



Keywords:

#experiments #data #science, #service, #mesh, #infrastructure #EOSCinPractice #cross-disciplinary

Bringing big science experiment data to the researchers' fingertips.

An EOSC in Practice Story where ESCAPE's Data Infrastructure for Open Science and CS3MESH4EOSC' Science Mesh connect.

The project involved

ESCAPE is one of the five thematically clustered European Strategy Forum on Research Infrastructures (ESRFI) projects supported under the European Union H2020 research and innovation programme (Grant Agreement 824064). It aims to establish a single collaborative cluster of next generation ESFRI facilities in the area of astronomy and accelerator-based particle physics in order to implement a functional connection with EOSC. This goal is driven by the observation that sciences are facing unprecedented volumes of data and files to manage. To facilitate researchers' work, ESCAPE enables **technical interoperability between the facilities**, that is, the "ability of different information technology systems and software applications to communicate and exchange data"¹. This minimises fragmentation, encouraging cross-fertilisation and developing joint capabilities in **astronomy, astrophysics and particle astrophysics communities**.

CS3MESH4EOSC offers Interactive and agile/responsive sharing mesh of storage, data and applications for EOSC. The project receives funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 863353. CS3MESH4EOSC addresses the challenges of the fragmentation of file and application services, digital sovereignty and the application of FAIR principles in the everyday practice of researchers. Initially, 7 major data services will be combined into **ScienceMesh** - a **federated service mesh** providing a frictionless collaboration platform for hundreds of thousands of users (researchers, engineers, students and staff). The service mesh offers easy access to data across institutional and geographical boundaries. The infrastructure will be gradually expanded and offered to the entire education and research community in Europe and beyond.



"Fundamental science is at the base of our world's functioning. You need fundamental science to boost innovation, that's why we want to bring big science data at the fingertips of everyday researchers"

Xavier Espinal, Data Infrastructure for Open Science work package leader, ESCAPE



The Users

This EOSC in Practice story targets both researchers involved in large science projects (via ESCAPE) and citizen scientists or users interested in accessing (a part of) big science experiment data for everyday research purposes (via CS3MESH4EOSC).

The Challenge

One of the ESCAPE services is the **ESCAPE DIOS** (Data Infrastructure for Open Science), a scalable **federated data infrastructure** allowing an open access data service for the ESFRI projects within ESCAPE and concerned with Exabyte-scale data volumes. The **ESCAPE DIOS** is a **flexible and robust Data Lake** in terms of storage, security, safety and transfer, as well as a basic orchestration machinery, which enables the combination of technology with high quality data from different communities and, therefore, the exploration of new areas in science. While ESCAPE DIOS supports complex science research projects, it remains hardly manageable for other types of research needs. The challenge this EOSC in practice story wants to overcome is to bridge these two worlds by making complex science data available easily also to a long tale of lower scale scientific experiments with much less demanding computing needs, so to empower every possible user and enable advancement in many areas, including research, teaching, citizen science.

This is why **CS3MESH4EOSC** and **ESCAPE** are working together to enable the ScienceMesh activities into the ESCAPE DIOS. This will bridge diverse scientific data (regardless of its size) with researchers, outreach activities and open science initiatives.

¹ Source: EOSC interoperability framework <https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1/language-en/format-PDF/source-190308283>



The envisaged solution

In its research efforts, ESCAPE is slowly being adopted beyond particle physics and transitioning towards the individual researchers or citizen scientists, the kind of audience targeted by CS3MESH4EOSC. The two projects are trying to extend their boundaries and capabilities in order to eventually meet halfway. The key point is the common technology. Both projects are offering their own **Data Lakes or sync & share services** via notebooks and analysis platforms. Exploiting this common element, the aim is to enable everyone to access complex data science in a way that hides all the complexity for the benefit of the user. Enabling common login, users can access both their daily workspace and advanced science experiment data in a common blended platform. ESCAPE is currently working on a pilot case to implement the final service. The Low-Frequency Array, also known as LOFAR, can be a possible catalyser, as it is both active in ESCAPE and is bringing applied use-cases in ScienceMesh. It is a large radio telescope network where its innovation lies on combining data signals from separated antennas, by digitising them and then transported to a central digital processor, which will combine it in a software that emulates a conventional antenna. LOFAR aims to make much easier to Astronomy researchers to share and process this data.

