SCIENCE-CLUSTERS.eu
Research Infrastructures for Open Science

Science Clusters

Position statement on operational commitment to EOSC and Open Research

List of authors:

ENVRI - Andreas Petzold, Anca Hienola
EOSC-Life - Jonathan Ewbank, Jonathan Tedds
ESCAPE - Giovanni Lamanna, Ian Bird
PANOSC - Andrew Gotz, Jordi Bodera
SSHOC - Franciska de Jong, Bonnie Wolff-Boenisch

1 March 2024

ENVRI Community
ESCAPE European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures
EOSC-Life
panosc photon and neutron open science cloud
SSHOpenCluster Social Sciences & Humanities Resources

Table of Contents

Executive Summary

The need for close collaboration between the European Strategy Forum on Research Infrastructures (ESFRI) and the European Open Science Cloud (EOSC) was addressed in 2019. Funded by the European Union (EU) within the H2020 framework programme, five projects linked EOSC with ESFRI-endorsed Research Infrastructures (RIs) and other world-class RIs, forming five Science Clusters[1].
These Science Clusters have taken different paths to become long-term structures: through Memoranda of Understanding (SSHOC), international Collaboration Agreements (ESCAPE), building on established consortia of thematic RIs (EOSC-Life) or in an on-going process, relying on existing cooperative frameworks or groups of domain-based RIs (ENVRI, PANOSC). They enable continued community-based commitment to support data-intensive research and open science.

The Science Clusters, representing distinct scientific domains, integrate multiple RIs and collaborate with pan-European e-infrastructures. Their actions, validated by the EOSC Future project (and to be further developed in EOSC Beyond), contribute to harmonized models for data access, tools, workflows, and training. Operating as a cluster of clusters in projects like OSCARS[2] and EVERSE[3], they align with Horizon Europe goals, ensuring the successful implementation of EOSC.

After five years of collaborative efforts, including interactions with the European Commission, EOSC Association, ESFRI-EOSC task force, and e-Infrastructure Reflection Group (e-IRG), there is a need for a fresh articulation of the Science Clusters' vision for the future, as reflected in this third position paper.

This position paper serves as a reminder that:

- The Science Clusters prioritize advancing researchers' access to data, tools, and resources for data-driven science. A cross-border open innovation environment for *FAIR* (Findable, Accessible, Interoperable and Reusable) data management and analytical services, respectful of *CARE*[4] (Collective Benefits, Authority to Control, Responsibility and Ethics) principles and attentive to digital security, contributes to the sustainability of ESFRI and pan-European RIs. Synergies between Science Clusters enhance researchers' efficiency and productivity through open-science standards and services.

This position paper:

- acknowledges the progress accomplished, updates the common work plan from 2021[5], expresses the Science Clusters' determination to establish thematic community-based 'EOSC Science Cluster Nodes', contingent upon resources and explains the potential implications on the interconnection with the EOSC EU Node and national nodes.
- Emphasizes adaptability of the Science Clusters to a unified vision capable of accepting different needs within EOSC and outlines their operational role within the EOSC Federation.
- Demonstrates the requirement for urgent support of a targeted Science Cluster destination action within Horizon Europe and future framework programmes in order to enable the operational role of the Science Clusters.

[1] https://science-clusters.eu/
[2] https://oscars-project.eu/
[3] https://cordis.europa.eu/programme/id/HORIZON_HORIZON-INFRA-2023-EOSC-01-02/en
[4] https://www.gida-global.org/care
[5] https://zenodo.org/records/4889503

1. EOSC Federation and EU Node

EOSC is envisioned and described as a federation of distributed systems that are combined into a system of systems, through multiple interconnected Nodes. Once implemented, the EOSC Federation would rely on Nodes that can be organised on a national, regional, European, Institutional, or thematic level. This evolving ecosystem of open, trusted, and federated EOSC Nodes would provide access to digital objects from existing (and emerging) RIs in Europe. Those RIs include pan-European or international organisations for research, the European Research Infrastructure Consortia (ERICs) and other world-class facilities and projects, that federate a very large community of European researchers and that promote the storage, sharing, processing, analysis and reuse of their digital objects (e.g., data, publications, software and workflows).

The EOSC Federation is expected to be inclusive, respect the autonomy of organisations and nations, and have significant scientific impact. Its interoperability framework requires the definition of technical as well as legal and organisational aspects before an operational EOSC can emerge. Defining and developing the implementation model of such a federation is a significant challenge that is crucial to the successful realisation of the EOSC objectives.

