



Authors:

ENVRI-FAIR - Andreas Petzold, Ari Asmi

EOSC-Life - Niklas Blomberg, Michael Raess

ESCAPE - Giovanni Lamanna, Ian Bird

PaNOSC - Andrew Götz, Rudolf Dimper

SSHOC - Ron Dekker, Ivana Ilijasic-Versic

Object: Science Cluster inputs to the EOSC Association for the INFRAEOSC destination in the next RI Work Programme (2023-2024).

Title: *“Fostering Open Science with Science Clusters of ESFRI research infrastructures through both community-based and interdisciplinary projects and common actions”*

Expected goals:

Implementation of EOSC

- Addressing the scientific needs for additional services, interoperability and alignment

Metadata and Data Quality

- Transversal actions towards FAIR stewardship and continued innovation of scientific software

Widening and Outreach

- Open innovation environment for research data, knowledge and services with engaged stakeholders and organisations

Sustaining EOSC

- A sustainable model for science communities to contribute to the EOSC

Preamble

Current landscape.

Research Infrastructures such as the ones on the ESFRI roadmap, are characterised by significant volumes of data they generate and handle. Thousands of researchers across scientific disciplines and other potential users are interested to use these data via Open Access policies. Data-intensive research, and effective data preservation for immediate and future sharing and re-use, are fundamental components of ESFRI projects and landmarks, contributing to their increasing role in responding to societal challenges such as climate change or global pandemics. Before the advent of the Science Clusters, researchers were confronted with a highly fragmented research data landscape and cooperation among ESFRI research infrastructures was community based, with limited scope and focus. The Science Clusters, namely ENVRI-FAIR, EOSC-Life, ESCAPE, PaNOSC and SSHOC, of ESFRI and other pan-European research infrastructures, are successfully helping to address the current situation. They are working with the European Open Science Cloud (EOSC), actively contributing to the setting up of dedicated FAIR services, and promoting practice for open science.

The Science Clusters today ensure strong connections between research infrastructures on the ESFRI Roadmap, between the Science Clusters themselves and jointly towards the EOSC. They are also seen as key for the transitioning of ESFRI RIs from servicing only their communities towards more integrated services in response to the Grand Challenges, and transforming their research products for economic and societal users. In addition, Science Clusters have anticipated the need for interdomain shared views for the benefit of EOSC and excellent science in Europe. They are therefore collaborating,

through regular meetings, and have produced common position papers, co-organised workshops and planned operative trans-disciplinary work programmes. Today the Science Clusters are an integral part of the EOSC, contributing to its development and implementation process, while their services and project outcomes are now forming pillars of the emerging EOSC fabric and contribute significantly to the content of EOSC. As important partners of EOSC, Science Clusters stimulate Open Science practices, cross-domain interoperability, and coordination of the scientific communities, and connecting the numerous Universities and national research institutes concerned by the ESFRI projects' partnership within each cluster.

The next phase.

In the Horizon Europe work programme, the Science Clusters are encouraged by both their own partner ESFRI research infrastructures, as well as by the relevant national institutes, to strengthen their mission in enhancing the participation of science communities by structuring and implementing European Open Data Commons.

The connection between EOSC and the Science Clusters aims to fill the EOSC Exchange with data and services and focus on interoperability. They also work maintaining and reinforcing the connection with corresponding domain-based research consortia and scientists committed to open science. Future projects should aim to support strengthening and advancing the current Science Clusters' roles of "developers", "aggregators" and "integrators" of data services and community-based projects for open science and innovation. Such projects would leverage crosscutting and cross-domain science, potentially connected with European sectoral data spaces¹, where scientific publications and major discoveries are linked to accessible digital objects for reproducible workflows; providing training, committing to education, and engaging citizens in science.

Upcoming proposals within Horizon Europe need to further address the stewardship of digital objects from research infrastructures, and build on previous and current developments of the ESFRI Science Clusters. This means engaging with new communities (even globally), and new projects that have been included in the recently updated ESFRI roadmap², in order to accelerate the uptake of these best practices more broadly. The next actions should include the integration of ESFRI Science Clusters' major thematic services as part of EOSC for domain-specific data management as well as the final implementation, support and operation of open-science virtual research environments for depositing, curating, and analysing data. In many areas, such VREs are expected to be created from the user communities, and the role of the Science Cluster is to build the infrastructure and services to support such activities.

The Science Clusters, meanwhile, are looking forward to collaborating with the EOSC Association. The five major topics³ addressed by the EOSC Association through established dedicated task forces have the consideration of the Science Clusters. Furthermore, such a collaboration would target in particular subjects in which the ESFRI research infrastructures can address the current fragmentation and/or limited impact such as: solutions for sustainable policies and services, schemes for credit systems and frameworks for rewarding scientists and supporting careers.

¹ Health, Industrial and Manufacturing, Agriculture, Finance, Mobility, Green Deal, Energy, Public Administration, Skills

² <https://www.esfri.eu/esfri-roadmap-2021>

³ Implementation of EOSC, Technical challenges on EOSC, Metadata and data quality, Research careers and curricula, Sustaining EOSC.

