EUMETSAT involvement in DestinE – The DestinE Data Lake

Lothar.Wolf@eumetsat.int

Background & EUMETSAT role

Policy framework and purpose

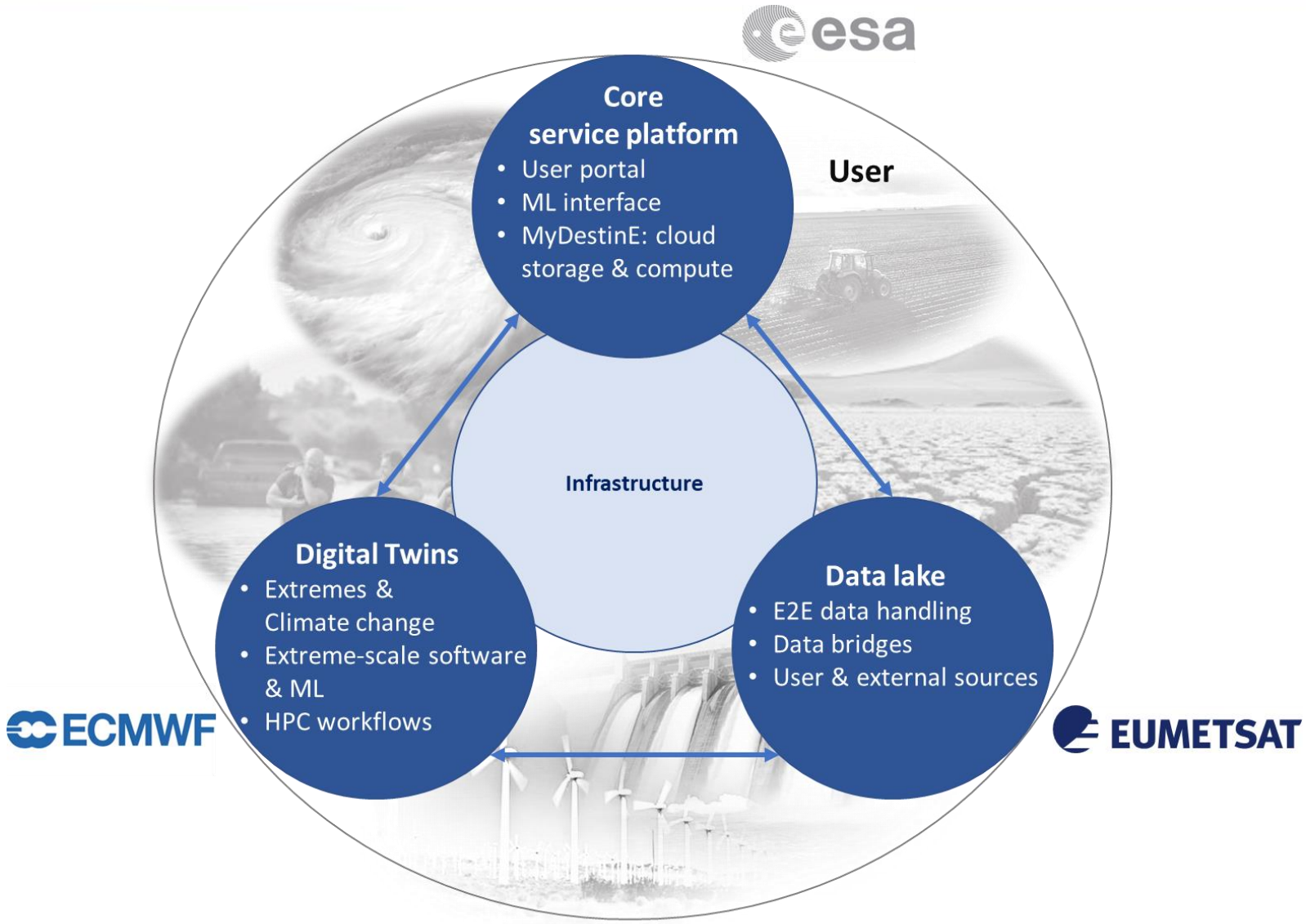
- Destination Earth is identified in communications of the European Commission on
  - Green Deal
  - Europe Digital's future
  - European Strategy for Data

- Purpose of DE according to EC scoping paper
  - "*Destination Earth (DE) aims at developing a very high precision digital model of the Earth (Digital Twin of the Earth) to enable end-users to assess not only the impact of environmental and other societal challenges but also the efficiency of the proposed solutions, incl. EU legislative measures*."