The EU Node of the EOSC Federation.

The concept of EU Node has been re-introduced recently[6]. The European Commission's DG Connect has publicly procured the management of services for the EOSC Platform to certain pan-European e-infrastructures.

The aim of such a procurement is *"…to build and deploy a fully operational enabling infrastructure for EOSC – referred to as the EOSC EU Node – providing access to a rich portfolio of FAIR data and professional quality interoperable services in all relevant domains from data handling to computing, processing, analysis and storing[7]…"*.

Scenarios describing governance and financing models for EOSC post-2027 have already been proposed and are under discussion[8]. Evaluations in progress include options that either unify or divide specific governance and funding responsibilities between the EOSC Association and the EC.

Stakeholders' consultations are planned during the early months of 2024. These will help the elaboration of potential models of governance and financing for EOSC. EOSC's implementation will enter a new phase, building on the EOSC Future project's achievements, to which the Science Clusters have contributed considerably.

All five Science Clusters, ENVRI (environmental science), EOSC-Life (life science), ESCAPE (astronomy and particle physics), PaNOSC (neutron and light source science) and SSHOC (social science and humanities), work closely with, and proactively support, the EOSC Association Board. They collaborate with the EOSC Federation and aid the thematic Nodes fulfil their role in the European open-science ecosystem[9].

[6] https://digital-strategy.ec.europa.eu/en/policies/open-science-cloud

[7] https://digital-strategy.ec.europa.eu/en/news/commission-announces-winners-eosc-procurement?pk_source=ec_newsroom&pk_medium=email&pk_campaign=Shaping%20Europe%27s%20Digital%20Future%20website%20updates

[8] https://eosc.eu/news/2023/12/eosc-governance-meets-to-advance-post-2027-planning/

[9] https://eosc.eu/wp-content/uploads/2023/11/20231112-Short-paper-on-the-EOSC-Federation-draft-v3.pdf

The EOSC Federation's future governance structure, rules of participation and inclusion criteria should provide the opportunity for assessing the extent to which the Nodes' scientific quality, potential for impact and coherence with the current research landscape contribute to user experience.

To be successful it is key to enhance the scientific value of EOSC, by incentivizing and rewarding collaborations to share knowledge and data as early and as openly as possible. The Science Clusters are well positioned for such an objective, rooted as they are in the research communities themselves. Furthermore, within a large European partnership approach there is currently no other framework as pertinent as the cluster model to guarantee the establishment of links within the EOSC ecosystem between scientists from a wide range of research organisations and disciplines, between partner RIs and Universities, academic groups, industries and citizen scientists.

The five Science Clusters, as they already declared in 2021 in their 2nd position statement document[3], support the ambition of implementing the EOSC Federation as a system of "Web of FAIR Data and Services for Science". Such a virtual structure would be incomplete and ineffective without "community-governed" open science commons co-developed and operated by scientists inspired by shared values, who meet and work together in a virtual global environment to produce open research. The implementation of such a virtual environment is fundamental in order to improve the FAIR data-research practice and results' assessment, to define and promote key enablers such as rewards and incentives, to further skills and education, and engage with society and research infrastructures. It is primarily where our researchers are engaged and is the vital connection point with EOSC.

2. A shared commitment in EOSC

The Science Clusters, ENVRI, EOSC-Life, ESCAPE, PaNOSC and SSHOC are far from being a homogeneous group. They are made up of multiple and diverse RIs and research communities that vary in the approach, FAIRness and maturity of their data, services and scientific research. Their scientific communities have distinct cultures, different approaches and organisations, methodological needs, levels of awareness of data management and services, levels of innovation and technology capacities. In spite of their differences, the five Science Clusters have taken up a similar approach in working towards the mission of building EOSC. Their diversity has not prevented them from cooperating together, aligning approaches in terms of data-intensive research requirements and practices, bridging cultural divides, exploring synergies and building a common vision about the operationalisation of EOSC.

In 2021, all five Science Clusters published a collective, prospective workplan for their sustainable establishment in the EOSC ecosystem. Over the last five years the five Science Clusters have successfully improved the FAIRness of science for researchers and important steps have been taken in building models for collaboration that help increase the trust in infrastructural support for research.

