

Exemplary Digital Services Enabling Open Science

Main Authors

- ◆ **Sasha Woods**
Earthwatch Europe
(Cos4Cloud)

- ◆ **Sona Arasteh**
Max Weber Stiftung
Eliza Papaki
DARIAH
(TRIPLE)

- ◆ **Gideon van den Berg**
ESADE
(CS3MESH4EOSC)

This joint document was produced with the support of Trust-IT Services, provider of the Horizon Results Booster, funded by the European Commission. Content was provided by projects and project groups that took part in the Horizon Results Booster.

Disclaimer

The information, views and recommendations set out in this publication are those of the projects that took part in the Horizon Results Booster and cannot be considered to reflect the views of the European Commission. The Horizon Results Booster is funded by the European Commission N° 2019/RTD/J5/OP/PP-07321-2018-CSSDEVIR-CSSDEVRI01.



Table of contents:

| | |
|--|-----------|
| Topic Overview | 5 |
| 1.1 The projects and EOSC | 5 |
| 1.2 Key challenges | 6 |
| Enabling Open Science through digital services | 7 |
| 2.1 Enhancing Citizen Science contributions in the EOSC ecosystem | 8 |
| 2.2 Advancing Open Science uptake in Social Sciences and Humanities | 9 |
| 2.3 Unleashing collaborative research potential with Science Mesh's seamless data platform | 9 |
| Project Group | 10 |

Executive Summary

The three projects which jointly produced this document - Cos4Cloud, TRIPLE, CS3MESH4EOSC - have contributed to the development and implementation of the European Open Science Cloud (EOSC) through the Horizon 2020 Framework Programme, focusing on different communities, via the establishment and early uptake of dedicated services.

Learn more about their results, bringing benefits in citizen science, social sciences and humanities, as well as to service providers for research.



1 Topic Overview

The European Open Science Cloud (EOSC) initiative aims to provide a federated and open multi-disciplinary environment where researchers, innovators, companies and citizens can publish, find and reuse data, tools and services for research, innovation and educational purposes.

EOSC is envisioned to foster a change of culture across scientific communities and research infrastructures towards Open Science, developing a ‘Web of FAIR Data and services’ for science in Europe upon which a wide range of value-added services can be built and exploited across countries and scientific disciplines.

1.1 The projects and EOSC

The core elements for EOSC to be able to operate in an heterogeneous landscape of e-infrastructures and service providers – the smallest common denominator – has been called the “Minimum Viable EOSC (MVE)”¹.

Cos4Cloud is the first use case and pioneer in integrating citizen science services in the EOSC. One of the project’s goals has been to create a Minimum Viable Ecosystem of Citizen Observatories (COs-MVE) and demonstrate that the MVE proposed by Cos4Cloud for COs can be scaled and the services developed extended to other domains – meeting the requirements of new EOSC users such as the public sector or the industry.

The **TRIPLE** project tackles a phenomenon especially prominent in the social sciences and humanities (SSH): The availability of research data and publications does not only depend on their accessibility but on the language and discipline they have been published in. The multilingual discovery platform GoTriple, developed by TRIPLE, is designed to enhance accessibility to SSH resources.

CS3MESH4EOSC has been developed to target the Cloud Services for Synchronisation and Sharing (CS3) community. This community is a group of providers, developers, and users of innovative storage and sync & share systems. CS3 has been widely deployed in research and education, primarily by e-infrastructure providers, NRENs (National Research & Education Networks), and major research institutions. These services are integrated into the daily workflows for hundreds of thousands of users. Crucially, however, these services remain largely disconnected, and they are developed and deployed in isolation from each other. Herein lies the entry-point for CS3MESH4EOSC, an initiative which tackles this challenge head-on.

By establishing this collaboration, Cos4Cloud, TRIPLE and CS3MESH4EOSC are working together to disseminate their practical results and showcase their potential impact on the uptake of Open Science practices across different communities.

