



Charting the course towards a concrete European Open Science Cloud: outcomes of the EUDAT Conference “Putting the EOSC vision into practice”, Porto, 22-25 January 2018

What does the EOSC vision mean in practice? What are the key factors for its success? What is the role of research communities? EDI & EOSC: What are the common goals and how should they work together?

These are few of the questions that the over 230 participants of the **EUDAT’s Conference “Putting the EOSC vision into practice”** strived to answer at the 3-day meeting held in Porto. Attendees included policy makers, service providers and research communities representatives from 25 countries working on various data challenges and disciplines.



The conference was opened by **Augusto Burgueño Arjona, Head of Unit "eInfrastructure", DG CONNECT** who presented EOSC as an important instrument to support the collaboration between e-Infrastructures & research infrastructures, and to promote open science: ***"EOSC has to be an inclusive ecosystem where horizontal and thematic service providers work together to meet the user needs"***.

The discussion on how to transform the EOSC vision into practice continued with a panel moderated by Annabel Grant, Senior Stakeholder Engagement Manager, GÉANT. Françoise Genova, Researcher at Centre de Données astronomiques de Strasbourg (CDS), Per Öster, Director, CSC & EOSC-hub Project Director, Grazia Pavoncello, ministerial representative at the Italian Ministry of Education, University and Research (MIUR) and Alex Vermeulen, Carbon Portal Director of ICOS ERIC, participated to the panel, together with Burgueño Arjona.

**Panel recommendations:**

- **EOSC should be mostly about interoperability and integration** - Previous investments (existing services) and experiences (competences and use cases) must be leveraged;
  - **EOSC should not be seen only as a technical problem** - EOSC is not expected to be a new platform but a federation of existing services provided by different organisations with different governance and funding mechanisms. EOSC needs to define clear terms of use considering access rights, conditions to access services and business models;
  - **EOSC has to be an "affordable", "inclusive" ecosystem** - EOSC aims at taking on board different countries, disciplines and institutions to offer all of them the same conditions to make data driven science. The access conditions need to be affordable from a cost perspective to encourage participation;
  - **Users must be in the driving seat** - EOSC should allow a smooth onboard of mature communities and should be an essential enabler for new communities;
  - **Research Infrastructures (RIs) have a double role in EOSC: as users and thematic service providers** - It is clear that EOSC is the opportunity for RIs to expand their user base. Multi-disciplinary use cases would be a good instrument to assess the value and the benefits that EOSC can bring to RIs and to attract them;
  - **Measuring usage of the EOSC services is fundamental to assess impact** - A Virtual Access mechanism is already in place. Measuring usage is not only fundamental in terms of impact but also to compensate the service providers. We need to fund services that are used;
  - **Data quality, trust, and user feedback are still open challenges.**
- "From an artistic point of view EOSC is a movie, which is art, but behind there is an industry that has to organize the infrastructure, the governance and the sustainability model."*** - Augusto Burgueño Arjona, European Commission



The conference continued with an inspiring keynote by Michael Wise, Head of Astronomy, ASTRON - the Netherlands Institute for Radio Astronomy, who presented the data challenges behind the Square Kilometer Array (SKA) Project. *“Based on current projections, the SKA Observatory, once operational, is expected to produce an archive of standard data products with a growth rate on the order of 300 petabytes per year. Any further processing and subsequent science extraction by users will require significant, additional computing and storage resources”*.

This presentation set the scene for a second panel chaired by Rob Baxter, EPCC, University of Edinburgh on “The European Data Infrastructure (EDI) and the Data Challenge” which focused on understanding the role of HPC in the EOSC and EDI landscape. Serge Bogaerts, PRACE, Giuseppe Fiameni, CINECA, Kimmo Koski, CSC, Sinead Ryan, Trinity College Dublin, and Michael Wise, ASTRON, contributed to the panel discussion.

**Panel recommendations:**

- **EDI should be the traditional vehicle to provide HPC resources to EOSC**
- **Different funding streams, common objectives** - Even if there are different funding streams (EuroHPC, EDI, EOSC), user communities and ICT providers need to work together to address societal challenges (climate changes, natural disasters, etc.)
- **EDI and EOSC should be supported by strong scientific cases, which are fundamental to demonstrate real impact. Ideally, these cases should be ideally multidisciplinary ones.**
- **Scientists managing exascale data need to focus on the science problem** - The data and HPC resources are just the instruments. ICT providers need to put scientists in the condition of doing science.



*Panel Plenary Session 2: The European Data Infrastructure (EDI) and the Data Challenge*

The panel recommendations were further discussed in a set of breakout sessions which explored overarching themes of crucial importance in the creation of a thriving data economy, such as legal issues, interoperability of services, the role of research infrastructures as thematic service providers and business models and sustainability of data infrastructures.

Nine co-located events organised by ENVRI, EOSCpilot, the EUDAT Working Groups on Sensitive Data, Semantic and Array Databases, GEANT, LIBER and SeaDataCloud complemented the EUDAT conference allowing participants to deepen specific topics and to establish new collaborations.

Finally, the conference was the opportunity to engage stakeholders to discuss the future of EUDAT and the **EUDAT Collaborative Data Infrastructure (CDI)**.

Counting already more than 20 members leading European research organisations, data and computing centres, the EUDAT CDI will continue developing and operating an interoperable layer of common data services to support research in Europe. The participation of the EUDAT CDI in the **EOSC-hub project** will also allow EUDAT to play a

concrete role in the EOSC ecosystem and will guarantee a continuous interaction with the user communities that were at the heart of the EUDAT strategy since its start in 2013.



*The EUDAT community at work*

### **About the EUDAT CDI**

With a network of more than 20 European research organisations, data and computing centres in 14 countries, the EUDAT Collaborative Data Infrastructure (CDI) is one of the largest infrastructures of integrated data services and resources supporting research in Europe. It is designed to address the full lifecycle of research data, representing a strategic solution to the challenge of data proliferation in Europe's scientific and research communities and it is realised through an ongoing collaboration between Service Providers and Research Communities working as part of a common framework for developing and operating an interoperable layer of common data services. The CDI services cover data access, data storage, data discovery and metadata, persistent identification, data management, authentication & authorisation, service management infrastructure and Research Data Management (RDM) training & consultancy. More information at [www.eudat.eu](http://www.eudat.eu)