



ESFRI SCIENCE CLUSTERS

POSITION STATEMENT

ON EXPECTATIONS AND LONG-TERM
COMMITMENT IN OPEN SCIENCE

JUNE 2021



ENVRI-FAIR - Andreas Petzold (Forschungszentrum Jülich), Ari Asmi (University of Helsinki) and Magdalena Brus (ICOS ERIC)

EOSC-Life - Niklas Blomberg (ELIXIR) and Michael Raess (INFRAFRONTIER)

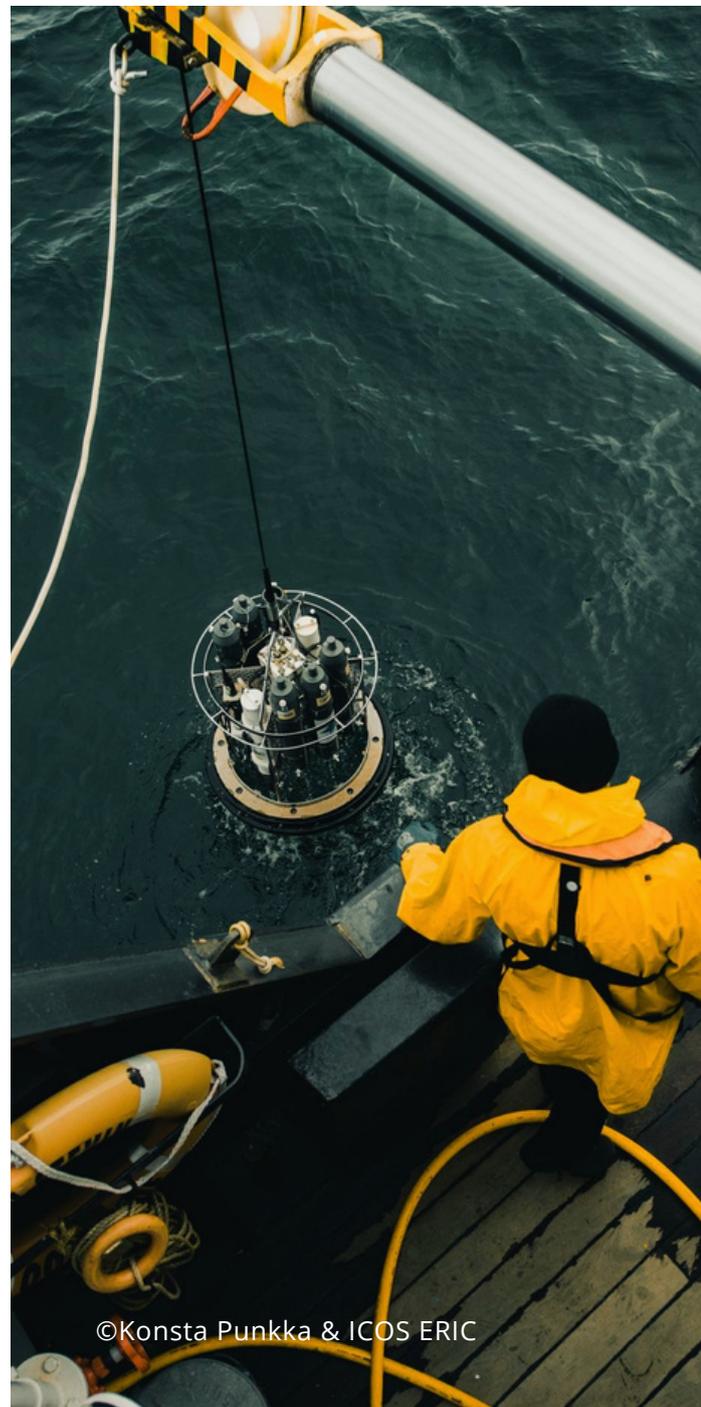
ESCAPE - Giovanni Lamanna (CNRS-IN2P3/LAPP) and Ian Bird (CNRS-IN2P3/LAPP)

PaNOSC - Rudolf Dimper (ESRF) and Andrew Götz (ESRF)

SSHOC - Ron Dekker (CESSDA)

| TABLE OF CONTENTS

- 03** Executive Summary
- 04** Science Clusters in the European research and innovation landscape
- 07** Common Statements
- 08** A work plan for the future of the Science Clusters
- 11** A domain-based Science Cluster analysis
- 12** ENVRI-FAIR
- 15** EOSC-Life
- 18** ESCAPE
- 21** PANOSC and ExPaNDS
- 23** SSHOC



THE SCIENCE CLUSTERS

are EU collaborative projects that were launched in 2019 to link ESFRI and other world-class Research Infrastructures (RIs) to the European Open Science Cloud (EOSC).

The main impacts of the Science Clusters' work programme concern:

- the improved access of researchers to data, tools and resources, leading to new insights and innovation for data-driven science both within and beyond the context of the domains in which the clusters are rooted;
- the creation of a cross-border open innovation environment for FAIR (Findable, Accessible, Interoperable and Reusable) data management for economies of scale, to develop synergies and raise the efficiency and productivity of researchers through open-science standards and thematic services;
- the enhanced co-developments to foster the cross-domain interoperability central to the EOSC goal.

The Science Clusters are an integral part of EOSC. Their services and outcomes are now forming the core of the emerging EOSC fabric. As important partners of EOSC, Science Clusters contribute to its development and its implementation process. Importantly, the Science Clusters form a natural collaboration between the ESFRI RIs' management boards partners in the clusters.

As EOSC matures and begins delivering data and services for European research, a discussion is needed to stimulate the Open Science practices, cross-domain interoperability and long-term coordination of the scientific communities covered by the five Science Clusters.

This position paper contributes formally to explain the urgent need of EC to support a longer-term role of the five Science Clusters to provide content to the EOSC, to enhance researchers' involvement in Open Science and to suggest potential cooperative pathways in the Horizon-Europe framework and along with the EOSC Association roadmap.

This is a second position statement document, written keeping in mind the various exchanges the Science Clusters had with the European Commission, the EOSC Association, the Directorates of the ESFRI RI partners, and based on the constant cooperative work between the management boards of the five Science Cluster Projects in the last two years.

This paper is aimed at highlighting:

- Expectations of the Science Clusters and the concerned research communities, pointing out a common structured vision and a series of suggestions for the future;
- The Science Clusters commitment in supporting the EOSC implementation and sustainability;
- A more detailed analysis from each Science Cluster that is provided for completeness.

