and on-demand operational production of climate simulations
Establishes an operational infrastructure to produce climate simulations in support of adaptation activities, both routinely (yearly or less) and on-demand.

DestinE – Weather-Induced Extremes DT

EXTREMES DT : A MAGNIFYING GLASS ON EXTREME WEATHER EVENTS

Global and **daily** simulations of extreme weather **4 days** ahead at **4.4km**

On-Demand regional simulations **2 days** ahead at **750m** to **500m**

Impact-sector models: user-relevant information for societal impacts

- First comprehensive evaluation of global medium-range forecasts at 4.4 km in near real-time – demonstrating clear benefits at local scale (TC, orographic precipitation)
- End-to-end workflows for the regional, on-demand, component for selected configurations set up; including impact sector models for selected use cases

KEY FEATURES

Km-scale Earth system models

Uses "km-scale" models to better represent extreme weather events and deliver information at scales where their impacts are felt.

Routine production of global simulations

Produces global simulations at 4.4 km resolution for 4 days ahead to predict extreme weather events worldwide.

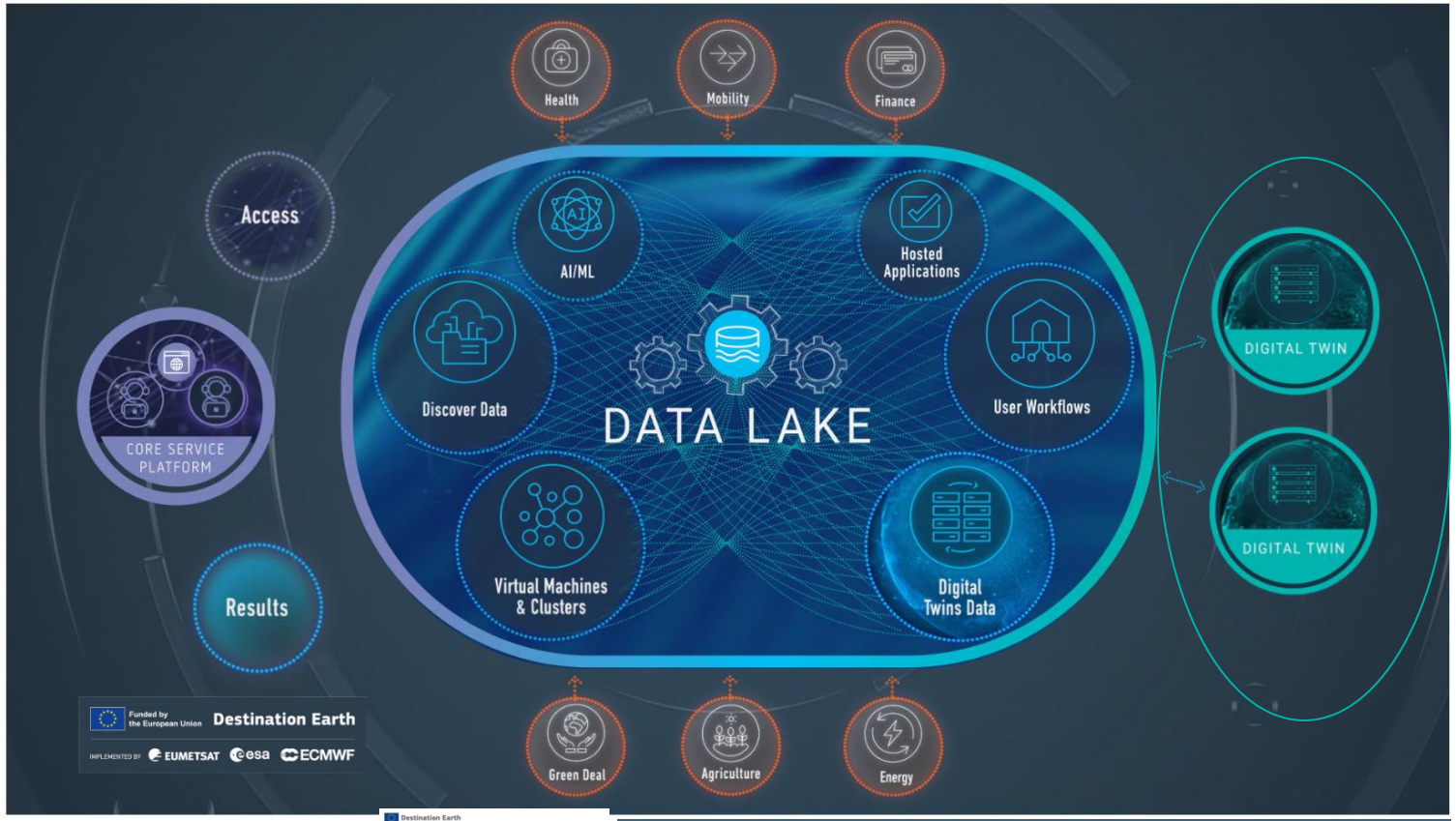
On-Demand refinement over Europe

Produces on-demand regional simulations at 750 to 500 m resolution for 2 days ahead to refine the representation of extreme events occurring over Europe.

