

EGU22-7330

<https://doi.org/10.5194/egusphere-egu22-7330>

EGU General Assembly 2022

© Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.



ENES Data Space: an open, cloud-enabled data science environment for climate analysis

Fabrizio Antonio¹, Donatello Elia¹, Andrea Giannotta¹, Alessandra Nuzzo¹, Guillaume Levavasseur², Atef Ben Nasser², Paola Nassisi¹, Alessandro D'Anca¹, Sandro Fiore³, Sylvie Joussaume², and Giovanni Aloisio^{1,4}

¹Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Lecce, Italy

²Institut Pierre Simon Laplace (IPSL), Centre National de Recherche Scientifique (CNRS), France

³University of Trento, Trento, Italy

⁴University of Salento, Lecce, Italy

The scientific discovery process has been deeply influenced by the data deluge started at the beginning of this century. This has caused a profound transformation in several scientific domains which are now moving towards much more collaborative processes.

In the climate sciences domain, the ENES Data Space aims to provide an open, scalable, cloud-enabled data science environment for climate data analysis. It represents a collaborative research environment, deployed on top of the EGI federated cloud infrastructure, specifically designed to address the needs of the ENES community. The service, developed in the context of the EGI-ACE project, provides ready-to-use compute resources and datasets, as well as a rich ecosystem of open source Python modules and community-based tools (e.g., CDO, Ophidia, Xarray, Cartopy, etc.), all made available through the user-friendly Jupyter interface.

In particular, the ENES Data Space provides access to a multi-terabyte set of specific variable-centric collections from large community experiments to support researchers in climate model data analysis experiments. The data pool of the ENES Data Space consists of a mirrored subset of CMIP datasets from the ESGF federated data archive collected by using the Synda community tool in order to provide the most up to date datasets into a single location. Results and output products as well as experiment definitions (in the form of Jupyter Notebooks) can be easily shared among users through data sharing services, which are also being integrated in the infrastructure, such as EGI DataHub.

The service was opened in the second part of 2021 and is now accessible in the European Open Science Cloud (EOSC) through the EOSC Portal Marketplace (<https://marketplace.eosc-portal.eu/services/enes-data-space>). This contribution will present an overview of the ENES Data Space service and its main features.