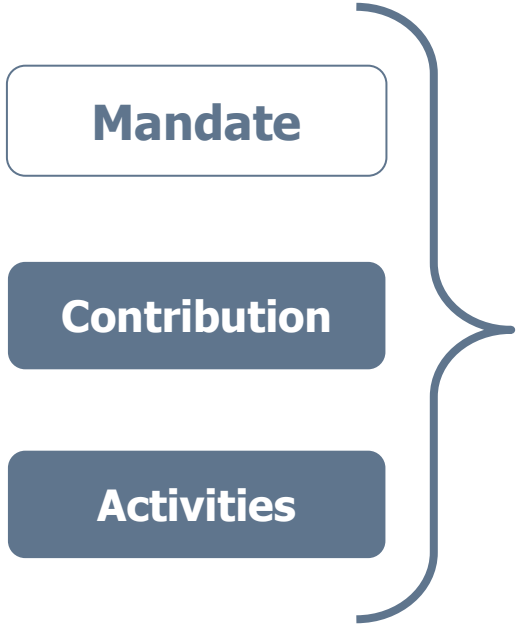
Destination Earth - core elements

Federated cloud-based modelling, simulation and predictive analytics platform	• (fast) access to data and high connectivity capacities • High Performance Computing (HPC), AI…
Data	• Copernicus/Earth Observation • environmental, users own data…
Digital Twins	• climate change adaptation, extreme weather • biodiversity, oceans…
Applications/services	• value-added services by/for private/public entities • available through the platform and digital twins

Three implementing agencies

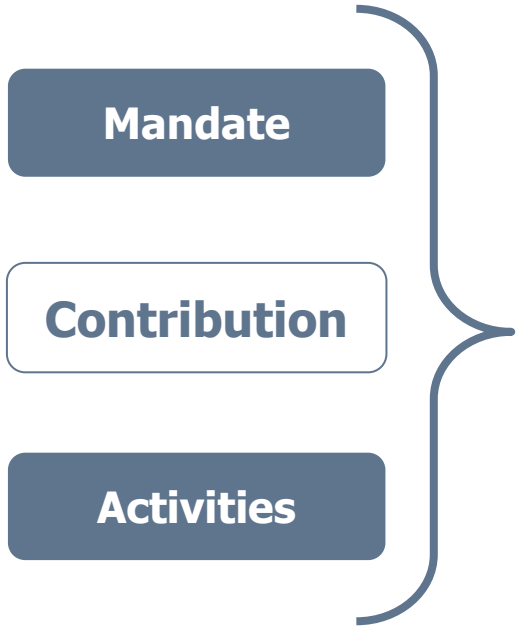


EUMETSAT Mandate



- Provision of **High Value Data Sets**;
- Design and exploitation of **elements** of the distributed and federative **operational infrastructure** of DE;
- Contribution to the **development** of “thematic” Digital **Twins** of the Earth in support to thematic lead entities;
- Contribution to the **exploitation** of “thematic” Digital **Twins** of the Earth in support to thematic lead entities;
- Contribution to DE-relevant collaborative **research** projects funded by the Horizon Europe programme in support to thematic lead entities.

EUMETSAT high level Contributions



- EUMETSAT **Big Data Services** in particular as Data Lake operator and with Data Tailoring Services including SAF contributions
- **European Weather Cloud(*)** jointly with ECMWF for data pre-processing steps in the value chain such as AI algorithms on the above mentioned Data Lake
- **Copernicus WEkEO(*)** with IT resources and services related to industrial/commercial applications

- EUMETSAT **harmonized data access service** (using standardized APIs and distributed data adaptors) which can be further scaled and extended to data interconnect with other data spaces and data lakes
- EUMETSAT **operational experiences** also with regards to establish **cloud & data federations** with suitable **governance concepts**

(*) Element already involves a partnership with either ECMWF, EC, ESA

EUMETSAT activities within Contribution Agreement

Mandate

Concept

Activities

- Design, Implementation & Operation of a self standing DestinE Data Lake & Warehouse
- Access to the operational EUMETSAT Data Lake with additional data sets and API interfaces
- Provision of the EUMETSAT Data Tailoring Services with new DE algorithms
- Extension of the EUMETSAT Harmonized Data Access Service supporting federation with other Data Lakes
- Provision of Compute Resources via the European Weather Cloud to host to AI and other pre-processors
- Provision of technical expertise in design, implementation and operation of state of the art cloud & data federations