Expected Impacts

1. The five ESFRI Science Clusters within the EOSC will continue to evolve, driven by the growing scientific need for intra-domain interoperability and alignment. While there will remain organisational diversity between the clusters building on the established federations of domain based (ESFRI or other world-class) research infrastructures, there are common needs within for a collaborative work programme to the benefit of the research communities.

Science Clusters as platforms for scientific interoperability in EOSC will extend on a longer term, either via cluster Consortium Agreements, or by leveraging the existing domain management boards to sustain their commitment in supporting the cross-fertilization, co-development and operation of the developed open-science resources, including thematic virtual research environments. This will ensure that the individual and collective RI data resources, computational and workflow services are exposed, connected and operated within the EOSC.

A dedicated “destination action” aimed at supporting the Science Clusters to their longer-term platform evolution would foster transnational cooperation. It will facilitate the sustainable cooperation among ESFRI projects and support the ESFRI science Clusters for their long-term collaboration.

2. In line with the objectives of Open Science, the five Science Clusters need to structure a specific transversal action towards FAIR stewardship and continued innovation (and legacy) of scientific software. Coordinated activities among the five Science Clusters would enable new and interdisciplinary research, leading to new insights and innovation for science and for society at large, committing research communities, Universities, and Institutes to a federated activity for training and skills development schemes.

- Encourage collaborative development and operation of common infrastructures, tools and methods for research software, best practices, techniques, and technologies, where appropriate and useful to the research infrastructures, and in accordance with the FAIR principles.
- Cross-cluster collaboration and cross-fertilization/synergies between the RI's and with academia and industry on new digital technologies and tools based on exa-scale data management (Artificial Intelligence/Machine Learning, Quantum Computing, Real-time/Edge Computing Data workflows etc.) with a new focus on education and training by research programmes facilitating early career development of new generations.
- Support research software collaboration across Science Clusters, technology interest groups/foundations, while leveraging existing international initiatives and encouraging national efforts from many countries on developing open-source solutions.
- Organize a regular conference series on broad cross-domain Scientific Computing covering all aspects and building on existing domain- and technology specific conferences (e.g. CHEP, HPC).

A dedicated structuring action within the next HE work programme is needed to realise this impact. It will leverage the current achievements of the Science Clusters and enhance the inter-domain potential for excellent science in Europe. The Science Clusters are ready to work with emerging national initiatives on data-intensive science to structure further the European part of a global coordinated effort.

3. Contribute to the creation of a cross-border and multi-disciplinary open innovation environment for research data, knowledge and services with engaged stakeholders and organisations.

- Support and help new research infrastructures develop and integrate with the EOSC multi-cluster ecosystem and collaborate with the EOSC Association. This activity would include mostly ESFRI roadmap infrastructures, but could also include key Member State run national infrastructures of international importance and use.
- Ensure cooperation of research infrastructure-specific and national platforms with EOSC; this could be synergistic and build on common developments and tools, or foster coexistence and collaboration. It can also stimulate a level-playing field within Europe.
- Leverage multi-domain data and a large spectrum of digital objects populating the main European Data Spaces to incentivize some explorative cross-cluster and cross-border open challenges for science, economy, technology, and society.
- Support the Science Cluster role in operationalising the European Research Area (ERA). Organise dialogue, share strategy with ERICs and other research infrastructures for a better link to governance of ERA aiming at matching policy priorities with needs of ERICs' operation and their exploitation by the scientific communities.
- Commit in a structured action in support of a stable Citizen-Science programme around missions for the exploitation of digital objects from the ESFRI RIs.

An action is needed within the next HE work programme to develop and implement the data spaces concept and to understand how the diverse research products of the Science Clusters will populate these data spaces. Although the European Commission is expecting diverse stakeholders committing to data spaces, the scientific communities engaged in the construction and operation of next generation ESFRI facilities need a dedicated support to be part of a more global and resilient plan in Europe. The need to combine science, economy, technology and society through effective visions and programmatic actions is crucial. Open Science and Open Innovation today implies unifying forward-looking approaches for any ESFRI facility integrating scientific needs, key-enabling technology advancement, socio-economic impact optimization, environmental impact minimization, energy and other resources management and industrial cooperation. Such an approach needs multi-domain data and will imply the generation of a large spectrum of digital objects populating the main European Data Spaces. A cluster-platform will guarantee the cross-border environment for addressing all these societal challenges. The cluster programmes can also support such developments in regional and Member State levels, bringing the EOSC and ESFRI products closer to the user communities, and increase the applicability to smaller national infrastructures, observatories, and research performing organisations.

4. Develop the sustainability model for the EOSC science communities; foresee a model for sustainable operation that makes use of the research infrastructures to run sustainable services in the EOSC.
 - Develop a sustainable business model for the EOSC and the research infrastructures participation therein
 - Involve research infrastructures and funding agencies in defining the procurement model as part of a sustainable business case, using experience from the current EOSC Future procurement action to guide the model.
 - Develop a mechanism by which computing and storage resources are made available to general scientific users (outside of the research infrastructure-specific resource pools).
 - Support the implementation of a framework and a policy for rewarding scientists committed to Open Science.

An action where Science Clusters collaborate with the EOSC Association towards the above impact is needed.