From weather to impact-sector information

Integrates impact sector models in the Digital Twin workflows, to transform meteorological data into user-relevant information.

DestinE Data Lake in a nutshell



DEDL is central component with three essential pillars

Distributed infrastructure

- Implementing near-data processing
- Data Bridges close to EuroHPC sites that host the ECMWF Digital Twins and big data providers such as EUMETSAT
- Storage of DestinE data

Discovery & Data Access

- Harmonisation of data access (HDA) to simplify data discovery & access
- External federated data spaces
- Digital Twin data (ECMWF):
- DestinE User generated data

Big Data Processing Services and Tools

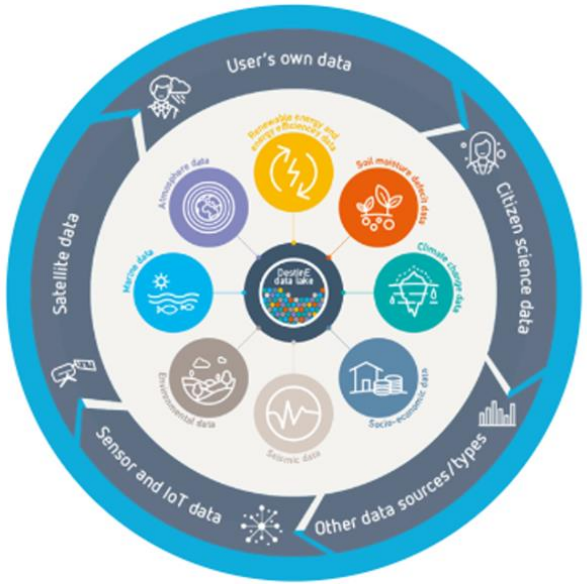
- Processing near data including distributed computing & workflows tools and services (i.e. ISLET)
- Supports & enables AI/ML applications