EUROPEAN RESEARCH COMMUNITIES ARE ON THE CUTTING EDGE OF MODERN SCIENCE

and a major factor of this success comes from the coordination and collaboration in building and operating Research Infrastructures (RIs) on a European level. RIs need to stay focused on the needs of their user communities and the services they have to provide. At the same time, there is an increased societal, economic and scientific drive towards multidisciplinary, big-data analytics, artificial intelligence and machine learning, and automated virtual research platforms. All this requires far more abstract and general solutions and harmonization across traditional disciplinary borders. The European Open Science Cloud (EOSC) initiative is a strong basis for such developments. However, EOSC is by necessity providing generic services, and built on relatively abstract concepts, providing services that often need to be refined with data, enriched with additional information and tailored for scientific use. Much of these service refinements can and will be done by the ESFRI RIs, but to ensure the economy, efficiency and multidisciplinary goals of doing so, this activity needs coordination and support. This role is currently covered by the Science Clusters.

We consider that the Science Cluster actions are a powerful enabler of European research excellence and competitiveness, and is one of the EC's most effective and innovative network tools. It underlines the international dimension of disciplinary and interdisciplinary dynamics and progress. The ESFRI and pan-European RI partners in the Science Clusters share the positive experience of their effective cooperative approach and are willing to pursue this successful experience in the long term.



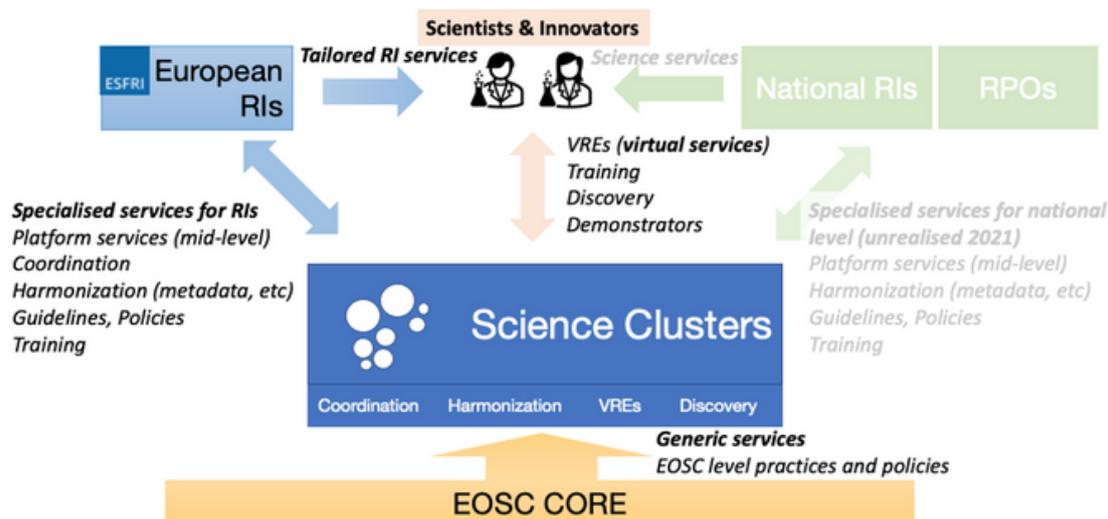


Figure 1 - Simplified diagram of Science Clusters as refining and transforming the EOSC Core services towards specific services and platforms for RIs, scientists and (potentially) EU member state institutions.

Position I: General Role and positioning of Science Clusters

The Science Clusters are engines for Interdisciplinarity. The five Science Clusters are implementing a cross-fertilization environment for the concerned scientific communities and this collaboration has effectively bridged sociological and technical barriers and begun strongly influencing a pan-European and cross-disciplinary approach that will drive scientific excellence and competitiveness. Thus, the Science Clusters contribute to shaping the future EOSC agenda, the European data spaces[1] foreseen in the European strategy for data, the Horizon-Europe framework evolution and in operationalising the European Research Area (ERA).

The Science Clusters occupy a unique position between EOSC, ESFRI RIs and scientific communities. The Science Clusters are the infrastructure platforms engaging ESFRI and other RIs for co-development, co-creation and co-operation actions to the benefit of and supported by large transdisciplinary scientific communities (Fig. 1). They also work in collaboration with outstanding forums (ex. EIROforum, ERIC forum and ESFRI) that bring together some world-class RIs for science and technology policy debates.

Position II: Science Clusters in EOSC and Open Science landscape

The Science Clusters have demonstrated to be uniquely capable to address the challenges of open research data and data federation by being at the same time data producers and consumers. The specific focus of the Science Cluster projects around Data-FAIRness and Open-Science challenges has been key for their success, due to the critical need of these services by the participating RIs to establish data management practices that enable interoperability and underpin scientific reproducibility. The cross-fertilization approach of the Science Clusters has shown its usefulness and efficiency for developing consolidated services for RIs and scientists as well as to address new emerging focus in open data science (namely within the European Data Spaces).

The Science Clusters are a key part in developing mid-level multidisciplinary tools and platforms, enabling specific scientist-centred services over generic ones. The implementation of Cluster, RI, or even user-created virtual research environments (VREs) are fundamental in order to enable directly usable scientific platforms. Cluster services can also improve the FAIR data-research practice as well as define and operate key

enablers such as, rewards and incentives, skills and education, society and economy widening and dedicated infrastructures.

The Science Clusters' track record of integrating infrastructure can be a significant asset for the EOSC. The recent establishment of the EOSC Association (EOSC|A) is a milestone towards the implementation of the EOSC. The fact that the majority of the EOSC|A members are Universities and Research Institutes, reveals from one side indisputably that science is intended as the major aim of EOSC and from another side that the European Union member states are willing to commit in the operational phase of EOSC in support of research communities. For the Science Clusters, it is of primary importance that the EOSC implementation roadmap, as well as its governance, make constant reference to domain-based scientific communities and to the clusters of (ESFRI) RIs supported by them. The Science Clusters work programmes leverage role and participation of EU Universities and Research Institutes.

Position III: Science Clusters in scientific and RI communities

The Science Clusters act as the key interfaces between the scientific communities, their infrastructures, and the EOSC. There isn't currently any other major framework as relevant as the Science Clusters to guarantee the links between scientists and multiple research organizations together within the EOSC ecosystem. The clusters and their member RIs are positioned to understand their community needs. Via the clusters the participating RIs can project the community needs and the disciplinary conventions, methodological frameworks, workflows into the EOSC. Conversely, the clusters naturally provide a scalable mechanism for the EOSC to engage and drive an Open Science culture by incentivizing and fostering collaborations between RIs to share knowledge

and data as early and as openly as possible.