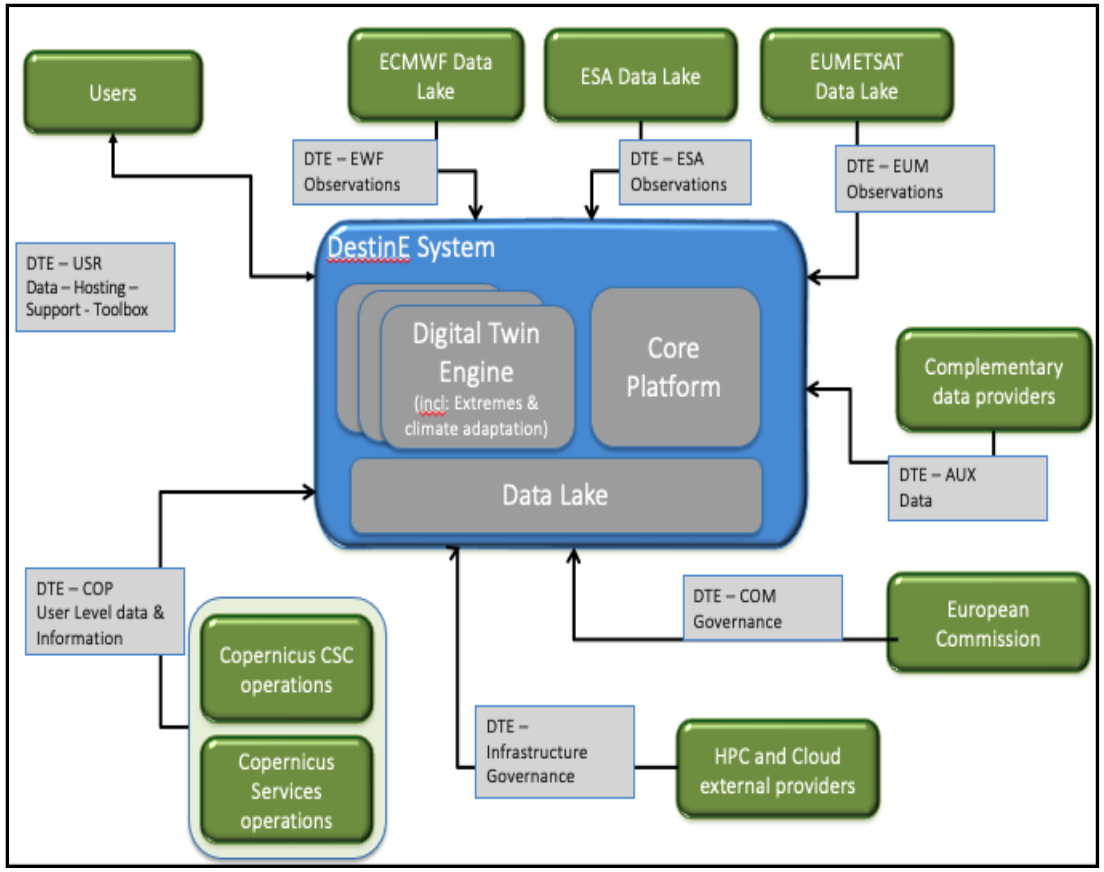
Example Digital Twins

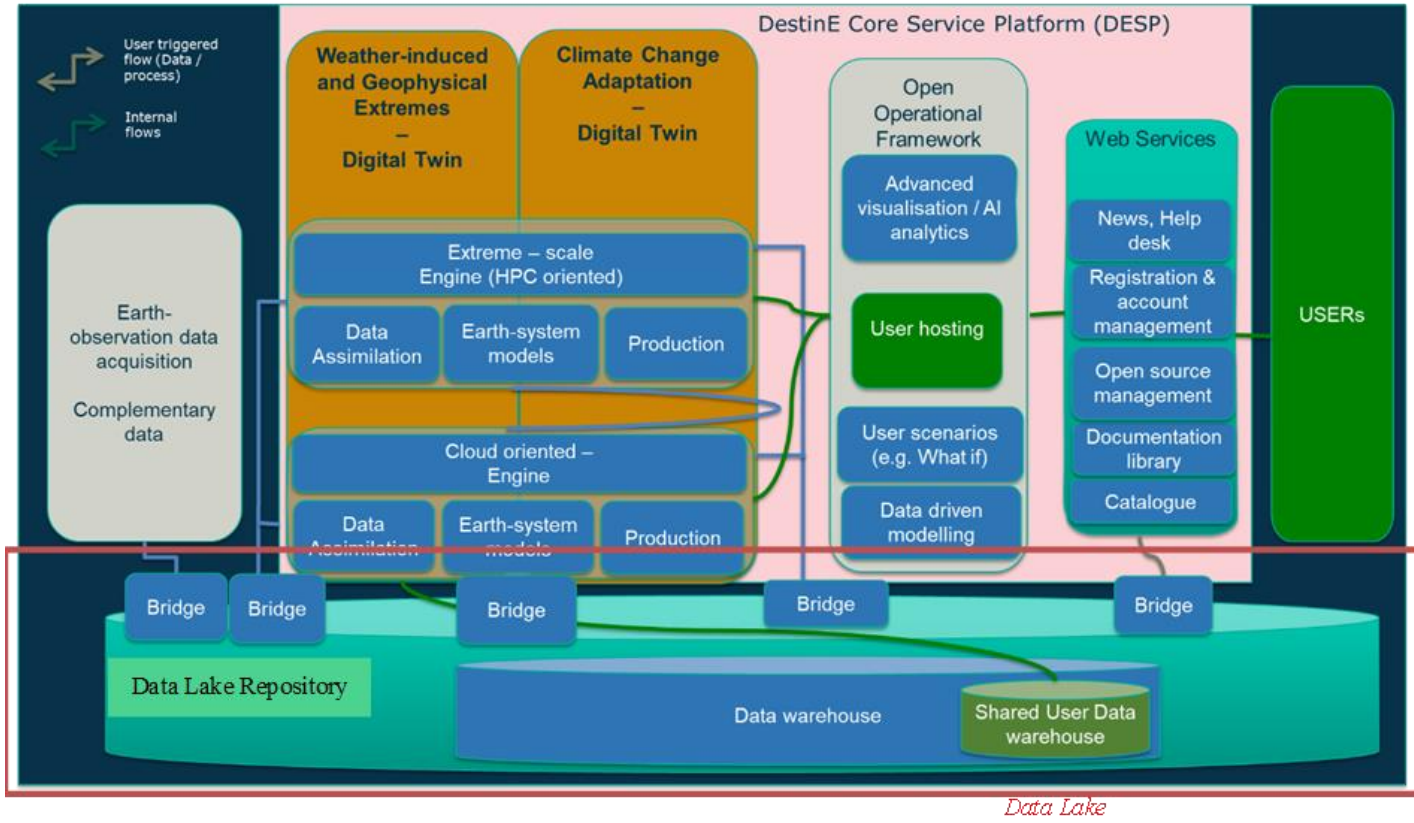
- Thematic lead
- Fast and flexible user interface
- Big data handling
- High performance and cloud computing