Achievements were centred around:

- improving access to data, tools and services for data-driven and data-based science;
- a reusable cross-border open innovation environment;
- increasing the efficiency and productivity of researchers through open data services and infrastructures for discovering, accessing, and reanalysing data;
- fostering the establishment of, and interoperability with, global standards;
- developing synergies and complementarity between the participating world-class research infrastructures;
- adopting common approaches to data management for economies of scale;
- working towards the creation of a pan-European data access mechanism and to the coordination activities for the EOSC on-boarding process.

Figure 1 provides a schematic view of the way that the Science Clusters have contributed to the creation of a pan-European data access mechanism and to the coordination activities for the EOSC on-boarding process.

Today, the Science Clusters are an integral part of the EOSC initiative. They contribute to its development and implementation. Their services, data and research outcomes now form the core of the emerging EOSC fabric. EOSC is built on relatively abstract concepts, and by necessity, provides generic services. These often need to be refined with specific datasets, enriched with additional, contextual information, and tailored for domain-specific scientific use. Many of these service enhancements can, and will, be carried out collaboratively by the ESFRI RIs. Therefore, despite their disciplinary diversity, they have a lot in common. Fulfilling the expectations of the five Science Clusters' constituent scientific communities needs a high-level of coordination and support that the current Horizon Europe INFRA-EOSC OSCARS and EVERSE projects should provide.
In addition to the joint preparation of several successful cluster-driven project proposals, the solid collaboration established between the Science Clusters is further illustrated by the common web presence now established at "*science-clusters.eu*"[1].

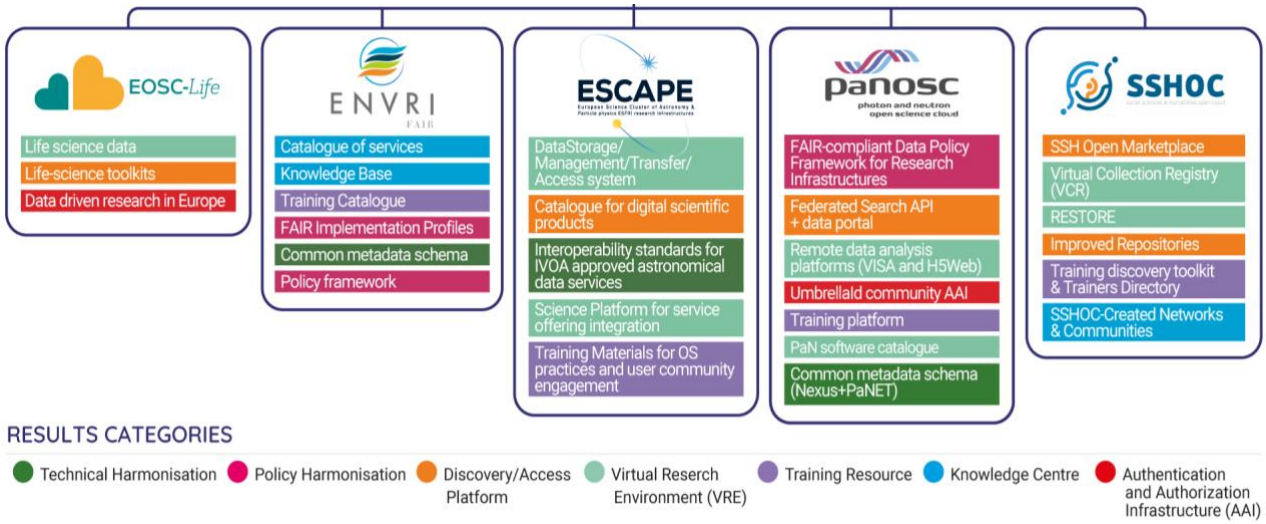


Figure 1: Science Clusters principal results which make up the EOSC ecosystem.

This showcases synergies as well as Science Cluster activities, allowing RI communities to share best practices and encourages other RIs to join the Science Clusters cooperative framework.
The ESFRI and non-ESFRI RIs participating in the five Science Clusters share a common view and call for a common policy and actions to drive interdisciplinarity and to maintain a strong link with scientific communities for the next phase of EOSC, namely by consolidating:

- the exchange of visions and insights with the EOSC Association to enable integration of existing RI data resources, standards, analytical and computational services as a strategic part of the EOSC Federation;
- the support of the Science Clusters in their role of "developers" as well as "operators" of composable data services for the community-based thematic EOSC Nodes for science and innovation, as well as supporting the widespread harmonisation of the scientific culture of their respective fields.