The Science Clusters provide vital links to the community and community governance. The five Science Clusters believe that the ultimate EOSC vision of an infrastructure of "Web of FAIR Data and Services for Science" will be incomplete and ineffective without "community-governed" open-science commons platforms co-developed and operated by scientists, inspired by shared core values and that meet and work together in a virtual global environment to produce open research.

The Science Clusters build and maintain key community-centred initiatives. The five Science Clusters have recently committed to co-steer shared actions to engage more researchers in the EOSC and in cooperation with pan-European e-infrastructure organizations in the H2020 EOSC-Future project. In this respect the Science Clusters are supporting consolidation actions within their current work programmes, are building multi-domain Science Projects (SPs) to demonstrate the scientific inclusiveness of the EOSC, and are engaging in making the EOSC a federation VREs for European Researchers.

The Science Clusters also provide a rapid and efficient platform for joining the RI services within or even cross clusters for important societal challenges. Recent COVID-19 projects demonstrate the applicability of Cluster level collaboration for extremely rapid response to sudden demands. Existing collaboration frameworks, not only on a technical but also practical level, create significant capability for science-based societal and economic resilience in Europe.

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

The five Science Clusters formulate hereafter the following shared statements:

The Science Clusters are willing to operate and adapt their thematic open-science resources in EOSC towards four urgent needs:

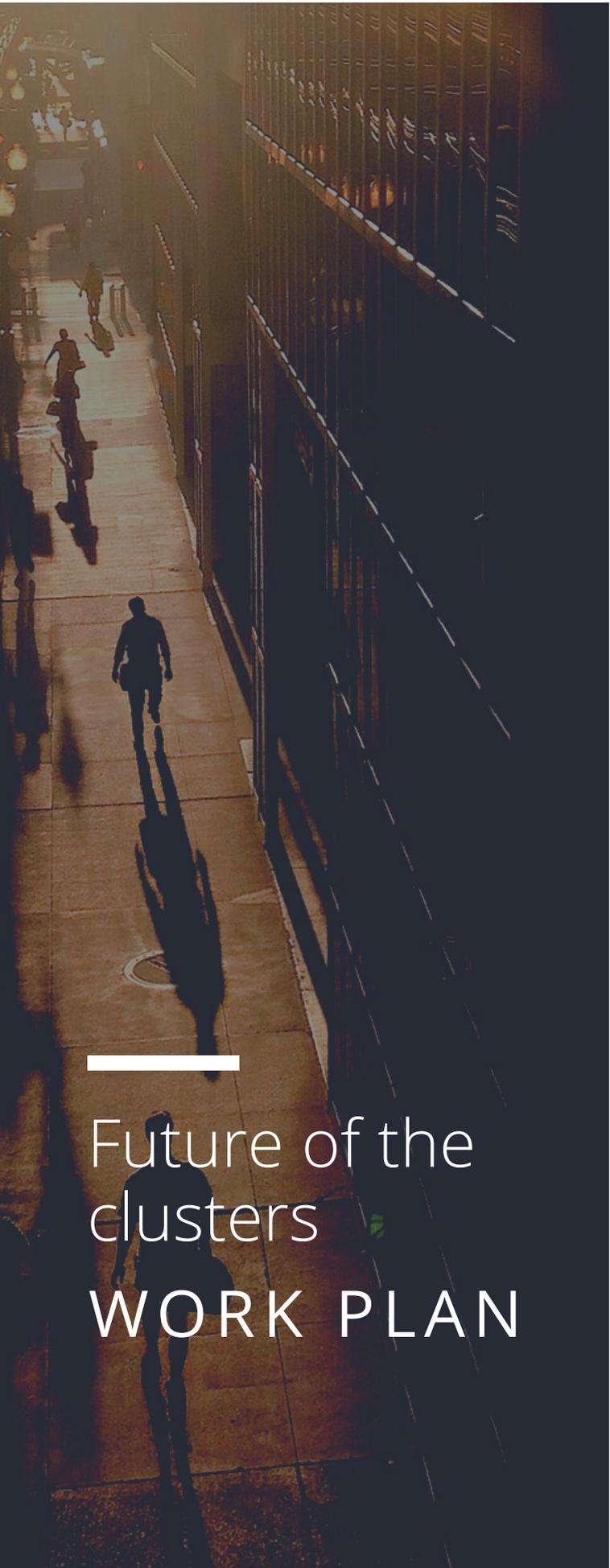
- (I) making science always central, solid and inclusive of its social and cultural dimension, as collective knowledge to be nurtured;
- (ii) leveraging crosscutting and cluster cross-domain projects connected with European sectoral data spaces;
- (iii) providing training, committing in education and engaging citizens in science;
- (iv) exploring the way to translate their excellence in data-research into economic value.

In order to fulfill such a mission:

- 1) The ESFRI and other RIs participating in the five Science Clusters, call for policy and actions to drive interdisciplinarity and maintain a strong link with scientific communities for the next phase of EOSC, namely:
 - (i) establishing a platform with the EOSC Association to strategically develop the existing RI data resources, standards and computational services as part of EOSC.
 - (ii) Horizon Europe actions to support the Science Clusters in their role of “developers” as well as “aggregators” of data services and community-based projects for science and innovation, as well as supporting wide-level scientific culture harmonization in their respective fields.
- 2) The European Open Science implementation agenda should target sooner and higher in priority the “EOSC science content enhancement” and the “science community cooperative environment implementation and operation”. For both the five Science Clusters offer to take a leading role, enhancing the open-science research and tuning the required services.
- 3) The EOSC compliance framework, as one of the core functions of EOSC|A, needs urgently to be addressed. The rules for participation as well as e-infrastructures provision are pending to allow the Science Clusters to operate the FAIR data-research for the users of their large number of ESFRI and other RIs.

Common STATEMENTS





The five Science Clusters share a prospective pathway towards their sustainable establishment in the EOSC ecosystem and expect to leverage the Horizon Europe framework for its implementation.

CLUSTERS AS PLATFORMS FOR SCIENTIFIC INTEROPERABILITY IN EOSC

The five Science Clusters within the EOSC will continue to evolve, driven by the scientific need for intra-domain interoperability and alignment.