Harmonised Data Access

Around **200** collections

One credentials

Various data providers



HARMONISED DATA ACCESS

DT collections

Discovery and access data

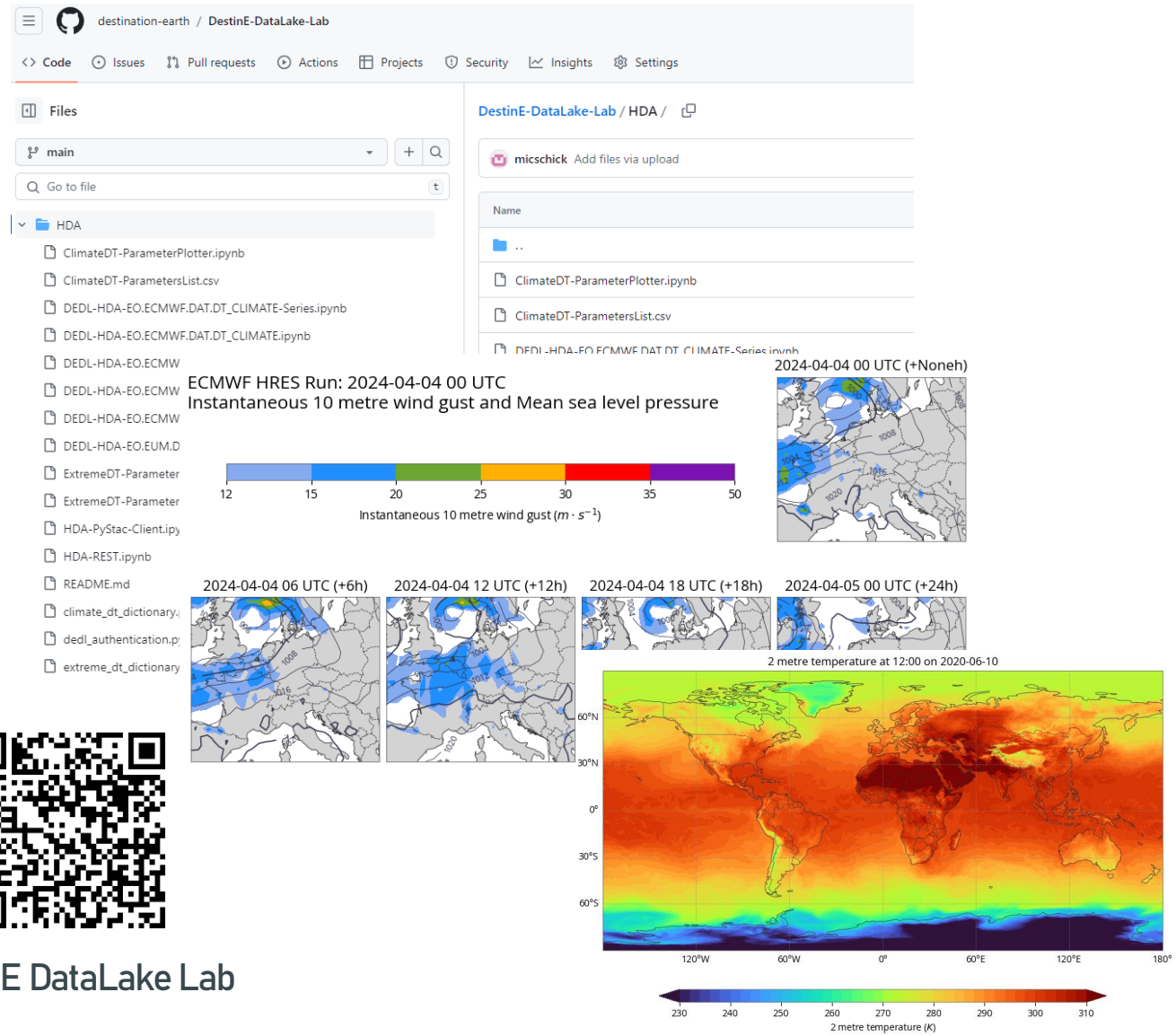
STAC

https://github.com/destination-earth/DestinE-DataLake-Lab

Harmonised Data Access

- Harmonised Data Access (STAC)
  - REST API client
  - PyStac-Client
  - EODAG (simple python API)
- Access DestinE Data Portfolio
- Datasets from multiple provider
- Protected datasets **(DT Data)**
- Quotas Group **based**

- Notebooks DestinE-DataLakeLab
  - From Insula, STACK
    - Climate DT plotter
    - Extreme DT plotter
    - HDA-PySTAC-Client
    - ....



DestinE DataLake Lab

DestinE Data Lake Cube

- Design and development of a new service exposing EUMETSAT and Copernicus datasets via a Cube interface (compliant to **OGC GeoDataCube** SWG recommendations)
- MVP supporting a subset of datasets to be delivered in June 2026
- Mixed approach: conversion of some datasets to Zarr + virtualized Zarr-like access to others
- Near-data processing capabilities: system and user-defined workflows
- Streaming API for efficient data loading in ML applications
- Service deployed and operated on EUMETSAT Data Bridge