The cluster of clusters approach.
All Science Clusters working together, through a "cluster of clusters approach", are focussed on the specific needs of their domain-based user communities, and the services that they provide to address them. In the context of OSCARS, they will consolidate their current achievements and continue to innovate in a cross-cluster interdisciplinary framework. They offer new opportunities for open science to a broader community, including the "long tail of science". For example, they enable researchers and laboratories who do not have access to the computational resources to manage and analyse large amounts of data themselves, and host or mentor open science projects and services. Such approaches can help avoid unnecessary or duplicate allocation of resources and energy intensive computations, thus reducing the environmental impact.
Such an approach will be realised through the operation of a cascading grant mechanism, and collaborations with national institutes and universities to establish outstanding operational entities in Europe, such as the community-based Competences Centres[10] (OSCARS) and the nodes of a Virtual Institute for Research Software (EVERSE).

The strategic roadmap of the five Science Clusters builds on the past work programme dating from 2021, as illustrated in Figure 2. General Objective (GO) actions target the foundation of the cluster initiatives, such as the deployment and the consolidation of thematic data infrastructures as parts of a federation. Specific

[10] In line with the definition of 4CH (https://www.4ch-project.eu/what-is-a-competence-centre/), in this paper "Competence centres" are to be associated with excellence, training and knowledge transfer, interdisciplinarity, standardisation, and a collaborative and virtual approach of different institutions or departments.

Objectives (SO) include the results of open science pilot projects from the Science Clusters, increasing the number of participating RIs, and targeting the enhancement of researchers' uptake of open science. Operational Objectives (OO) include sustainable operation of "open platform" infrastructures and fostering inter-domain FAIR-data intensive science projects, by permanently establishing an "Operational" sustained mission of the Science Clusters in the EOSC Federation.

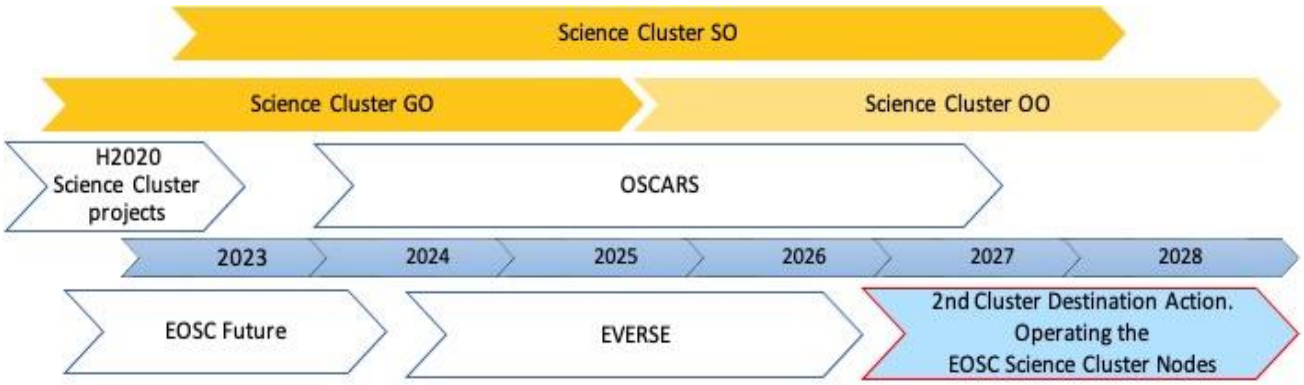


Figure 2 - The Science Clusters' workplan addresses General, Specific and Operational Objectives (GO, SO & OO).

The evolution of the Science Clusters' workplan is inspired by the following objectives:

- **Consolidate clusters and strengthen co-development of common inter-cluster data services.**

Some of the most successful achievements of the Science Cluster projects are the key data services within EOSC for "data provision, discovery, and exploitation", e.g. catalogues, analysis platforms, FAIR data archiving, solving the challenge of research communities/themes. There is also a clear need to support the community-based consolidation of individual cluster services, as well as to fulfil the reference role of the Science Clusters at a pan-European level, to foster inter-cluster projects.

- **Delivering content to EOSC and demonstrating its impact on science.**

The five Science Clusters, while consolidating their thematic data infrastructures, commit to target and pursue specific scientific objectives that will produce open science results, and mentor new impactful trans-RI science projects. This approach is also inspired by the success achieved with the Test Science Projects within EOSC Future.

