The ENES Data Space

Fabrizio Antonio
Research Engineer at CMCC Foundation

Dissemination level: Public

Disclosing Party: Project Consortium

Recipient Party: European Commission

EGI-ACE Project Final Review

12 September 2023

EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.

ENES Data Space: Motivation & Goal

Motivation: Tackle main **key challenges and practical issues** related to **large-scale climate analysis**

- Input data from **multiple models**
- **Data download** is a big barrier for climate scientists
- **Client-side** & **sequential approaches**
- Several **data analysis tools** and **libraries** needed
- **Strong requirements** in terms of **computational** and **storage resources**

Goal: Deliver an open, scalable, and cloud-enabled **data science environment** for **climate data analysis** on top of the **EOSC Compute Platform**

The **ENES *Data Space*** aims at providing an entry point to:

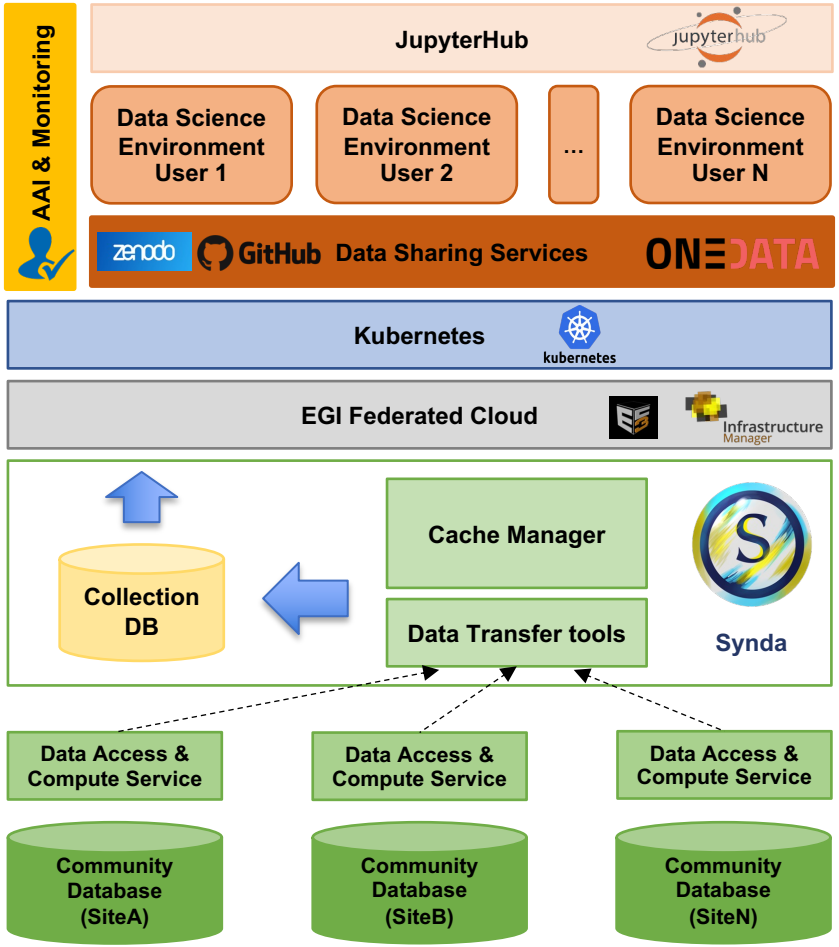
- **Datasets** → most relevant from **ESGF** (e.g., CMIP); pre-staged; open
- **Storage** & **Compute** resources → provided by **EGI**
- **Data Science Software Stack** → to address a wide spectrum of analysis needs (mainly Python-based)
- **Jupyter-based gateway** → to devel/share/(re-)use apps → **FAIR** principles

→ **Ultimate goal**: promote **Open Science** and a more sustainable, effective, and **FAIR** use of **Data** and **Services**