While there will be organisational diversity between the clusters, building on the established federations of domain-based (ESFRI or other world-class) RIs there is a common need within each cluster for a domain-based 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 VREs. This will ensure that the individual and collective RI data resources, computational services and VREs are exposed, connected and operated within the EOSC.
- The Science Clusters, via partner RIs, engage with the scientific communities at large and, where applicable, existing consortia/groups, to

Future of the
clusters
WORK PLAN

develop specialized data services.

- The Science Clusters will engage with RIs, European Commission and EOSC|A to define the way to sustain the developed cluster services, to establish the sustainability scheme for their operation and lead (in the respect of the positioning of individual national institutes) the expected key operative roles within each domain-based resources and VREs.
- All (five) platforms will pursue the high-level cooperative scheme as well as the cross-disciplinary co-developments to explicitly foster the cross-domain interoperability central to the EOSC goal.
- The platforms, internally and collectively, will study, define and set up a series of new task forces to respond to new or emerging challenges, after consultation with the European Commission. They will also structure their commitment to Horizon Europe to take part in the European Data Strategy at large. Cross-cluster initiatives are expected to emerge easily leveraging the platforms. The European Commission is expected to acknowledge and support the “cluster-platforms” without which it will be difficult to leverage their pivotal roles in achieving the EOSC ambitions.

SUPPORTING THE “EOSC ASSOCIATION MANDATE” AND THE EUROPEAN COMMISSION

The Open Science endeavor in Europe is the original collective goal of the Science Clusters and the current focus of their respective H2020 projects. The EOSC Association mandate is supported by the Science Clusters. Clusters’ partners are meanwhile represented individually in different bodies of the EOSC governance (and in particular as members in the EOSC|A). The Science Clusters would bring in addition the collective domain-based scientific community expectations and potential.

Science Clusters should be acknowledged in their role as domain-based “open science commons enablers”. Therefore, the five Science Clusters consider it imperative to be linked directly to the EOSC Association Board of Directors, supporting it in relevant surveys and Science and Technology pre-analysis that should influence the evolution of the EOSC roadmap. This will also help in a stronger involvement of the EOSC|A General Assembly and build a solid consensus with national stakeholders.

It is expected that the Horizon Europe framework programme will be able to support the Science Cluster-platforms roadmap, specifically in their roles as “enablers, operators and aggregators” of open data services/commons for science and innovation. The Science Clusters are ideal for fostering Open Science, because their members are naturally concerned and extremely active in all its different declinations: (i) Incentives & Rewards; (ii) Research Integrity; (iii) Research Outputs; (iv) Education & Skills; (v) Scholarly Communication; (vi) EOSC implementation; (vii) FAIRness of data; (viii) Innovation; (ix) Citizen Science; (x) Society at large. Furthermore, the Science Clusters through their services and thanks to the support of their partner (ESFRI) RIs, can easily pool further data and services in a wider dimension and coordinate clusters’ co-participation to new projects addressing European sectoral data spaces. For such reasons, the Science Clusters seek a direct

link with the European Commission, in the perspective of engaging with future funding schemes in Horizon Europe.

PROMPT NEED FOR A NEW CLUSTER WORK PROGRAMME

The five Science Clusters aim to address the Open Science challenges shared by their partners and the concerned community. These challenges are technical, operational, sociological and scientific. The indisputable success of the Cluster experience and the encouraging results achieved mid-way through the current projects call for some next prompt actions along two main “Pillars”.

Pillar A - Inter-cluster common data services co-developments

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 research communities/themes challenges. It is of high importance that the next funding opportunities in the Horizon Europe RI Work Programme cover this expectation and leverage the cluster roles in Europe by envisaging inter-cluster projects. The sustainability of the five clusters is built through the structuring of their wide cooperative virtual research environments (VRE). The five Science Clusters open-science environments are not intended as a pathfinder to EOSC (cf. Candela et al, 2013 DOI: 10.2481/dsj.GRDI-013), but as the implementation goal for an effective and

successful operation of EOSC. Support for a prompt organisation, consolidation and a sustainable operating approach is considered by the Science Clusters as one of the most relevant next steps in their agendas. The implementation and operation of the VREs means also “the cooperative framework and the validation bench” through which Science Clusters commit to connecting EOSC content with technical aspects together with the pan-European e-infrastructures.

Pillar B - Delivering Content to EOSC

The five Science Clusters, through the construction of their own VREs, and the Science Projects in the EOSC-Future project, will be ready for concrete “Open Science Objectives”. They will be part of the current and future Horizon Europe Data Spaces, R&I pillars and established priorities. Dual approaches of thematic-based Open Science Projects (OSP) and Cross-Cluster Open Science Projects (COSP) are envisaged, enabling, for example, the Industrial and Manufacturing or the Green Deal to be tackled by focusing on topical aspects within one cluster supported by complementary contributions from all clusters. The five Science Clusters rely on the maturity of the communities as well as on the engagement of the ESFRI RIs to adopt and support this perspective. The added value of Pillar B lays in the capacity of the ESFRI RIs partners to leverage:

- The wide innovation potential with industry (including SMEs);
- The affirmation and extension of a prominent European role more internationally;
- The capacity of structuring education and training-by-doing offers to facilitate work insertion of young generations as well as to participate in the European resilience plan;
- The role in supporting Science & Technology Policy for Society as well as data/cloud business capacities (e.g. by engaging with GAIA-X);
- The facilitation of leveraging any potential OSP and COSP for regional developments.

A DOMAIN-BASED CLUSTER ANALYSIS

In ‘Solutions for a sustainable EOSC: a FAIR lady report’[2], it is stated that the resources and thematic services identified as ready for integration by the Research Infrastructure clusters should be prioritized for onboarding (in EOSC) because they represent high-quality resources valued by research communities. It also states that the Research Infrastructures have already established communication channels, and it mentions that at the cluster level best practices are being developed to ensure the interoperability

of federated services.

This aligns with what is being claimed here, namely that the widespread adoption of EOSC by research communities will clearly depend on the success of the work currently in progress in the five Science Clusters, their achievements and their longer-term valuable application.

In the following, each cluster details the result of its own analysis about the “sustainability” of its level of engagement and cross-domain impact for the uptake of open-science policies within EOSC.