ECMWF | **EUMETSAT** | **ESA**

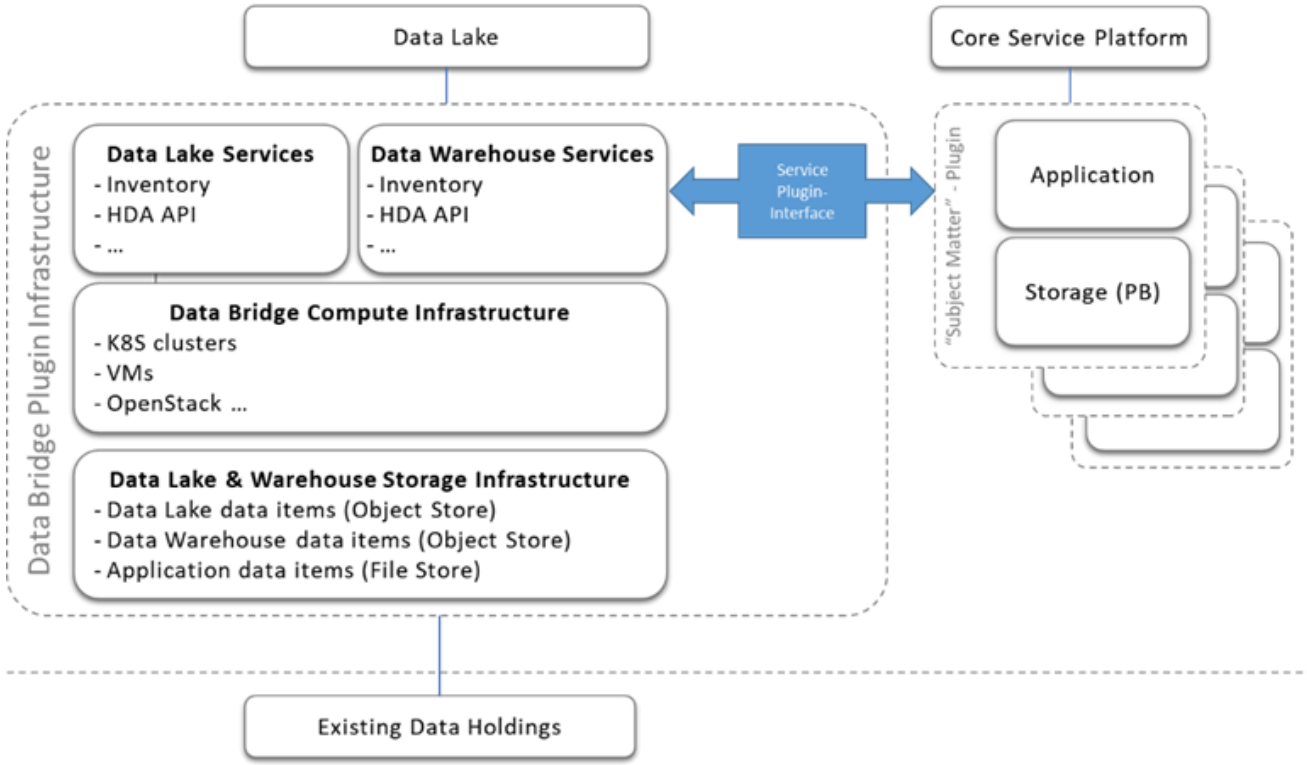
Technical Overview and Synergies

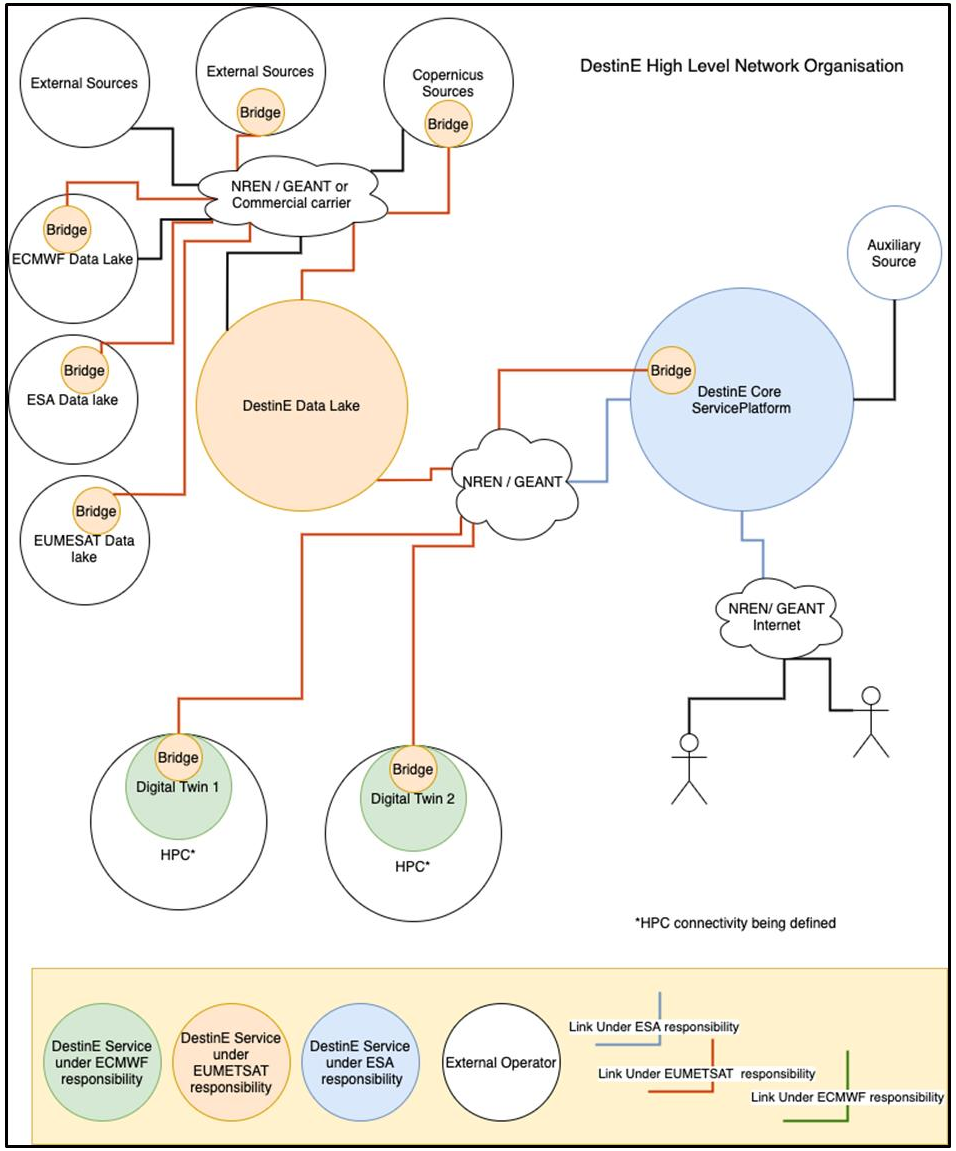
OVERVIEW OF DESTINATION EARTH ARCHITECTURE





High Level Data Lake & Data Warehouse concept





OVERVIEW OF EUMETSAT RESPONSIBILITIES IN DESTINATION EARTH

