



Contribution ID: 29

Type: **Presentation**

## Open Data analysis with EOSC-hub services

*Tuesday, January 29, 2019 2:00 PM (10 minutes)*

The EOSC-hub project mobilises providers from the EGI Federation, EUDAT CDI, INDIGO-DataCloud and major research e-infrastructures offering services, software and data for advanced data-driven research and innovation. These resources are offered via the Hub – the integration and management system of the European Open Science Cloud, acting as a single entry point for all stakeholders.

Several of the use cases in EOSC-hub will enable scientific end-users to perform data analysis experiments on large volumes of data, by exploiting a PID-enabled, server-side, and parallel approach. Users expect also easy to use interfaces like Jupyter Notebooks for interacting with the system.

This talk presents an ongoing effort to turn these needs into a production-ready service that provides a FAIR-approach to the researchers analysis workflow by leveraging and integrating the following services from the EOSC-hub catalogue:

- Check-in for handling user authentication and authorisation to services with federated AAI
- DataHub for providing a data access and management
- B2HANDLE for PID minting for research objects available in DataHub
- Notebooks as an interactive interface to carry out the end-user tasks, including data analysis of a public dataset referenced by a PID and publication of the results back with a new PID.
- B2FIND for making data discoverable by harvesting data using OAI-PMH

Thanks to the use of federated AAI and interoperable protocols, this work will be further extended to integrate new services from EOSC-hub to support computing intensive workloads, other data-related services, and potentially any other service with community-scoped functionality as required by the different use cases.

**Primary authors:** Mr GRENIER, Baptiste (EGI Foundation); FERNÁNDEZ, Enol (EGI Foundation)

**Presenters:** Mr GRENIER, Baptiste (EGI Foundation); FERNÁNDEZ, Enol (EGI Foundation)

**Session Classification:** Open Data Ecosystem and CS3

**Track Classification:** Open Data Ecosystems and CS3