- **Science Cluster EOSC implementation.**

By working together within and across the Science Clusters as described, we can work with our EOSC Partners to deploy and to operate sustainably the community-based composable virtual research environments (VREs), the domain-based Competence Centres, the system of thematic nodes for continuous innovation of EOSC Exchange services and the training schemes (including the newer Virtual Institute for Software to be established by the EVERSE project). In this way, we can provide a significant step change in establishing open science practices within the research communities, and society more broadly, bridging EOSC and other sectors of the global data space.

3. Statements and call for actions

The ESFRI and pan-European RI partners in the Science Clusters share the positive experience of an effective and cooperative approach and are committed to continue their roles in ERA and to promote open science. The Science Clusters have enabled the interconnection among the partner RIs for scientific and technological objectives that would not have been possible without the EOSC initiative. The continuous commitment of the Science Clusters in EOSC and their sustainability participate in the sustainability of ESFRI and other world-class RIs.

The Science Clusters have elaborated a common vision about the evolution of the implementation of EOSC, with the "EOSC Science Cluster Nodes" providing high-level services to support open science and research excellence.
The five Science Clusters unanimously share a prospective workplan for their sustainable positioning within the EOSC Federation and actively solicit **a dedicated destination action within Horizon Europe and future framework programmes with similar scope and impact as the ones under Horizon 2020** that in 2018 allowed the uptake of cluster initiatives. A second destination action in the EC framework would be fully justified if it were to support the transition of the Science Clusters to a fully operational phase.

The five Science Clusters note:

- The procured EU Node services will be instrumental for multi-disciplinary collaboration, support for the long tail of science and the fostering of standardisation in data stewardship practices. The EU Node infrastructure and services should, however, be driven primarily by the stated needs and the continuous evolving requirements of researchers, including those of the scientific communities that are already collected by and channelled through the Science Clusters and their participating RIs.
  In this respect the Science Clusters propose to be formally involved in providing requirements and assessing the value of the procured EU Node services. The European Commission can rely on the Science Clusters to assess the EU Node services' quality and their real scientific impact.

- An EU Node should pick up and operate the commonalities that are found to be useful by the different scientific communities, while respecting their diversity and relying on the Science Clusters' cooperative programme to select and populate these "commonalities":

  i. To maximise the scientific impact, mandating the operators of the EU Node to procure "common services" should be avoided (apart from a very few obvious ones: authentication and authorisation infrastructure (AAI), access to computing, data model standardisations, etc.).

  ii. Before a component/API/guideline is included as part of the EU Node portfolio, its utility needs to be demonstrated, if a successful interoperability framework is going to be deployed.

  iii. Open science and service cascading grant calls for combining scientific domains would be very strategic in guiding what is required.

  iv. In terms of governance, the EU Node legal and central authority would be called to establish methods to assess the scientific values of its core services, the ownership of which can be moved to national entities or existing external e-infrastructures. In this respect, the inclusion of the "cluster of clusters" system within the central legal entity governing the EOSC Federation would guarantee that both the implementation and operations remain strongly connected to the user experience.

- The EOSC agenda (or more largely the agenda of Open Science in Europe) needs to concentrate on its scientific impact. It should target sooner, and with higher priority, the "EOSC science content enhancement" and the onboarding of scientific communities via the "implementation and operation of open science platforms"; for all these objectives the five Science Clusters offer to take a leading and structural role.

- A dedicated "EU Open Science programme" should leverage further cascading grant calls that will be suitable in the long run to promote cross-border and cross-domain cooperation, the uptake of open science by researchers, and the prioritisation of society's scientific needs in the EU framework. The Science Clusters are ready to drive and assume the responsibility of running such an EU programme that would also maximise the scientific impacts of their partner RIs.

- If one assumes that the EOSC EU Node, owned by the EC on behalf of the EU member states, corresponds to, or will be led by, the future EOSC legal central authority, and if such a central authority will be also in charge of leading the "EU Open Science programme", the "cluster of clusters" framework would become a "system of thematic EOSC Science Cluster Nodes" comprised within the central authority and participating to the governance of EOSC. With such a role for the Science Clusters, the remit of the central authority is extended to governing and operating:
  - ◊ the continuous evaluation and evolution of high-level services for open science;
  - ◊ new open science projects;
  - ◊ the sustainability of the data challenges of all RIs (that are within the ESFRI roadmap and other world-class ones), the onboarding in the open science ecosystem of new RIs and the participation in excellent research.