- Design and development of the Data Lake followed by its end-to-end deployment, testing and operation;

- Implementing, testing and operating the interfaces
  - to the partners´ data repositories, in particular integration of Copernicus interface and access to all related data;
  - to the DestinE Core Platform;
  - to the Digital Twin Engine for accessing the generated output and ensure its operational support to the Digital Twins;

- Providing the online inventory of all the data available through DestinE data Lake;

- Establish the corresponding industrial procurements to accomplish above points.

Relation and Synergies

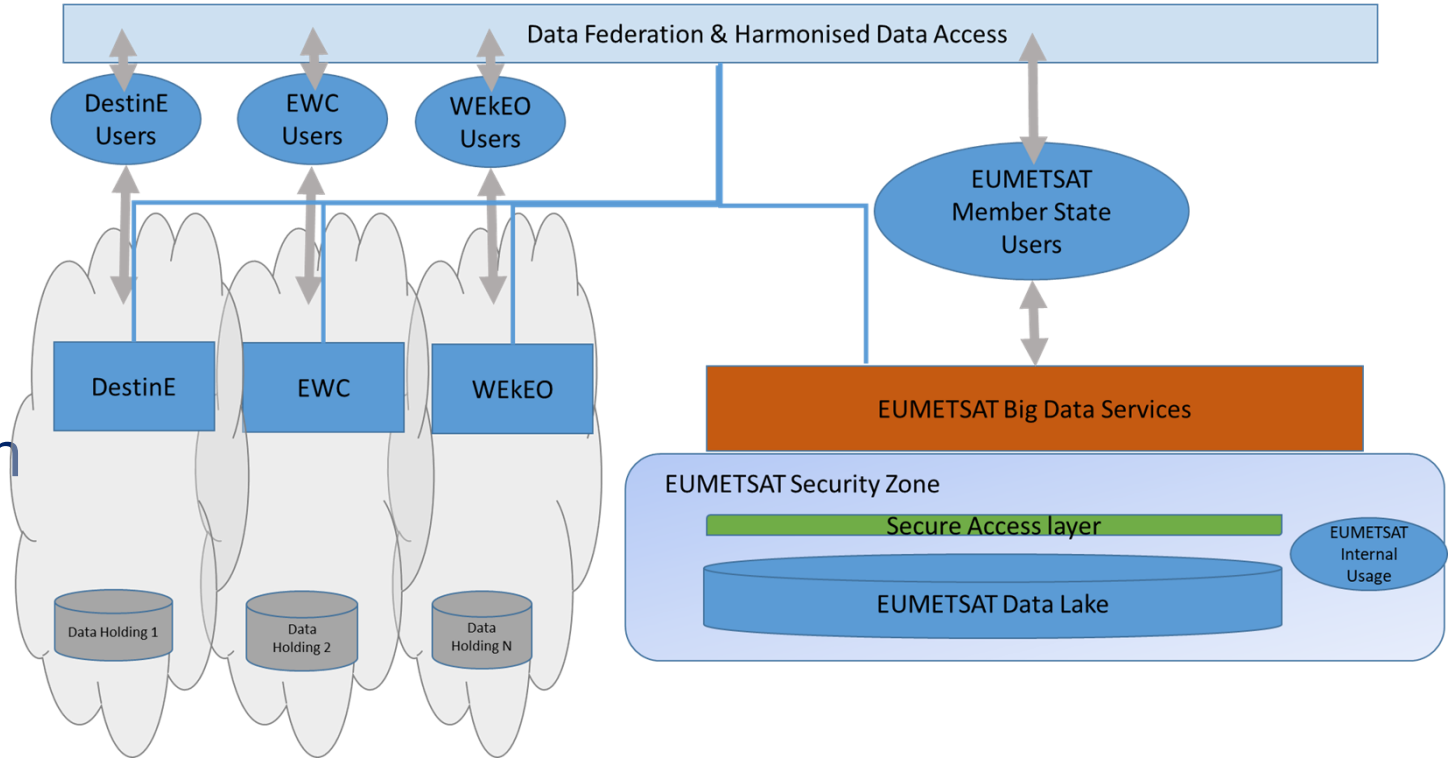
- DestinE is part of a coherent Big Data Strategy of EUMETSAT