1. EOSC Multi-Annual Roadmap 2023-24 https://eosc.eu/sites/default/files/2022-05/20220523_MAR_02_GL.pdf

1.2

Key challenges

- Citizen science has the potential to support both progress towards and monitoring of the SDGs, by increasing availability, analysis and exchange of data, and provision of opportunities for science education, increasing public engagement, and collaboration across society.
- Discovery of resources, in particular in the SSH domain, can often be fragmented, scattered across different platforms and hidden behind institutional access policies and paywalls. This creates further issues with regards to collaboration among researchers and providing researchers with a broader audience, enhancing thus the societal impact and reuse of this research.
- While cloud computing has revolutionised scientific research by enabling researchers from diverse domains to collaborate remotely on challenging scientific problems, the lack of standardisation and interoperability among cloud services has created fragmentation and hindered the users' ability to collaborate and exchange information freely.
- Research organisations are often concerned about sensitive data that may be compromised by unauthorised parties. In addition, there is the necessity to ensure that they have control over their data and that it is stored in compliance with relevant regulations. This is compounded by the fact that there is a growing demand for data to be made openly accessible for reuse by other researchers, while also maintaining privacy and security standards. Secondly, research organisations are also wary of vendor lock-in, where they become dependent on a single cloud provider for a specific service. This issue is very much related to the challenge of interoperability. A lack of interoperability creates downstream challenges for end-users, with the often-high costs of switching from one provider to another leading to vendor lock-in.
- The exponential increase in the amount of data produced by the scientific community in the past decade has led to the demand for better data management and stewardship practices. To help scientists leverage this data, FAIR (Findable, Accessible, Interoperable, and Reusable) data principles have been called for. Ultimately, applying FAIR principles helps maximise the impact and value of research outputs. Moreover, not applying these principles in everyday research practices causes several overlapping problems, often with financial ramifications. In fact, the lack of FAIR data has been estimated to cost the EU approximately EUR 10.2 billion per year (European Commission 2019).

2

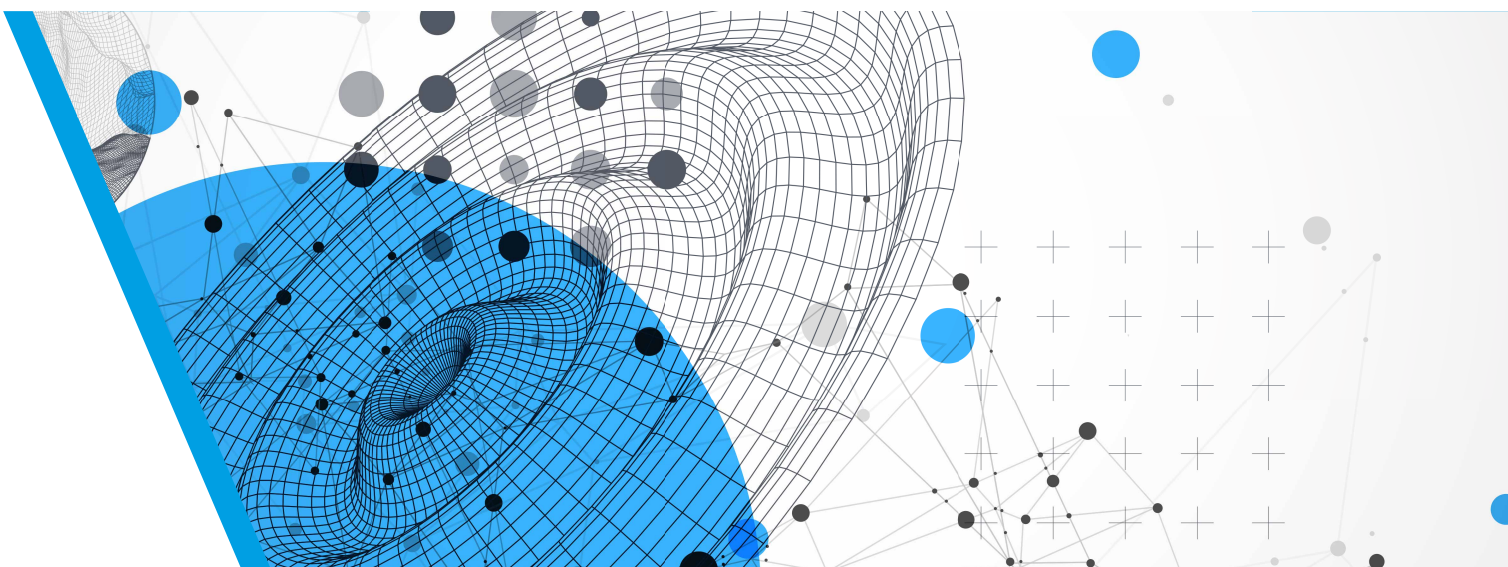
Enabling Open Science through digital services

Ensure that citizen science services continue to be made available on the EOSC to:

- Improve citizen science as an approach for decision-making and SDG monitoring
- Create and strengthen networks of citizen observatories
- Increase integration of citizen science data with existing SDG monitoring and reporting

Truly enforce Open Science and Open Access by:

- Allowing users to discover and visualise search results through a series of innovative services and tools
- Developing software following open standards and making it available under permissive open source licences. The advantage of the open source approach is that the software can easily be migrated to other systems and thus lock-in effects are avoided. This also means that the innovative visual discovery services developed in TRIPLE can be easily connected to other data sources and integrated in other platforms and infrastructures.
- Enabling visual discovery services, such as in GoTriple, that do not only include scholarly articles, but enable a fully integrated visual search with a wide variety of scholarly outputs, including research data and scientific projects. This provides an important contribution to the digitisation of science that makes research outcomes beyond scientific publications increasingly important.