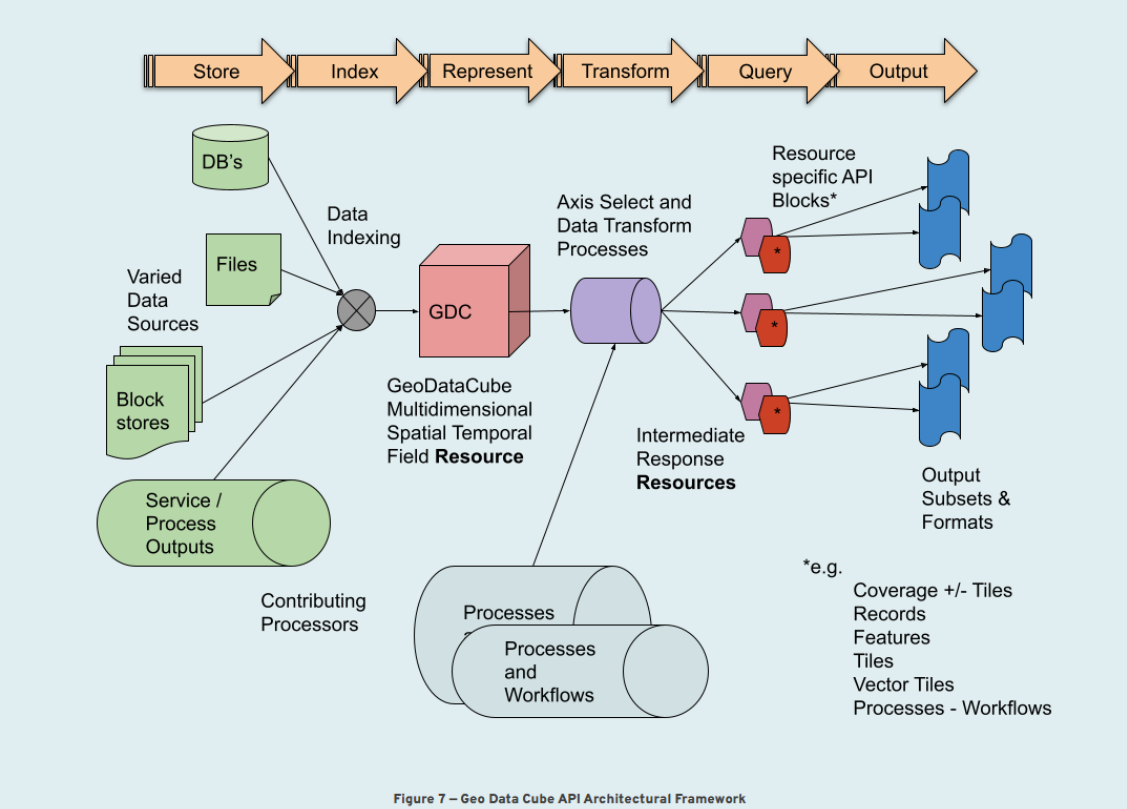
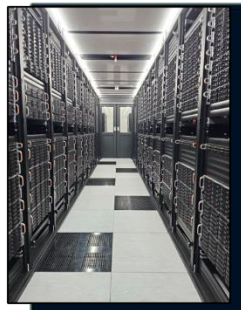


Figure 7 – Geo Data Cube API Architectural Framework



Data Lake distributed infrastructure

LUMI HPC — FINLAND, Kajaani — June 2023 — Nominal

Leonardo HPC — ITALY, Bologna — January 2024

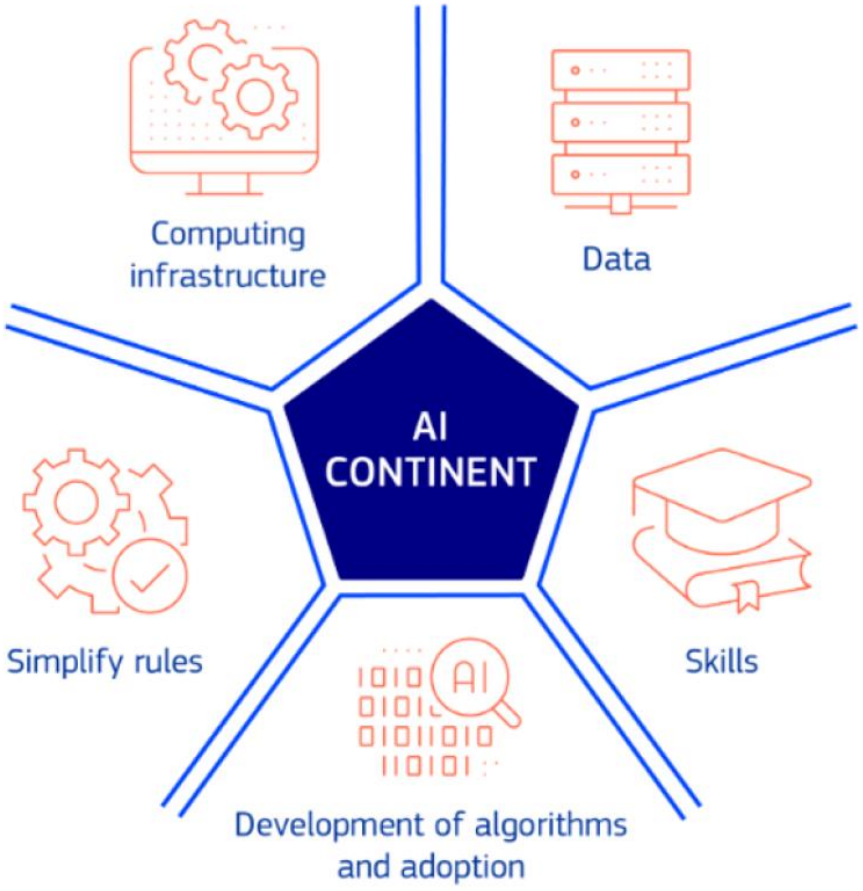
Mare Nostrum HPC — SPAIN, Barcelona — March 2025