  The sustainability of EOSC and its evolution will be guaranteed since the Science Clusters would necessarily engage with the user communities and have the vocation to provide what they need to conduct their science.

Recommendations from the Science Clusters:

1) The Science Clusters call for future funding support.

   - In the expectation that the community-based competence centres that will be implemented within the first phase of the OSCARS and EVERSE projects prove their utility, the Science Clusters highlight the need for future funding support for their sustainable deployment, structuring, management and operation.

   - The highly composable VREs of the Science Clusters will move from the prototype phase to the operational phase through the hosting of the OSCARS cascading grant projects. While the Science Clusters will ensure the continuous development and maintenance of the VREs, support is required to deliver the HPC needs to service the VREs' continuous operation. A new funding action is needed to guarantee that these VREs will be sustained by the future central EOSC authority. Such support will enable the commitment of the Science Clusters to operate the reference domain-specific platforms for the "Web of FAIR Data and Services for Science" that EOSC has the ambition to be.

   - The cross-fertilisation and the cooperative actions (brought in by the RIs within each Science Cluster) need to be maintained. Funding support by the EC is solicited for that. Indeed, continuous development, upgrade and innovation is part of the scientific research process. Without long-term

support of the Science Clusters for their work programmes, there will be only limited involvement of user communities and of the partner RIs, which are the data producers and provide fundamental feedback and validation of the deployed EOSC core services.

- With the involvement of international organisations in the Science Clusters, taking into account the global participation of universities and research centres in the partner RIs, and with the aim of extending the successful EOSC approach beyond the European framework, grants to strengthen links with third countries will be strategic. Supporting the international and inclusive role of Science Clusters would benefit the global dissemination of research on FAIR data and help engage developing countries.

2) The Science Clusters call for the EC to acknowledge the fundamental role of Science Clusters and endorse their full integration in the implementation of the EOSC Federation.

3) The Science Clusters look forward to relying on, supporting and participating in the evaluation of the procured EOSC services for the EOSC EU Node, and call for alignment of the evaluation actions with national initiatives.

4) The Science Clusters call to be an integral part of the governance of the central EOSC/EU Open Science authority, and to have a key role in steering the central EU open science programme funding.

5) Strengthening and supporting RIs in ERA would require that the legal authority of EOSC supports and relies on the Science Clusters' role to incentivise an economies of scale approach for the sustainability of ESFRI and other RIs, as well as for security in data management; to maintain and develop innovative training schemes; develop a science strategy in the EU, co-elaborated with EU member states and the EC, for specific science priorities and innovative paths; structure and strengthen science diplomacy, and ways for scientific knowledge to serve public debate as well as decision-makers; linking EOSC to other sectoral data spaces by leveraging crosscutting and cluster cross-domain projects connected with these data spaces and assessing cooperative schemes with SIMPL[11] procured services.

6) The operational role of the Science Clusters has to be seen in the EOSC Federation through the implementation of five thematic 'EOSC Science Cluster Nodes', each of which will be represented by a legal entity. (Such legal entity could be also a national institute willing to host the thematic Node by consensus of cluster partners, and to contribute to strengthening the links with universities and other academic/research organisations). Such EOSC Science Cluster Nodes would then be integrated within the governance of the central legal authority of EOSC, with the specific and prominent remit to ensure the scientific scope of EOSC.
   More precisely, a future EU central authority would rely on the thematic Science Cluster Nodes to manage:
   - ◊ cascading grant calls for open science for a period spanning the transition phase, and potentially in the long run;
   - ◊ a highly composable virtual research environment per scientific domain for full hosting of open science projects;
   - ◊ the federation of data archives;
   - ◊ the cooperative schemes with EOSC national Nodes, which could be transdisciplinary, on either specific domain-based open science programmes (including curiosity-driven national access to pan-

[11] https://digital-strategy.ec.europa.eu/en/news/commission-awards-eu41-million-contract-develop-infrastructure-common-european-data-spaces

European RIs actions) or multidisciplinary projects including those bridging with industries, as well as with other sector data spaces.

