

# Destination Earth Data Lake Big Data Processing Services - Edge Services

Patryk Grzybowski<sup>1</sup>, Marcin Ziółkowski<sup>1</sup>, Aubin Lambare<sup>2</sup>, Christoph Reimer<sup>3</sup>, Michael Schick<sup>4</sup>

<sup>1</sup>CloudFerro S.A., Warsaw, Poland; <sup>2</sup>CS Group, Le Plessis Robinson, France; <sup>3</sup>EODC, Vienna, Austria; <sup>4</sup>EUMETSAT, Darmstadt, Germany



## Destination Earth (DestinE)

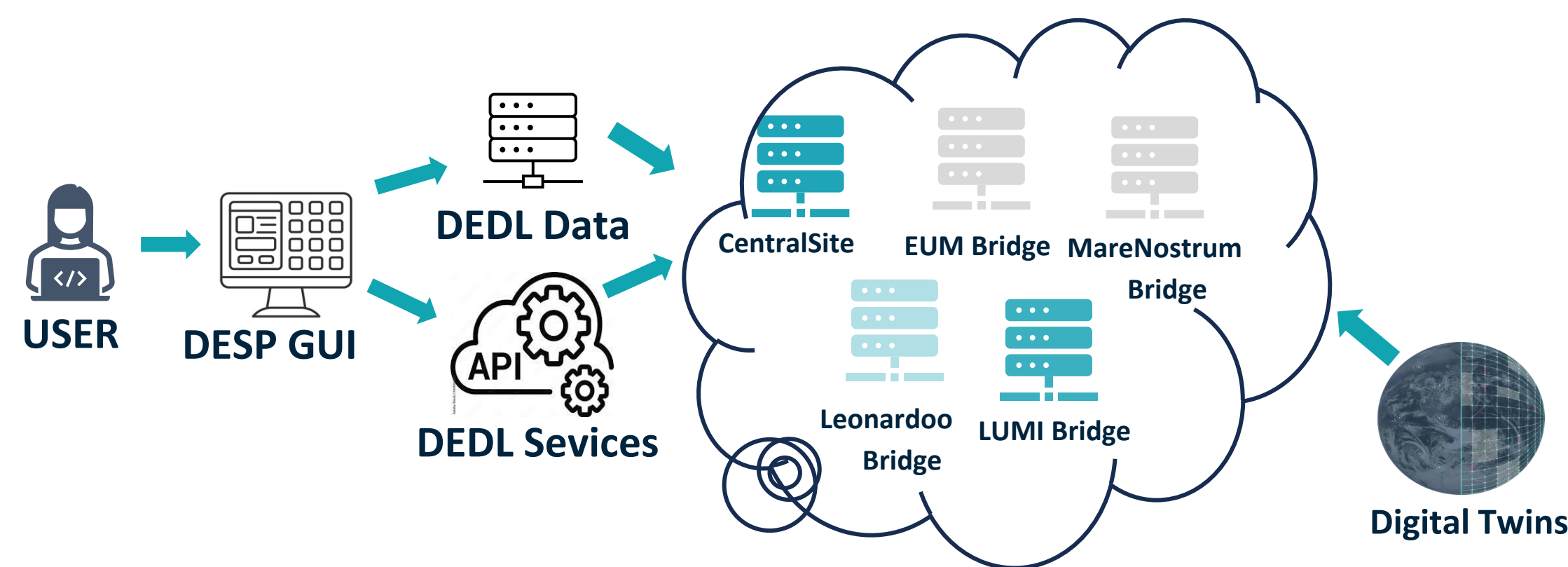
### SUMMARY

**Destination Earth (DestinE)** is a flagship initiative of the European Commission to develop a highly-accurate digital replica of the Earth to model, monitor and simulate natural events, hazards and human activities.