[2] <https://op.europa.eu/en/publication-detail/-/publication/581d82a4-2ed6-11eb-b27b-01aa75ed71a1/language-en/format-PDF/source-175468053> (page 13)



A domain-based
CLUSTER ANALYSIS

ENVRI-FAIR



Background

The cluster of the European Environmental Research Infrastructures (ENVRI Cluster; see ENVRI.eu) covers atmospheric, marine, solid earth, ecosystems and biodiversity research, properly structured and organized to support research urgently needed for achieving the Green Deal and in support of reaching the Sustainable Development Goals. The majority of the more than 20 RIs gathered in the ENVRI Community were established in the ESFRI framework. The participating RIs provide data, research products and services from key areas of the Earth system. The data, products and services provided by the ENVRI Cluster are crucial European contributions to the integrated global observation system for monitoring the state of the Earth system and climate. **The ENVRI Cluster represents the core component of the European environmental research infrastructure landscape with the ENVRI community as their common forum for collaboration and co-creation.**

The ENVRI community, however, represents not only mature RIs on the ESFRI level but includes also those RIs in the environmental sector which are evolving and maturing, and includes furthermore national and regional projects and infrastructures. In that respect, the ENVRI community is developing towards an integral representation of Earth system science in Europe, even beyond the ESFRI level.

The EU-funded science cluster project ENVRI-FAIR builds on the mature and operational

Environmental Research Infrastructures that include principal European producers and providers of biodiversity and environment-related research data and services. Leading RIs are ESFRI Landmarks (EISCAT-3D, EMSO ERIC, EPOS ERIC, EURO-ARGO ERIC, IAGOS AISBL, ICOS ERIC, LifeWatch ERIC) and ESFRI Projects from the Environment domain (ACTRIS, DANUBIUS-RI, DiSSCo, eLTER, SIOS) or the Health & Food domain (AnaEE).

ENVRI-FAIR targets the development and implementation of both technical frameworks and policy solutions that make subdomain boundaries irrelevant for environmental scientists and prepare Earth system science for the new Open Science paradigm. ENVRI-FAIR will ultimately create the ENVRI-hub as the platform of the environmental sciences community in EOSC and for delivering environmental data and services through the EOSC. The backbone of the project are the operational ICT infrastructures of the participating RIs. Technical solutions, software, and services developed in the framework of ENVRI-FAIR will be transferred to the maturing RIs in the environmental sector to ensure coherent RI evolution and implementation and to reduce their implementation costs.

The architecture and functionalities of the ENVRI-hub are driven by the applications, use cases and user needs, and will be based on three main pillars: (1) the ENVRI Knowledge Base as a Wiki-based resource for knowledge, services and assets; (2) the ENVRI Catalogue as a machine-actionable interface to the ENVRI ecosystem; and (3) subdomain and cross-domain use cases as demonstrators for the capabilities of service provision among ENVRIs and across Science Clusters.

The ENVRI community as the cluster of environmental RIs sees many pathways for significant contributions to the objectives of the Biodiversity Strategy 2030, the overall European Green Deal and Horizon Europe in general. The

ENVRI mission builds on the understanding that societal impact is a key feature of environmental research. Given the climate emergency and many other urgent environmental challenges, the ENVRI community supports a clearer focus on societal impact in future research funding. As an outcome of the completed ENVRIplus cluster project, the ENVRI Community has already shown the clear alignment of the ENVRI with major grand challenges. The Community is now ready to deliver significant impact to the European societies, e.g. on climate change mitigation and adaptation, promoting food security, halting biological diversity loss, and promoting ecosystem services at land and in oceans, security in the face of natural hazards, or air quality.

The ENVRI Community is convinced that research based on the ENVRI resources, products and services can lead to guidance and support for policymakers as a basis for the European Green Deal. The European Commission strives for Europe to be the world's first climate-neutral continent by 2050 and presented the European Green Deal to achieve this. The ENVRI Community fully supports this ambition and is observing the environment, generating data and tools that can guide, enable, and monitor our progress towards this. The ENVRI portfolio of services and expertise, with research on greenhouse gases and all other components of radiative forcing, on biodiversity and ecosystem integrity at land, in freshwater, coastal systems and in the oceans, as well as on air and water quality, ranging from the Mediterranean Sea to the Arctic, is providing an excellent base for cutting-edge scientific support of all aspects of European science.

The ENVRI Community is closely related to four out of the five European research and innovation missions that will be part of Horizon Europe. In line with ESFRI, the environmental RIs are seen as essential ERA building blocks to accomplish the objectives being at the heart of the missions and

to support the research around them.

The ENVRI community is ready to contribute and deliver solutions to the respective grand challenges, namely adaptation to climate change including societal transformation, healthy oceans, seas, coastal and inland waters, climate-neutral and smart cities, and soil health and food. Together, the Community can provide an infrastructure worth more than one billion euros of capital investment ready to support these missions.

The ENVRI cluster and the ENVRI-hub as its platform in the EOSC rely on the sustainability of the participating RIs as well as on the sustainability of the ENVRI Community as the Science Cluster of the Environmental Domain.

Collaboration Mechanisms

The field of environmental sciences is large. It covers highly diverse research domains and is related to a complex fabric of environmental challenges in societies. The ENVRI community reflects this complex landscape with many RIs being, by their nature, cross-, multi- or interdisciplinary, and some being specialized in one domain. Some are dedicated to observations, while others perform experiments, and there are those RIs that are mainly data infrastructures. Some are established RIs that have already achieved a high internal organization and have a legal structure, while others are still projects and networks that are undergoing the process towards a mature RI.

The ENVRI Cluster has established the Board of European Environmental Research Infrastructures (BEERI) as its governing body to account for the complex landscape of environmental sciences. BEERI was established during the cluster project ENVRIplus and is currently maintained as part of ENVRI-FAIR. Today, BEERI is available as a body for consultation and science-based input to policymaking in the fields named above. The ENVRI Community and its governing body

BEERI, are currently exploring additional models of access that would enable scientists to use RIs as platforms for mission-related research. The body also discusses potential pathways towards sustainable collaboration among the participating RIs and the further development of the ENVRI-hub as its platform in EOSC beyond the duration of the cluster project ENVRI-FAIR.

The baseline of potential collaboration mechanisms discussed within BEERI is the sustainable continuation of independent research infrastructures, which have developed and will further develop and expose research services and tools in their respective scientific fields. The environmental research infrastructures organised in the ENVRI community prepare a Memorandum of Understanding to foster cooperation, joint activities, exchange of experience and mutual support in its framework and encourage cross-disciplinary research activities as well as the continuing operation of the ENVRI-hub.