EUMETSAT Data Center — GERMANY, Darmstadt — Dec 2024

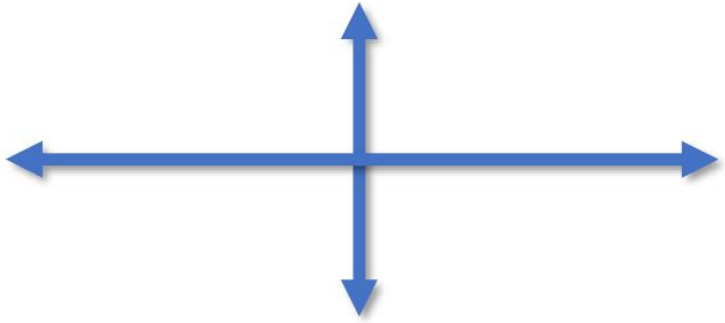
CloudFerro Data Center — POLAND, Warsaw — September 2022

Totals	
Storage (PB):	120
CPUs:	20450
GPUs:	66
Datasets:	+200

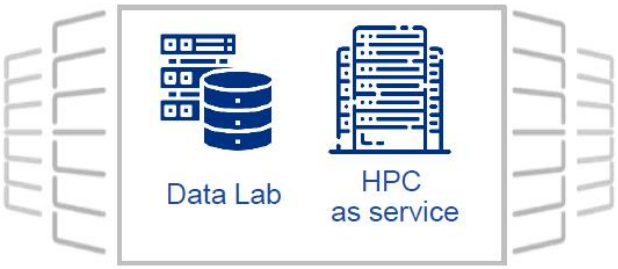
DestinE + AI Factories + DataSpaces



Destination Earth



AI Factory



Data Spaces

Data Spaces Considerations in Context of DestinE

Several initiatives on connecting Data Spaces

- DestinE, EOSC, CDSE, WEkEO ...
- SIMPL => Smart Middleware Platform

from “FAIR/O”

- Finable, Accessible, Interoperable, and Reusable /open license

towards “FIRST”

- Findable, Interoperable, Reusable, **Streamable**, **Transformable**

Harmonise/ Standardise

AI requires

- Analysis Ready Data => AI Ready Data
- Model Context Protocol Server (MCP)

Loosely coupled Data Spaces (federated vs monolithic)

Process near data and Transform & Stream
only what is needed !

Pangeo

DASK

xarray

Open Geospatial Consortium

STAC SpatioTemporal Asset Catalog

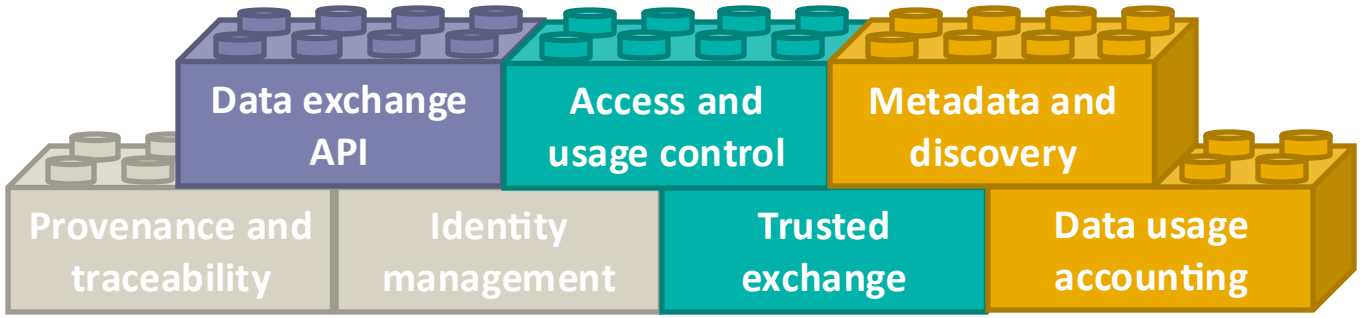
Zarr

python

Destination Earth

SIMPL is the smart middleware that will enable cloud-to-edge federations and **support all major data initiatives** funded by the European Commission, such as the **common European data spaces and DestinE**

SIMPL - Smart Middleware Platform
- Simpl-Open – middleware itself
- Simpl-Labs – Infrastructure & test beds
- Simpl-Live – Deployments of Simpl-Open




SAVE THE DATE!
Fifth Destination Earth
User eXchange
9-10 June 2026
Brussels, Belgium
Funded by
the European Union
Destination Earth
implemented by
ECMWF
esa
EUMETSAT
in partnership with
RMI