

EOSCPilot Project

The European Open Science Cloud will offer 1.7 million European researchers and 70 million professionals in science and technology a virtual environment with open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines by federating existing scientific data infrastructures, today scattered across disciplines and Member States.

The EOSCpilot project¹ has been funded to support the first phase in the development of the European Open Science Cloud (EOSC). It will:

- Propose and trial governance frameworks for the EOSC and contribute to the development of European open science policy and best practice;
- Develop a number of demonstrators functioning as high-profile pilots that integrate services and infrastructures to show interoperability and its benefits in a number of scientific domains; and
- Engage with a broad range of stakeholders, crossing borders and communities, to build the trust and skills required for adoption of an open approach to scientific research.

Interoperability and Service Architectures

EOSC services will need to be distributed - a decentralised System-of-Systems (SoS) based on collaboration and coordination, with components independently provided and managed by local, regional, national and international organizations. The EOSCpilot project will demonstrate the usage of existing digital infrastructures in a combined manner, across disciplines and borders to realise the both the goals of the FAIR principles for sharing data, and equivalent principles to sharing software, methodologies and all aspects of the research lifecycle. It will determine the interoperability needed for an efficient use of IT services and equipment, and of our precious scientific data, which obstacles we are facing, and what solutions are already available.

EOSC will provide a comprehensive and evolving set of services supporting an “open science”-friendly knowledge production lifecycle which will enable the outputs of the research process can be created, deposited, analysed, published and preserved, and can be discovered, accessed, used and reused. Such EOSC services will include:

1. scientific instruments;
2. data;
3. applications, workflows, software;
4. storage, compute and network connectivity;
5. written knowledge (e.g scientific publications, educational and training resources);
6. services for enabling federated access, like federated identity service provisioning, authentication, authorization, and accounting; and
7. collaborative services enabling the sharing, use and reuse of digital capabilities.

These services will need to support a broad spectrum of user groups from researchers to citizens, according to the policies of their funding agencies. The EOSC service portfolio will need to grow incrementally according to stakeholder-defined principles and user needs. EOSC will benefit from a clear and transparent portfolio presentation of all EOSC services,

¹ <https://www.eoscpilot.eu>

harmonised access policies and methods, based on clear points of access and support, integrated body of policies for access and use, and the human networks supporting its services.

EOSC will also require a lightweight service management layer including policies for service provisioning, processes, procedures and related resources capabilities. This management will span organisations to support collaboratively-provided services and is intended to support provision of long-term, high-quality services.

Science Demonstrators

EOSCPilot will follow an evidence-based approach, using a number of Science Science Demonstrators to show the relevance and usefulness of EOSC services and how they enable data reuse, and drive EOSC development. Current science demonstrators focus on Environmental & Earth Sciences, High Energy Physics, Social Sciences, Life Sciences, Physics, whilst additional areas will be "on-boarded" through Open Calls. These demonstrators will investigate what is preventing the efficient common use of the e-infrastructures, and determine the current showstoppers and solutions found by some communities and infrastructure providers. Through these we will demonstrate that obstacles in data and infrastructure interoperability can be overcome – and will discover additional challenges, of which we might not have been aware.

Governance and Policy

The EOSC governance must be capable of supporting the definition, management and coordination of EOSC components and service providers without imposing a strict hierarchical model or supply chain. This will involve defining organisational, operational and managerial interoperability, whilst recognising that the SoS will evolve and adapt over time.

The realisation of EOSC will be a rich ecosystem consisting of many different stakeholders and a heterogeneity of services and technologies which will need to operate across a wide range of organisational, community and national borders. Moreover, to successfully operate EOSC must encompass many existing structures and services both at the national and international level. To be successful, any governance for the EOSC must adopt the same principles of openness, transparency and inclusion that EOSC seeks to encourage amongst the research community.

The governance will need to accept policies reflecting the community-based principles, rules and procedures in place with built-in incentives for service provisioning and sharing, and use and non-discriminatory policy-based access; and policies defining the principles of engagement for both service providers and users. Services will need to conform to similar principles to the FAIR principles for data, which would include additional principles such as Managed (professionally delivered and supported) and Sustainable. On the consumer side, conditions for access to services will need to be clear and transparent, and that functionality and levels of services will need to be defined.

The "governance" framework must enable and encourage engagement from the key stakeholder communities such as European e-Infrastructures, Data/Research Initiatives, Service and cloud providers, Research funders, Cloud community, Research Communities and Institutions, Research Infrastructures, Policy makers.

The "governance" framework must also enable interoperability and co-ordination within a number of different domains: legal interoperability, interoperability of organisational

processes, technical interoperability, operational interoperability, data\information interoperability, skills interoperability, research challenge interoperability².

Governance and engagement will need to be optimised for the different communities and domains of interoperability; so difference mechanisms and processes will be needed. The “governance” framework would therefore, itself, be a rich ecosystem of different engagement mechanisms, processes, and principles dependent on the community and the domain. EOSCPilot will develop and trial evidence-based solutions to this problem by

- analysing principles of engagement from already established organisations in order to establish a minimum set of principles and practice which will be in harmony with widely accepted working practices;
- seek consensus from a broad community of expertise to identify high priority policy drivers and constraints and evidence based ways to address these in the areas of open science and open scholarship; data protection and assurance, procurement; and ethics;
- analysis and recommendations as to different funding and business models including the need for framework contracts compliant with the EC Public Procurement Directive (Directive 2014/24/EU) to will allow easy access to commercial solutions that meet the community’s needs in an efficient way under competitive terms, and the legal structures needed to manage these effectively;
- consolidating the above with input from the technical, service and interoperability demonstrators to form a governance framework;
- working closely with the European Commission to pilot, trial, build an evidence base and establish consensus on the governance framework with a broad community.

Skills and Community

As important as the technical and governance components will be, the importance of providing expertise should not be overlooked. Building upon existing work, EOSCPilot is establishing a skills framework that will help infrastructures, institutions, and other stakeholders to find, access and benefit from relevant skills development activities. Such activities, including those we pilot in EOSCPilot, will support development of data stewardship and the related roles required to sustain data science services exemplified by the Science Demonstrators. This will inform EOSC alignment and coordination of human resource infrastructure, in parallel with its similar efforts for Research and e-Infrastructures.

Conclusion

EOSCPilot has been funded by the European Commission to support them in defining and implementing both the initial phases and longer term vision of the EOSC. Our role is to develop evidence-based solutions to the many challenges that EOSC must overcome, and build consensus on the future direction of the EOSC. We welcome the opportunity to engage with both the Commission and early adopters to rigorously analyse the issues, provide evidence based proposals, and build community consent on how to realise the vision of the European Open Science Cloud.

² Based on European Interoperability Framework for the Digital Single market - https://ec.europa.eu/isa2/eif_en