Three components are included within DestinE:

- Core Service Platform (DESP)
- Data Lake (DEDL)
- Digital Twins (DTs)

### Destination Earth Data Lake (DEDL) Design



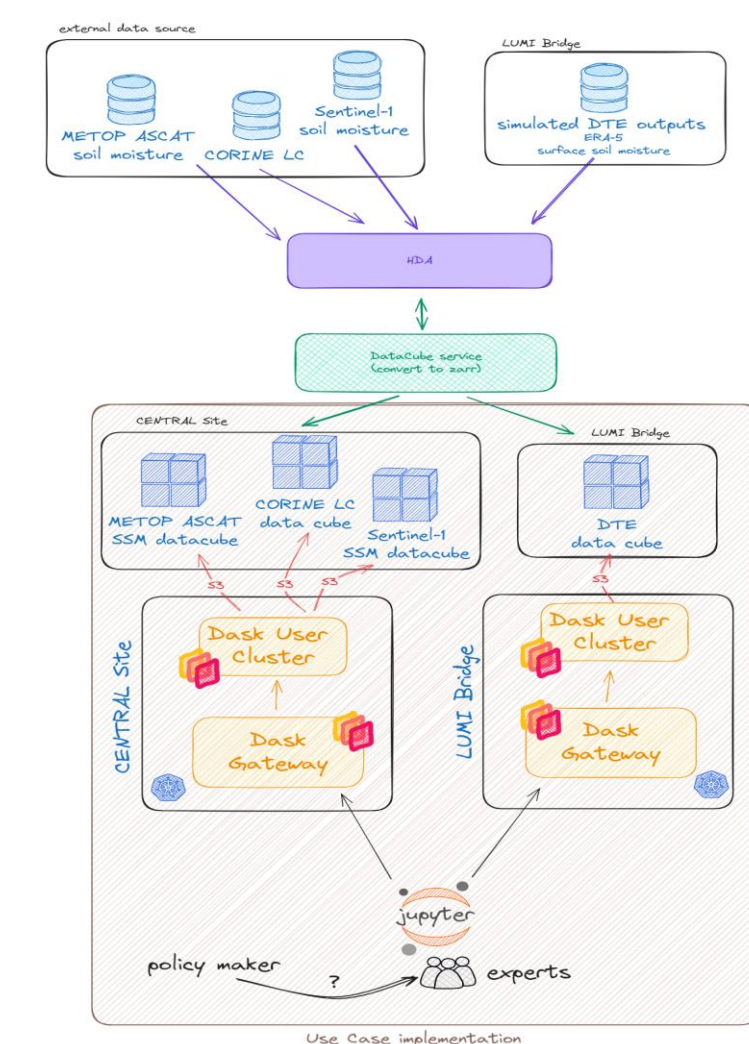
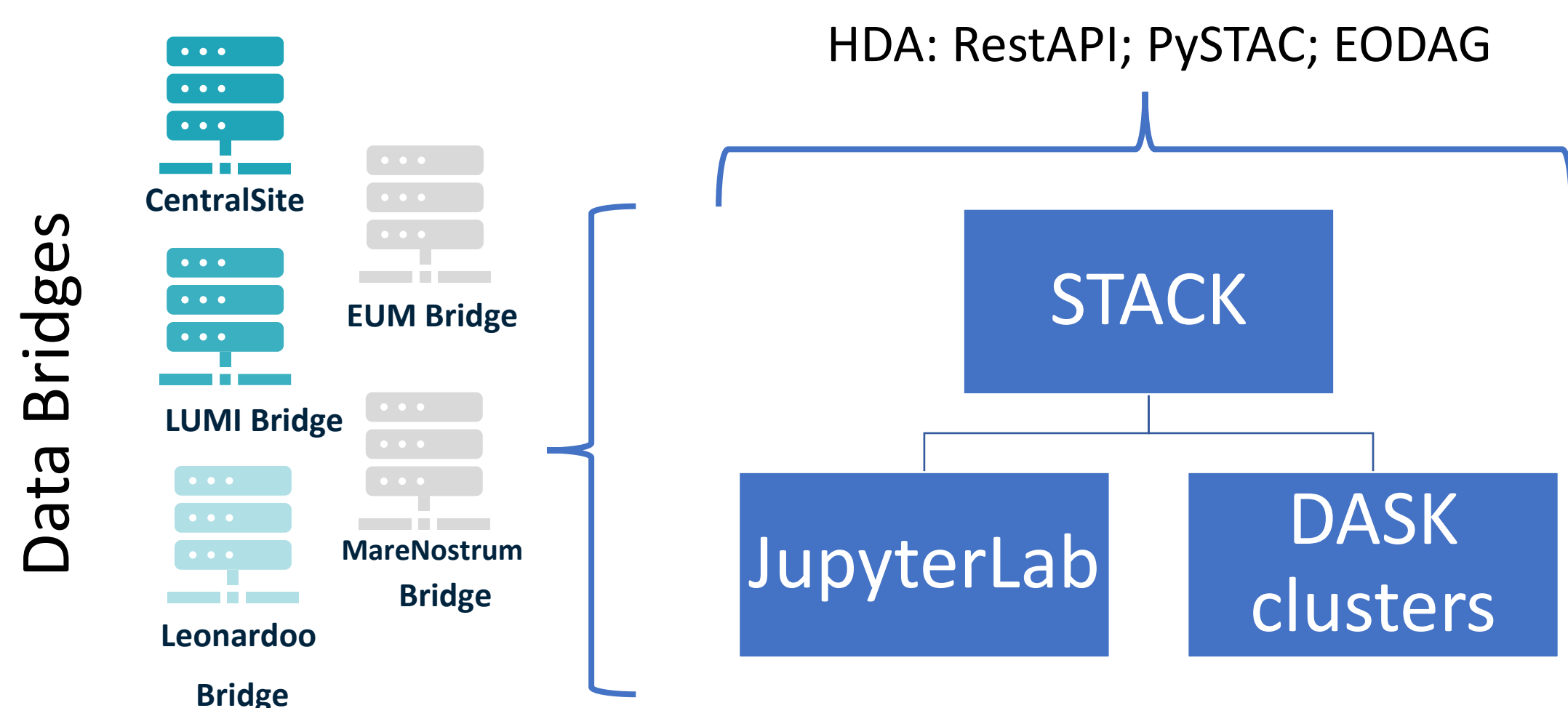
### Services provided by DEDL