BEERI also recommended individual research infrastructure legal entities should become members of the EOSC Association, while it is not intended to constitute a legal entity representing the ENVRI Cluster. Nevertheless, the ENVRI Cluster is seen as the core contributor to EOSC from the environmental sciences. In that respect, the ENVRI Cluster will be the main actor for cross-cluster collaboration and the development of truly interdisciplinary research in Europe.

Given the current level of organisation, a continuing funding scheme for further collaboration of the Science Clusters within Horizon Europe is an indispensable prerequisite for the sustainability of the ENVRI Cluster cooperation, the further development of the ENVRI-hub and the consolidation and further development of cross-cluster interdisciplinary scientific activities.

Aspects of Collaborative Actions

The ENVRI Community Members engage in undertaking joint activities towards the following

objectives:

- supporting each other by sharing experience and good practices in an open spirit;
- elaborating and promoting a common strategic vision;
- increasing the communication and the visibility of environmental Research Infrastructures, their services, their role, and their needs to diverse groups of stakeholders, in Europe and globally;
- providing training and building capacity of the RI staff;
- coordinating the response of the environmental RIs to urgent matters, like the recent impact of COVID-19 restrictions on climate monitoring activities;
- coordinating the RI participation in the R&D calls to ensure good coverage of the necessary services;
- engaging in ENVRI-related projects and maintaining the results obtained in previous projects.

Beyond the collaborative actions targeting the coherent evolution of the community of environmental sciences in Europe, the ENVRI Cluster will share the above-listed experiences, technologies, services and training resources with the entire group of Science Clusters.

Notably, the following resources and services will be shared and further developed with the Science Clusters in Europe:

- The ENVRI Training Platform and resources published there;
- The ENVRI Knowledge Base and resources published there;
- FAIRness assessment tools developed in collaboration with the Go FAIR Foundation, e.g., the analysis and use of FAIR Implementation Profiles;
- The Dashboard on the State of the Environment developed as a link to the SSHOC Science Project on Climate Neutral and Smart Cities;
- The ENVRI-FAIR Science Project on Biodiversity in a Changing Climate in close collaboration

with EOSC-Life.

- Cross-cluster cooperation has the potential to facilitate the continuation of activities that were initiated by different projects, such as RI-VIS, ENRIITC, etc.

EOSC-LIFE



Background

EOSC-Life (eosc-life.eu) brings together the 13 Life Science Research Infrastructures (LS RIs) within ESFRI in a project that will create EOSC for the life sciences, an open collaborative space for digital life science. The project co-creates and integrates the EOSC federated core, while simultaneously creating, adapting and adopting the services and policies for Open Science that helps researchers in the life sciences to manage, publish, analyse and reuse data.

The LS RI provide access to advanced instruments and research facilities - helping researchers to describe biology from single molecules to ecosystems and long-term population cohorts. The LS RI are all distributed organisations, each bringing together national facilities and centres into a connected European entity with harmonised access, quality and data management. EOSC-Life builds on the outcomes from previous cluster projects: in CORBEL the LS RIs established a foundation of collaborative scientific services that support users throughout the execution of a scientific project: from planning and grant applications through to the long-term sustainable management and exploitation of research data. EOSC-Life takes the next step and further

harmonise user access, unify data management for biomedical research in Europe and will embed the combined infrastructure capabilities into the scientific workflow of advanced users.

EOSC-Life prepares European life science for a new way of working

By populating EOSC-Life, we are not merely fulfilling a project's goals—we are building tools and training people to make this the new normal for life science data. European scientists should be able to collaborate and reuse data regardless of where they are based. EOSC-Life provides the seed funding for a cohort of "RI Data Experts" - developing data expertise within the Life Science Research Infrastructures, helping to build a community of experts across the European research infrastructures. Training programmes, workshops and hackathons help to prepare our users for a new way of working and is supporting RIs in transferring their face-to-face training to a remote setting by organising a remote learning series such as the online tutorial on ResOps, cloud native tools and technology for researchers.

EOSC-Life provides access and support to cloud resources for the life science community and offers deployable platforms for workflows. In addition, we establish capabilities and identify suitable cloud providers for secure hosting of sensitive data and have a helpdesk system for project partners to provide support for technologies within the project (e.g. cloud deployment, Galaxy). We are creating an access and user management system to enable multi-RI applications and workflows that build on existing approaches and support access to sensitive data with their specific requirements. **The Life Science Login** (LS AAI) enables researchers to use their home organisation credentials or community or commercial identities (e.g. ORCID, LinkedIn) to sign in and access data and services they need across multiple platforms.

Toolbox for sharing of sensitive data. Sharing sensitive data is a specific challenge within EOSC-

Life. For that reason, a toolbox is currently under development, providing pooled information on recommendations, best practices, software tools etc. to researchers who wish to share and/or use sensitive data in a cloud environment in general, and the European Open Science Cloud in particular. The sensitivity of the data may arise from its personal nature but can also be related to intellectual property considerations, biohazard concerns, or the Nagoya Protocol.

Landscape mapping on sharing and re-use of health data. EOSC-Life is currently mapping the national landscapes on sharing and re-using health data to understand the implications for Life Science RIs and their services, especially with regards to data protection, appropriate safeguards (e.g. de-identification techniques), risk-based approaches, data ownership, and conditions for sharing and re-use of health data. The results will allow us to assess EU countries' preparedness for EOSC.

EOSC-Life populate the EOSC with life science data

EOSC Cloud data deployment and FAIRsharing.org EOSC-Life Collection now contains over 100 diverse data resources (each containing thousands of datasets) in FAIRsharing.org that follow FAIR principles. Provenance standard for life science data is being developed under ISO 23494 to describe the history of data in life sciences in distributed environments, in order to assess reusability of data for further research and to improve reproducibility of research results. The standard also supports compliance with the Nagoya Protocol for equitable benefit sharing.

Clinical Research Metadata Repository and the COVID-19 Clinical Trial Data Repository.

EOSC-Life partners have been working on the development of the Clinical Research Metadata Repository (MDR), including COVID-19 data, allowing researchers to access clinical studies and related data objects. These include, for example, protocols, information sheets and consent forms,

data management plans, statistical analysis plans, case report forms, results, publications, descriptive metadata, etc. MDR contributes to making clinical research data from all disease areas FAIR by increasing data Findability. Browse and search metadata on clinical trials—as well as all related documents. The COVID-19 Clinical Trial Data repository is part of the European COVID-19 Data Portal.