7) The thematic EOSC Science Cluster Nodes will encompass their domain-based Competence Centres and together would formally commit to making EOSC the “data space for science”. This will imply a move towards “open data, software and hardware” and more broadly to open access to RIs. The system of thematic EOSC Science Cluster Nodes will lead to further initiatives such as supporting a series of outstanding international conferences and camps on “Open Research”. Furthermore, the synergies and coordination between the Science Cluster Nodes with EOSC national Nodes would be planned for the success of those missions.

8) The usefulness of the EOSC Exchange layer would be guaranteed since it would become the particular remit of the EOSC Science Cluster Nodes to assess and maintain the portfolio of high-level services through the Science Clusters’ repository of services, software, workflows, tutorials, training supports and expertise. The EU central authority is expected to ensure the quality of its components, while the five Science Clusters are ideally positioned to own the “Exchange for Science” layer and to provide that as a service to EOSC, assuming total responsibility on content policies and quality of components that the Science Clusters and EOSC would offer. This is complementary to other hardware and e-infrastructures’ Exchange horizontal services, like compute capacity.
   The ownership of the “EOSC Science Exchange” layer has already been taken over by the 'cluster of clusters' approach within OSCARS. Among other objectives, this approach is consolidating a multidisciplinary action (e.g., through EVERSE and other projects) on Artificial Intelligence (AI) and Machine Learning (ML) for the benefit of the partner RIs and universities involved. This is essential for the data mining or data visualisation techniques that are so important for all Science Clusters and strategic for bringing some relevant services into the EOSC Science Exchange layer to help address the challenge of understanding and interoperability of 'big data'. Furthermore, such a role of the Science Clusters will be highly relevant for establishing common, transdisciplinary terms and methods for the “validation” of AI software repositories at the EU level, indispensable to assess the quality of future scientific results that will be massively based on AI workflows. Coordinated and supervised AI solutions in science will then favourably impact FAIR data management, help in democratizing the scientific analysis process, support more citizen science projects demanding FAIR management of high-level science data products, and foster the usage of HPC platforms, thanks to a more in-depth training on high performance software programming for the application of AI for research and innovation objectives.

   In such a perspective, the EOSC National Nodes and other national initiatives are expected to play a role, since in many cases the Science Clusters services are produced by the collaboration among researchers and their universities and national institutes. Therefore, Science Cluster Nodes would foster the role of the EOSC National Nodes in the sustainability of some of the EOSC Exchange services. The future operations of the five thematic EOSC Science Cluster Nodes would leverage **a matrix framework** in which vertical competences from each Science Cluster domain, at the service of the concerned scientific community, and for the sustainability of partner RIs, will be interconnected through horizontal cross-fertilisation axes among the five Science Clusters in order to adopt similar solutions, such as (i) data-management services; (ii) science data-analysis platforms; (iii) common solutions for quality and impact of open source software (e.g. AI methods); etc.. Through such an approach, the five EOSC Science Cluster Nodes will commit to own and deliver cross-domain EOSC Exchange services.

4. Some specific expectations

In the following, some of the main domain-based expectations formulated by the RI participating in OSCARS are highlighted; the commonalities will be leveraged in OSCARS and consequently strengthened in the future implementation of the matrix of EU Nodes.

ENVRI

The ENVRI community is set to collaborate on the establishment of a prototype EOSC Science Cluster Node for environmental science data across Europe, featuring state-of-the-art data integration services through ENVRI-Hub. This will enable a comprehensive amalgamation of diverse environmental datasets, propelling the Node to the forefront of data harmonisation. Furthermore, ENVRI's commitment to developing domain-specific tools and services will ensure that the Node addresses the unique demands of environmental research. Alongside this, the ENVRI community will build its Competence Centre, fostering an environment where exemplary practices and cutting-edge research elevate the EOSC-ENVRI Node's capabilities. This Centre will act as a hub of innovation and expertise, contributing to the robust web of data and services within EOSC. By harnessing previous strides made in data federation and by navigating the challenges of harmonising diverse datasets within the ENVRI community, this initiative promises to enhance the data infrastructure significantly and provide a crucial blueprint for the implementation of the EOSC Science Cluster Node framework. Financial support for the development and maintenance of an EOSC ENVRI Node is crucial to facilitate the effective integration of research data and resources across Europe, enhancing scientific collaboration and discovery.

