

How EIROforum could contribute to the EOSC

Data culture – promote open science and FAIR data principles across our user communities

As the interest in open science increases within the research communities we support, the need to train scientists and other concerned parties on how to introduce open science concepts and FAIR data principles into daily research work becomes more pressing. The EIROforum members are willing to incorporate open science and FAIR principles into their already existing training programmes. However, developing or expanding services and preparing data for consumption by audiences that are not involved in the data collection requires additional effort and time that is frequently not foreseen in the budget of research infrastructures.

Research data repositories – ensure the openly available scientific datasets continue to be curated over the long term and maintained through certified digital repositories

As curators of trustworthy datasets that are widely used by the global research community, the EIROforum members recognise the importance of certification programmes for digital repositories and their long-term preservation. Adoption of internationally agreed certification programmes should be a goal for EOSC that can be progressively achieved over several years. Ensuring long-term preservation of digital repositories goes far beyond the management of the research data itself requiring changes to processes across the full research lifecycle with implications that should not be underestimated if the '*long-term sustainability of open research data*' as stated in the EOSC declaration is to be achieved.

EIROforum members manage production-level Research Infrastructures and are willing to share their experiences on operating data-services for the international community over several decades, seamlessly handling technology, service and evolution of requirements, with other research infrastructures.

Service deployment – champion the integration of commercial cloud services in to the EOSC

The EOSC declaration recognises the role commercial operators will play in service deployment. Since 2010, the EIROforum members have collectively invested in the adaptation of commercial cloud services to support their scientific programmes via the Helix Nebula initiative and more commercial providers. We see the use of commercial services as progressively expanding and have adopted a hybrid cloud model in order to profit from the innovation and expanded capacity they can bring. Establishing common quality levels and

standards for all service providers, public or commercial, will facilitate integration and create market opportunities. The EIROforum members are prepared to expand their activities by involving more organisations that are willing to invest and to broaden the scope of the commercial services to be deployed.

High Performance Computing and the European Data Infrastructure

The EOSC declaration indicates that “*supercomputing and data infrastructure could support the European Open Science Cloud by providing data access and advanced computing and data management services.*”

EIROforum recognises the importance of HPC and is willing to work with the supercomputing community to accelerate its transition to support data intensive science and on-demand access. The EIROforum members are experiencing a growth in the variety of architectures that are deployed to support data intensive science by public and commercial operators. This heterogeneity needs to be taken into account in the planning of HPC services made available via EOSC and will increase the effort required by the user communities to develop and maintain software capable of exploiting such a variety of architectures.

The democratisation of HPC, requiring easier means of access and integration into the cloud model, also has implications for the proposed European Data Infrastructure which must include Europe’s major data providers and become an integral part of the EOSC.

Innovation

EIROforum experience confirms the importance of innovation as highlighted in the EOSC declaration. Existing public private partnerships, including CERN openlab, Helix Nebula, the ATTRACT programme and more recently LEAPS, are vehicles for innovation that bring together research infrastructures with industry. These collaboration models, that operate beyond the traditional grant-based research and development approach, can contribute to EOSC by leveraging joint investment from the public and private sectors.

Governance model – take an active role in the governance of EOSC

The EOSC declaration states that “*The role of ESFRI and EIROFORUM research infrastructures and organisations in the EOSC will be enhanced, Member States and the European Commission made significant investment; research infrastructures should be 'the steward of the community of standards' and provide scientists with a ramp-up for the utilisation of the EOSC.*”

We think more detailed information about the EOSC needs to be shared with the Members States in order to initiate or strengthen their engagement in the EOSC. EIROforum and the research infrastructures could use their strong connection with Member States, that provide

both financial and political support to the scientific communities driving them, to facilitate the engagement of Member States in the EOSC.

Given the EOSC's broad and ambitious scope, EIROforum considers it essential that DG RTD, CNECT and GROW work closely together on the subject and is willing to continue the joint meetings it initiated with these directorates in November 2016.

EIROforum members are key actors in data intensive science and curate some of Europe's most important and popular datasets. Such datasets and associated services will be major assets for the EOSC, but access to and exploitation of the data needs to be carefully planned. The EIROforum members are willing to work with the EC to define a mutually acceptable role in the governance of the EOSC.

Author(s)

Professor Iain W. Mattaj, FRS, FMedSci, Chair of the EIROforum Council

Bob Jones, EC projects coordinator, CERN