Extension of the COVID-19 Data Portal. The COVID-19 Data Portal brings together relevant datasets submitted to major centres for biomedical data, with the aim to facilitate data sharing and analysis, and to accelerate coronavirus research. Within EOSC-Life the COVID-19 Data Portal is extended to mobilise open biomolecular data (500,000 records from the biomolecular and literature domains are available openly to users), to mobilise new SARS-CoV-2 data (currently >160,000 viral isolates with raw sequence data; 75% of the world's data flows through the Data Hubs) and to connect to clinical and epidemiological data.

EOSC-Life package tools and workflows for use in the cloud

EOSC-Life is developing expertise in the cloud deployment of software and workflows across all domains of the life sciences. This will help Life Science Research Infrastructures to develop their computational infrastructure and make it FAIR.

WorkflowHub a registry of scientific workflows to make them more findable. An open community of developers and users has been formed and is growing, as the WorkflowHub Club. During the COVID-19 pandemic EOSC-life provide tools and workflows to tackle analyses of COVID-related data in an open and reproducible way. In collaboration with the global Galaxy community, a registry for COVID-19 workflows have been made available on public Galaxy instances worldwide. **RO-Crate** is a community effort to formalise packaging of research data with structured metadata, based on Schema.org. EOSC-Life has been instrumental in developing the RO-Crate community and

specifications, in particular with Workflow RO-Crate and we are also aligning RO-Crate with FAIR Digital Objects through Research Data Alliance and the DISSCo SYNTHESIS+ project.

Collaboration Mechanisms

The Life Science RIs are the foundation for sustainability of EOSC-Life services, the life science infrastructures on the ESFRI roadmap are established entities and form the long-term actors in the life science ecosystem with strong anchoring in the user communities and national facilities.

In 2015 the LS RI established the LS RI Strategy Board, governed by an MoU between the individual infrastructures, as a framework for long-term strategic collaboration and joint work to "serve excellent science in Europe by provide access to world-class facilities, samples, instruments, services and data". Through the LS RI Strategy Board the infrastructure directors meet at least four times per year and have regular "retreats" aligned with the EOSC-Life project and work-package leaders.

In addition, the joint services provision between the LS RI is supported by pairwise MoU that build on the build on the outcomes from CORBEL project (2015-2020). The purpose of these MoU is to sustain and improve the procedures for facilities access established in CORBEL and serve as a basis for joint access projects aligned with Horizon Europe priorities and missions.

The experience from the rapid response to support COVID research and manage cross-disciplinary data flows during the pandemic shows the long-term value of maintaining close collaborations active interfaces between the RI. Societal challenges such as pandemics, climate change and food security links human biology to developments in environments, ecosystems and agriculture. In addition, life sciences are rapidly

developing - every year sees the invention of new technologies, instruments and models - responding to the changing user needs need ongoing investments to ascertain support to advanced European projects.

Aspects of Collaborative Actions

Building on EOSC-Life and the developing European landscape future challenges and collaborative actions should address

- Continuously populating EOSC with data through widespread FAIR, cloud-based publishing and access. In EOSC-Life we have to date catalogued over 100 high value research data sources and established new services to address critical gaps (e.g. clinical trial data). This has only skimmed the surface - large-scale data production happens at thousands of facilities across Europe every day. Data catalogues, data discovery standards and meta-data indexes will be developed in close alignment with HE Missions and Pillars. There needs to be a joint effort with the community to ascertain that harmonised approaches are established across the life sciences and avoid inadvertently driving divergence through the strong thematic projects.
- Provenance, reproducibility, traceability and secure management of sensitive data. Issues around data sovereignty, quality management, data security and provenance cuts across the whole of life-sciences - from research with data from human volunteers to maintaining compliance with the Nagoya protocol for biodiversity and bioresources. These are common, core issues for all the LS RI and will need investment to bridge the different national implementations and derogations. The issues also bridge into other thematic domains (e.g. social science) and the emerging European Health Data Space and will require a joint approach across thematic projects.
- Packaging tools and workflows for use in the cloud. EOSC-Life has developed expertise in the cloud deployment of software and workflows across all domains of the life

sciences with a core set of common services. Success of EOSC will depend on the abilities of scientists and support staff within all national facilities to embrace new technologies and build the skills and capabilities required to work within a pan-European cloud setting.

- Alignment of cost-recovery models for cloud-based storage and compute with trans-national facilities access. Experience from EOSC-Life shows that cost models from national centres and cloud providers are not always well aligned with European funding requirements. Realising the digital single market for research data will need a concerted effort to align administrative and business support across countries and institutes.

NuPECC. The large professional community engaged in ESCAPE-related science extends to tens of thousands of scientists.

The partners in ESCAPE recognise the strong synergies and potential commonalities which are there at several levels: the communities themselves are often overlapping, with multiple cross-overs between communities and within their research institutes; there are often common funding agencies for astronomy and particle/nuclear physics in many countries; and the computing facilities that all of these ESCAPE partners use are often host to both astronomy and particle physics experiments. Thus, the natural synergies of the science domains are also reinforced by these structural factors.

ESCAPE aims to address the Open Science challenges shared by its partners and the community. These challenges are technical, operational, sociological and scientific. Open Science allows scientific information, data and outputs to be more widely accessible and harnessed.

The recent update to the European Strategy for Particle Physics received input and strong support from the astronomy, astrophysics, and nuclear physics communities. In particular, in terms of computing, data management, and software tools, there is a recognition that the broader community faces similar challenges of managing Exabyte-scale datasets, complex software and a complex, distributed and heterogeneous computing infrastructure that is essential for optimising data processing and analysis across the available resources contributed by the RIs and their partner institutes. This apparent complexity (distributed computing) has nevertheless been shown to be an excellent way to optimise the use of the available funding in many countries - combining local access to data and facilities with a major contribution to international (often global) research partnerships.

The new generation RIs in ESCAPE are going to

ESCAPE



Background

European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures (ESCAPE) brings together a large fraction of the European research infrastructures (RI) in Astronomy, Astrophysics, Particle and Nuclear Physics. These RIs are ESFRI facilities and landmarks such as CTA, ELT, EST, FAIR, HL-LHC, KM3Net and SKA as well as other pan-European research infrastructures such as CERN, ESO, JIV-ERIC and EGO.