Why do I need EOSC?

Firstly, establishing a free Data Lake and Data Management service in EOSC could help support the long-term existence of such services coming from the collaboration between ESCAPE and CS3MESH4EOSC, which would be potentially otherwise gone at the end of the projects. Moreover, Joining EOSC allows the following advantages:

- » Become an integral solution easily deployable for European projects, experiments and collaborations
- » Single point of access to big data from science experiments
- » Easy interaction with data, all complexity is hidden
- » Fast run time

The impact on society

This EOSC in practice story ensures increased inclusiveness in data access and management. Moreover, fundamental science is at the base of our world's functioning and has (although indirectly) an impact on virtually anything we deal with in daily life, from touchscreens to smartphones, internet, and nanotechnology with its related industrial applications. That's why lots of innovation can come from **sharing fundamental science experiment data**.

Contribution to cross-disciplinarity

Cross-disciplinarity is addressed on a double level in this case. First, in the aspect of pure science, via enabling data access to a multitude of different users. Second, from the technology point of view. Currently, projects and initiatives involve different techniques, systems and technologies, e.g. HPC, commercial clouds, private cloud from the universities, etc. Harmonising the way we access these technologies and the way they communicate with each other could and should boost European Research and Innovation.

Possible risks & limitations

Coordination efforts are needed to ensure the successful launch, run and maintenance of the service. The community needs to be engaged and attracted to use such service so to guarantee its future sustainability.

Sustainability for an EOSC in practice

To overcome the mentioned risks, early engagement of vast communities of final users is fundamental. Such communities could be initially unaware of their needs for data. But it is no secret that we all are indeed and will increasingly be in need for data. Let's just think about the fact that nowadays we take more pictures with our phones and need more Gigabytes in our drives. Our pictures are also larger and occupy more space than they did only a couple of years ago, because cameras are more precise and sensitive and produce better quality images. The same thing is happening with scientific instruments to measure nature: detectors, telescopes and antennas. It is not hard to imagine the computing and data processing needs many of us will face in the future. Connecting the services of ESCAPE and CS3MESH4EOSC is just a first brick that could potentially lead to many more synergies and unlock the potential of data sharing in our society.

The pilot test under development is expected to be completed in June. Regarding ESCAPE, the final aim is to have ESCAPE DIOS and all the other services produced by ESCAPE as a resource on **EOSC Exchange**. At the moment, ESCAPE services and tools are already being widely adopted by different science projects. This demonstrates the shared Data Lake infrastructure is a promising solution. And once ESCAPE's outcomes will be available via the EOSC Portal, engaging community usage will be key to the future development and further progress of the service. CS3MESH4EOSC will start working on the ScienceMesh integration into EOSC by June 2022, to be finalised by January 2023.

Future funding model scenarios

ESFRI and Research Infrastructures that will be adopting Data Lake tools will be maintaining and bearing related costs themselves to support their experiments, as the scale they operate is large.

Regarding the outputs ESCAPE is producing with CS3MESH4EOSC, the scenario is still uncertain, as the collaboration spontaneously started and was not planned at the moment of project funding. A basic infrastructure produced by ESCAPE will be continued in the framework of EOSC-Future. The rest of the services produced by ESCAPE and CS3MESH4EOSC could be transferred to EOSC when the project ends. The financial details and the possible economic support coming from EOSC itself or other interested stakeholders are yet to be decided. The collaboration with CS3MESH4EOSC, which has planned to produce a deliverable on sustainability in the final months of the project ending in December 2022, could strengthen both projects' position in the dialogue with EOSC and help drafting a comprehensive business plan.

Useful material related to this story

- » [Science Mesh Workshop "Global Platform for Scientific Collaboration"](#)
- » [Scientific disciplines embracing no border Research Environment: ESCAPE & Science Mesh](#)

Want to learn more about DIOS and the other ESCAPE services? [Access here](#)

Want to become a ScienceMesh adopter yourself? [Access here](#)

Liked this **#EOSCinPractice** story? Follow [@EOSCFuture](#) for more!