EUM Big Data Services → WEkEO → EWC→ DestinE

- EUMETSAT promoted the “Federation” of cloud-based systems, based on Harmonised Data Access across geographically distributed locations:
  - ✓ Access to distributed data integrated with EUMETSAT´s **Big Data Services**
  - ✓ Demonstrated by **WEkEO** since June 2018 (V0) - fully operational since Q1/2020;
  - ✓ Proven suitability to address EC’s needs for Copernicus → will be sustained in COPER 2.0
  - ✓ Forms also the basis of the **European Weather Cloud**
- The role of EUMETSAT in DestinE:
  - Recognises EUMETSAT’s **leading experience** in this area
  - Ensure that EUMETSAT’s integrated solution is also central to DestinE.

SYNERGY AND COMPLEMENTARITY WITH OTHER EUMETSAT BIG DATA INITIATIVES

- The various external activities and environments such as WEkEO, EWC and DestinE can be considered as self-standing;
- The EUMETSAT Big Data Services form the common basis - They also ensure that activities and data within EUMETSAT remain protected from interference in terms of access and data policy;
- The Data Federation layer allows to realise the synergies expressed above.



Programmatic aspects

Major Tasks – aligned with the main responsibilities

- Scope of EUMETSAT's contribution: establish a multi-cloud data lake and provide operational support to Digital Twins of Earth and the DestinE Core Platform;
  - Design and development of the Data Lake followed by its end-to-end deployment, testing and operation;
  - Implementing, testing and operating the interfaces
    - to the partners' data repositories, in particular integration of Copernicus interface and access to all related data;
    - to the DestinE Core Platform;
    - to the Digital Twin Engine for accessing the generated output;
  - Provide the online inventory of all the data available through DestinE Data Lake
- Perform the industrial procurements

Phases

- Phase I: July 2021-December 2023
  - Preparation, testing, piloting and implementation of a first set of Destination Earth services and of the first Digital Twins
- Phase II: January 2024- December 2025
  - Further development of Destination Earth and potential integration of further digital twins
- Phase III: January 2026- December 2027 Full operationalization

Overview of the Legal Framework

- EUMETSAT activities in support of DestinE for the period 2021-2027 need to be implemented by way of a Third Party Programme in accordance with the EUMETSAT Convention

- Agreements:
  - EU will conclude three separate Contribution Agreements with EUMETSAT, ESA and ECMWF;
  - EUMETSAT, ESA and ECMWF will conclude a tripartite cooperation arrangement addressing technical and operational coordination and implementation aspects.

Q & A

Questions ☺?