2.1

Enhancing Citizen Science contributions in the EOSC ecosystem

Citizen science is work undertaken by educators and scientists together with citizen communities to advance science, foster a broad scientific mentality, and/or encourage democratic engagement; allowing people in society to join the debate about complex modern problems (Ceccaroni, 2017).

Cos4Cloud has developed innovative technologies to support citizen science observatories by improving:

- data quality, via simplifying expert verification, or using AI to improve identification
- data interoperability, by facilitating integration between data sets
- data accessibility, by processing data and ensuring open access
- and the sustainability and maintenance of citizen observatories, for example, through low-cost technology that can be used, adapted and replicated (Woods, 2022)

Cos4Cloud made these services available via the EOSC Marketplace², a virtual space aimed at the European scientific community, so that anyone interested in creating or improving their citizen observatory can use them.

2.2

Advancing Open Science uptake in Social Sciences and Humanities

Open Science and its infrastructural components will only be truly open if they enable open research and data workflows in all disciplines and scholarly domains, not only in the STEM fields that traditionally have a stronger impact on the development of Open Science than the SSH domain. Integrating flagship domain-specific data services to the EOSC service portfolio is a crucial step towards making broad disciplinary inclusion a reality in EOSC.

Over the course of its lifespan, the TRIPLE project has developed an innovative multilingual and multicultural discovery platform for the social sciences and humanities (SSH), called GoTriple³. It provides a single access point that allows everyone to explore, find, access and reuse materials such as literature, data, projects and researcher profiles at European scale.

The TRIPLE project supported SSH in several ways. In particular, TRIPLE:

- imagines new ways to conduct, connect and discover research;
- facilitates interdisciplinary projects and large-scale scientific collaborations;

1. cos4cloud-eosc.eu/services

2. gotriple.eu

- supports scientific, industrial and societal applications of SSH science;
- promotes cultural and linguistic diversity within Europe;
- connects researchers and projects with other stakeholders: citizens, policy makers, companies, enabling them to take part in joint research projects and to address some of their issues.

The GoTriple platform is among the very few European discovery services designed with the specific needs and epistemic traditions of SSH disciplines in mind. By including GoTriple on the EOSC Marketplace, the European SSH communities at large can discover and reuse SSH resources across disciplinary and language boundaries. The GoTriple platform is envisioned to play a crucial role in breaking down the silos that currently challenge multidisciplinary research both within and across the SSH domain. It brings together digital scholarly objects of all kinds from a wide range of databases, data repositories, publishing and aggregation services to promote findability and reduce fragmentation within SSH.



Unleashing collaborative research potential with Science Mesh's seamless data platform

CS3MESH4EOSC's main asset - the Science Mesh - addresses the challenges of fragmentation in file and application services, digital sovereignty, and the application of FAIR principles in the everyday practice of researchers. Using vendor-neutral Application Programming Interfaces and protocols, the Science Mesh will create an interoperable federation of data and higher-level services to enable friction-free collaboration between European researchers. Ultimately, the Science Mesh will engage with researchers by integrating with their local sync-and-share service providers, enabling them to easily implement and embrace Open Science practices.

The Science Mesh key technology consists of an interoperability platform (or 'federated layer') which enables different EFSS services to communicate with one another. Through this federated layer, the Science Mesh focuses on enabling four specialised data service categories:

- **Open Data Systems:** end-users will be able to organise data via tagging and metadata assignment, creating a published and referable dataset that will ultimately be deposited to an open data repository, archive, or library for curation and long-term archiving. The Science Mesh will make use of persistent identifiers (e.g., DOIs and ORCID iDs) to facilitate this.
- **Collaborative Documents:** end-users will be able to use collaborative editing applications via the Science Mesh to simultaneously edit various types of documents.
- **Data Science Environments:** end-users will be able to access remote execution environments to replay (and modify) analysis algorithms without a need to set up upfront accounts in the remote system. This would be accessible via the web interface at the remote sites of researchers, allowing them to work on algorithms and data processing programs

interactively.