EOSC-Life

The Life-Science-RIs (LS RI) within the EOSC-Life consortium cover a broad range of life science disciplines, with a particularly disparate range of competences, experiences and requirements regarding FAIR data and open science. In the past, these differences constituted a barrier to effective cross-discipline collaboration. Participation in EOSC-Life showed the LS RI that they urgently need new methods for (meta)data harmonisation and mobilisation, for meta-analyses and for analyses across their currently isolated sub-disciplines. The establishment of cluster Nodes is one of the necessary steps towards (meta)data harmonisation and will promote exchange and aggregation services. These promise to address community needs better and more efficiently than the continued development of specialized tools, promoting open-source practices, catalogues of innovative and replicable tools, as well as digital commons for scientific workflows and FAIR Digital Objects.

They will also ensure that Semantic Interoperability (SI) methods, services and governance rules align with the framework proposed by the EOSC SI task forces. Within OSCARS, as in EOSC-Life, there will be an effort to contribute to the creation of highly portable and integrated VREs, including those adapted to the particular options and specifications required when working with interdisciplinary and sensitive data, including patient clinical data, while also integrating catalogues of domain-specific semantic artefacts. Finally, the future EOSC-Life cluster Competence Centre, providing support and training programmes and federating data stewardship through RIs and institutions will serve as a prototype for EOSC Science Cluster Node core services. This will help to build common data harmonisation and sharing methods and increase the potential for the heterogeneous data generated by the EOSC-Life community to be re-used creatively, efficiently and productively. Ultimately, these actions will help drive innovation in Europe.

ESCAPE

The ESCAPE partner RIs together with EU member states' domain-based national institutes will implement the links with data repositories from different RIs' specific platforms. ESCAPE will operate the community-based competence centre for: data management through the data lake infrastructure; deploying practices and methods for open-source catalogue of innovative software for scientific workflows; the high-level services for composable RI-specific analysis platforms, as well as services and an infrastructure for integrated VREs for data interoperability and reusability.

The ESCAPE community will assemble, within a domain-based EOSC Science Cluster Node, the competence centre, the coordinating hub of the virtual research institute for scientific software, the software catalogue, an initiative for the large collaborative development and sharing of AI and HPC workflows, its domain-based VRE, as well as the operational links to existing infrastructures, such as the Astronomical Virtual Observatory, the CERN Open Data portal, the ESO archived data, dedicated Citizen Science platforms, such as Zooniverse, and will aim also at bridging with ESA.

The technological capacity of the ESCAPE community as well as its ICT infrastructures are also made available to all the partner RIs and the concerned scientific community, and will contribute to the evolution of the EOSC EU Node core services.

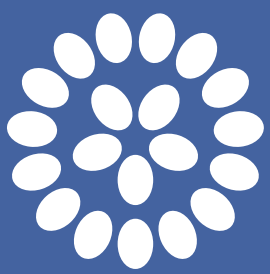
PANOSC

The PaNOSC community will work together to implement a prototype EOSC Node for data from the Photon and Neutron sources in Europe. The node will create links with the other science cluster data repositories and scientific data repositories e.g., the Protein Data Bank, Human Organ Atlas, to contribute to building the web of data as part of the core of EOSC. The PaNOSC community will contribute to the EOSC Handbook of operations, as a use case for other national and European facilities and communities who want to create an EOSC node. PaNOSC will build on the work started during the PaNOSC H2020 project for federating data. The challenges of federating heterogenous data produced by the PaNOSC community is similar to the one faced by federating the science clusters and will serve as a test bed for the EOSC Science Cluster Node concept.

SSHOC

SSHOC aims to transform the data landscape in the social sciences and humanities by creating an integrated, cloud-based network of interconnected data infrastructures that facilitate both data-based and data-driven research, and by providing community support through a thematic competence centre.

The SSHOC community will focus on improving access to SSH data and tools, workflows and training. Five priorities structure the work programme of SSHOC: (i) Federation of heterogeneous data, supporting researchers' needs and the crossing of boundaries of scientific domains and languages, by exploring ways of maintaining standards (for building catalogues and enabling data access), and furthering the interoperability of metadata and data, including AI technologies, (ii) Access models for sensitive data and mature tools and services for potential onboarding to the EU Node, (iii) Advancing domain-specific metadata descriptions and standards to enhance interoperability through the common SSH Open Marketplace, (iv) Multilingual user interfaces across the SSHOC service portfolio and (v) Sustainability and long-term preservation measures to keep data FAIR over time.



SCIENCE-CLUSTERS.eu

Research Infrastructures for Open Science