#### DEDL User Services

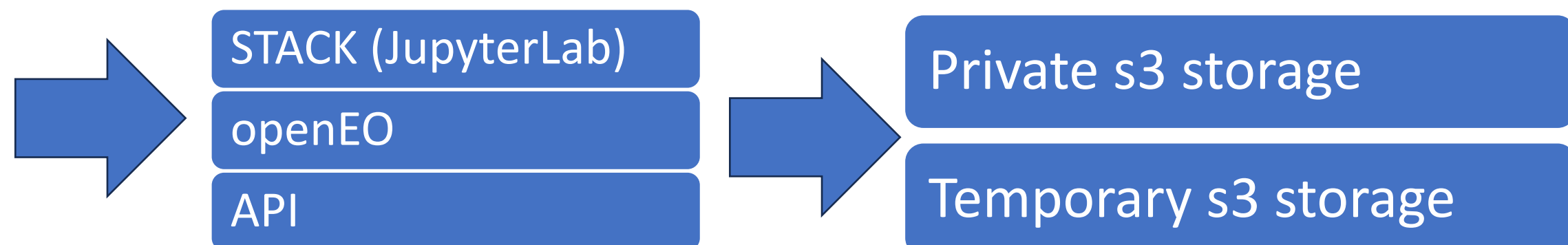
Discovery Services	Data Access Services	Edge Services	User Service Desk
Discover Data	Federated Datasets	Cloud Infrastructure (Islet)	Help Desk
Discover Services	Fresh Data Pool	Application (Stack)	
	DT Outputs	Processors (Hook)	
	User Generated Data		

- HDA API for data discovery and data access:
  - Around 170 datasets
  - STAC protocol
  - One pair of credentials
  - One API
- Storage services including Fresh Data Pool for intermediate data, DT outputs and user's private data
- **Edge services:**
  - **Cloud computing infrastructure (Islet Compute) and storage (Islet Storage)**
  - **JupyterHub (Stac)**
  - **processors (Hook).**
- Help Desk

### DEDL Edge Service



- Pre-defined hooks**
- Data harvest
  - LAI
  - CARD BS
  - CARD COHINF
  - SNAP C2RCC
  - SEN2COR
  - MAJA
- Users' custom hooks**



- Virtual environment based on the OpenStack
- Preconfigured images
- VPU and GPU
- K8s clusters



- Access to s3 storage
- Manage s3 storage