- **On-demand data transfers:** this data service is dedicated to the integration of large data transfers in the Science Mesh, allowing efficient data-based collaboration on on-demand basis - opposite to 'a priori' planned data processing workflows

Through its federated layer and specialised data services, the Science Mesh intends to combine

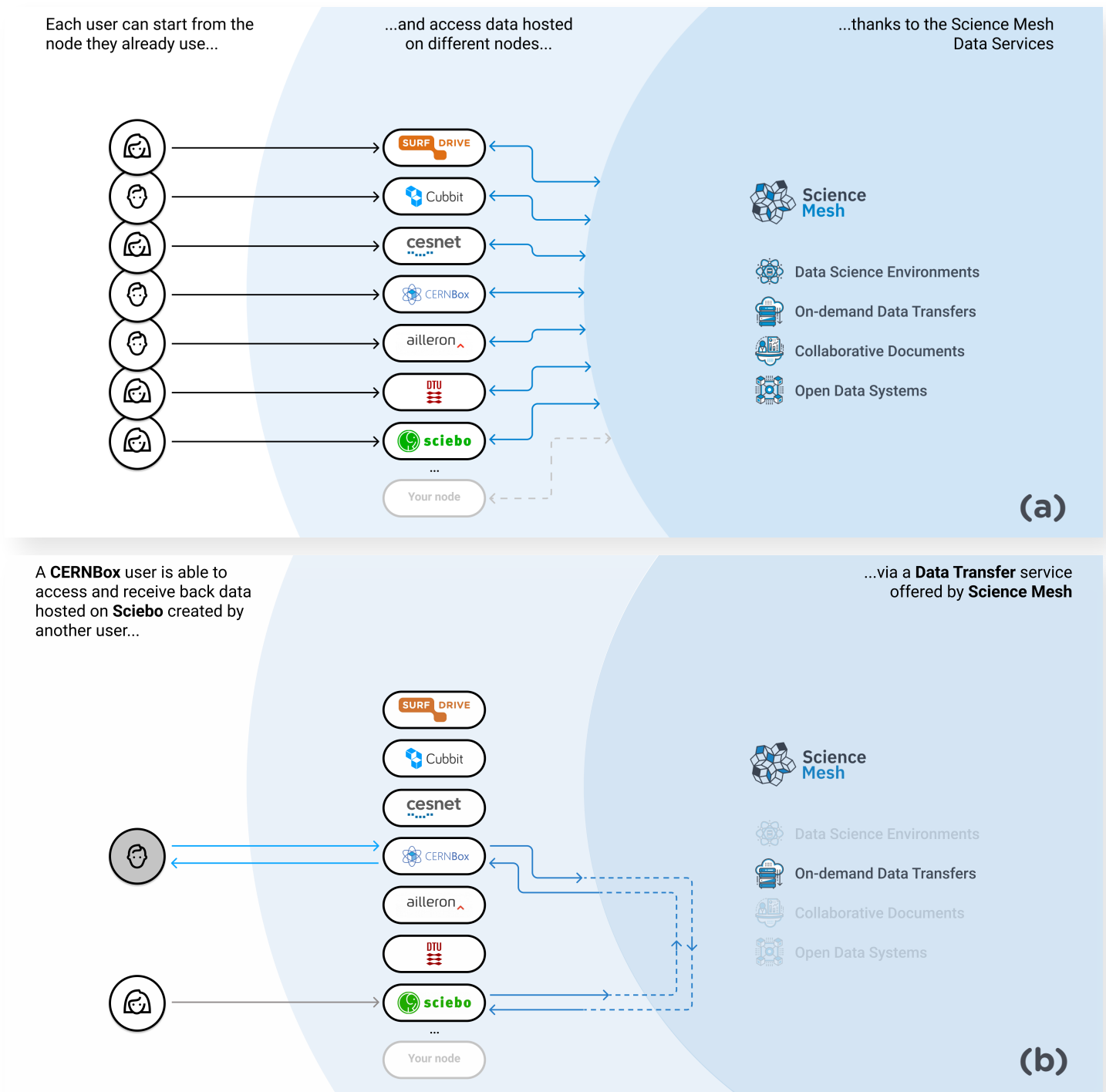


Figure 1. Schematic diagram of the Science Mesh in action. (a) The overall structure, with end-users accessing their data on nodes they already use. Then, through the federated layer, end-users can collaborate with users working through different nodes. In addition, all users have access to four powerful data services. (b) An example of a CERNbox user collaborating with a Sciebo user and sharing data using the On-demand Data Transfer service



Project Group



Cos4Cloud

cos4cloud-eosc.eu



Triple

project.gotriple.eu



CS³MESH⁴EOSC

Connecting European Data

cs3mesh4eosc.eu

Capture QRcode
or follow this URL

horizonresultsbooster.eu





HORIZON RESULTS BOOSTER

April 2023