ENES Data Space

ENES Data Space architecture

WP6



EOSC compute services & interfaces

EOSC data sharing services & interfaces

WP6

Infrastructure as a Service (IaaS) Cloud

WP3 & WP6

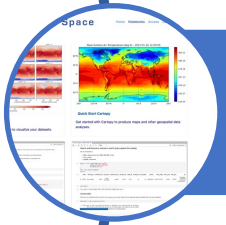
Data collector and Cache Service

Community (legacy) Infrastructure
Community-specific
data access services & interfaces

Climate research workflow

ENES Data Space

Data Search & Discovery

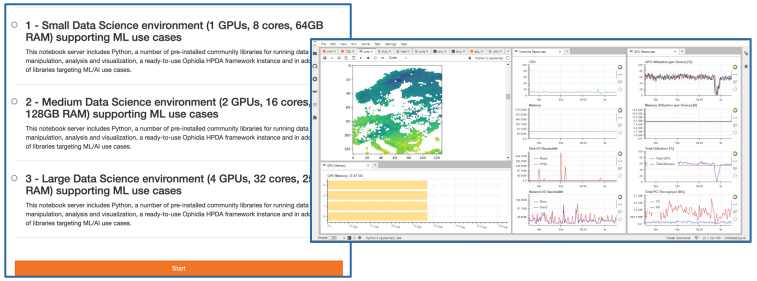


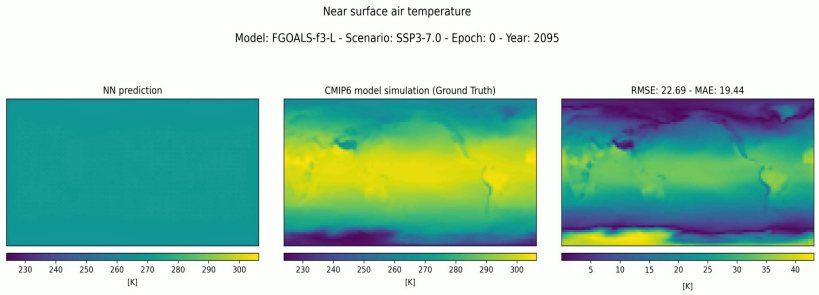
Data Analysis & Visualization

Reusability & Sharing

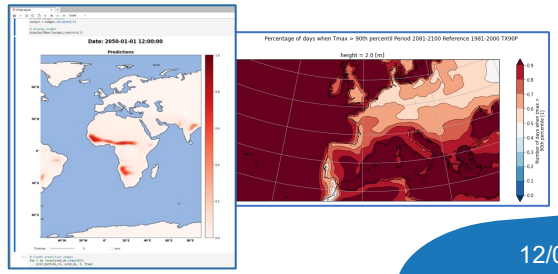
Support to AI-based applications

- Data Space infrastructure recently enhanced to include **accelerators** (e.g., **GPUs**)
- Several profiles including the most common **Python libraries for Machine Learning**
- **GPU usage dashboard** for real-time visualization of NVIDIA GPU metrics
- Several use cases supported
  - Climate use case from the **EGI-ACE Call for Use Cases**
  - Research activities carried out by the **CMCC Machine Learning Research Group**
  - Support to **Digital Twin applications**
    - Analysis of extreme events on future climate projections
    - Large amounts of **CMIP6 data** for models training and validation
    - Proof of concept of a first DT − the **Wildfires DT** − developed by CMCC exploiting the ENES Data Space **features**





interTwin



12/09/2023

DEMO

DOI 10.5281/zenodo.8325047

Useful links

EGI-ACE: https://www.egi.eu/projects/egi-ace/

ENES Data Space: https://enesdataspace.vm.fedcloud.eu/

ENES portal: https://portal.enes.org/

EGI-ACE *receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101017567.*

IS-ENES3 *has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 824084*

Thank you!

Contact: **egi-ace-po@mailman.egi.eu**
Website: **www.egi.eu/projects/egi-ace**

EGI Foundation

@EGI_eInfra

EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.