The ESCAPE project is supported and supervised by national institutes of the European Union member states that are organized in thematic consortia such as ApPEC, ASTRONET, ECFA and

operate more as open facilities, while others are incentivised by European science policy to move quickly towards Open Science, accelerating scientific discovery and disseminating results and knowledge with the community at large. ESCAPE has a central effective role in leveraging uses cases and transversal science projects to enhance the participation of researchers in Open Science and a unique ambition of enabling a multi-probe and cross-domain Open Science Commons research environment. The RIs in astronomy, astrophysics, particle and nuclear physics are willing to engage actively and collectively for current and future key societal challenges (Open to Society).

Thus, we see a strong motive for a cluster construct such as ESCAPE to play a significant role in building and making use of the inherent synergies between the RIs and their communities. This in turn will enhance the value of the significant long-term investments of the funding agencies by sharing facilities, tools, experience and forward-looking approach on societal and data challenges across the ESCAPE community.

It is clear that there are other coordinating bodies with which many of the RIs interact; ESFRI itself has an overall coordination role, the EOSC Association for planning the Open Data Science implementation in Europe, while the thematic and national agencies coordinate through consortia such as ECFA, NuPECC, ApPEC, AstroNET etc. However, we see a distinct role for ESCAPE - to build on synergies across the broad domain in large scale data management, computing, software, tools, and services, and the expertise of the communities, as well as to coordinate new open-science and open-to-society transversal projects among researchers. ESCAPE's scientifically and technically synergistic role is wider and more transversal than any single one of these coordinating bodies (who nevertheless are often part of ESCAPE's governance). This is a

unique combination and opportunity, which is also currently perceived by other ESFRI and world-class next generation RIs willing to join the ESCAPE programme and cooperative context.

Another aspect of the role of ESCAPE is to act as a single voice representing our community, towards funding and engagement opportunities such as with the European Commission, and specifically towards the European Open Science Cloud framework. We see ESCAPE as a strong voice in ensuring that EOSC can respond to the needs of our large communities, remembering that Europe is often only one part of our research partnerships and that we need to ensure ongoing global collaboration and interaction. ESCAPE is already currently moving forward more globally, with further international collaboration.

Thus, we see ESCAPE as a good vehicle for long term sustainability of this collaboration. We would rely on the thematic agencies to encourage new RIs to join ESCAPE. In addition, broader partnerships with other clusters would be beneficial.

Collaboration Mechanisms

The continuation of an ESCAPE partnership is strategically beneficial. Several possibilities on how a sustainable ESCAPE partnership could be constructed are considered. The current analysis progresses towards the following potential schemes:

- A large thematic consortium is expected, recognising common interests for cross-fertilisation Open Science and Open-to-Society actions and extending the existence of the cluster collaboration agreement after the ESCAPE project duration.
- Keep the (ESFRI) RIs at the strategy and steering roles would enable resources commitment to common projects that are useful for the RIs' implementation and operation.
- We might create a new collaboration with a revised agreement, more inclusive of further RIs aiming at joining the ESCAPE cluster and reflecting how the ecosystem has evolved,

potentially with the longer term in mind.

- MoU between all partners has been shown to work at large scale for other initiatives in our own domains (e.g. WLCG in particle physics, VO in astronomy, etc..) and it would help in supporting the evolution and operation of the ESCAPE Open-Science Virtual Research Environment (VRE).

ESCAPE as thematic platform operating its VRE within the EOSC relies on the sustainability of the consolidation of its current and future work programme. A funding scheme for the Science Clusters within Horizon Europe is an indispensable prerequisite for the sustainability of ESCAPE. It is key for further development of cross-cluster interdisciplinary scientific activities as well as strategic to leverage the further extension of the cluster scientific community engagement in response to the European Data Strategy.

Aspects of Collaborative Actions

There are many areas where we have already seen the benefits of a close collaboration in our thematic domain. These are all significant achievements and, in some cases, even structured operational roles for a long-term sustainable infrastructure for all of our RIs.

- Common exabyte-scale data management tools, in particular the data lake concept as a foundation for a FAIR data repository and enabler of Open Science.
- The catalogue of software and research products created by ESCAPE, and enabling the sharing and collaboration on open source scientific software, tools, data and knowledge.
- New collaborative cross-border software, workflows and methods developed by and for the benefit of the community at large.
- ESCAPE scientists are both data users and producers, therefore they are able to leverage the cluster cross-developments for education purposes and innovation schemes (in cooperation with SMEs).
- The HEP Software Foundation, initiated in the

context of WLCG but that aspires to a broader scope. This has successfully engaged the HEP community in all aspects of collaborative software and tools, and enables a collaborative mechanism to work on new problems, such as AI/ML, and to build further commonality. This structure could easily be broadened to the ESCAPE community as there is nothing inherently HEP-specific about it, and it has been recognised as a powerful voice in the community. Such evolution would be able to guarantee the implementation of an infrastructure for thematic software heritage to foster digital objects FAIRness at large.

- Solving the AI/ML need for labelled training data with the help of volunteers through Citizen Science experiments, enormously expanding the scientific EOSC users by orders of magnitude, while embedding education for the science-inclined public. This leverages open data for the benefit of society at large. Again, this structure could easily be broadened beyond ESCAPE.
- The VO alliance to leverage the ESCAPE framework to extend its practices and services to a larger scientific context and furthermore to interface the new large-scale data challenges that astrophysics and astroparticle physics will face with the advent of next generation RIs.
- Creating the capacity to build a credits system and to support the development of a policy framework for rewarding of scientists committed to working in open science.
- Building a long-term sustainable framework for cross-RI combined analysis, multi-messenger astronomy as a particularly important example.
- Bringing in the areas of detector/telescope R&D and technology at large as part of the ESCAPE platform, where the benefits of managing digital objects in a consistent way with the rest of the community may be advantageous and may open paths for innovation, business and society.
- The need to combine science, economy, technology and society through effective

visions and programmatic actions is crucial. Open Science and Open Innovation implies today unifying forward-looking approaches during the preparatory phase of any new Big Science facility integrating science needs, key-enabling technology advancement, socio-economic impact optimization, environmental impact minimization, energy and other resources management and industrial cooperation. Such an approach leverages multi-domain data and will imply the generation of a large spectrum of digital objects populating the main European Data Spaces. The cluster-platform guarantees the cross-border environment for addressing all these open challenges for society.

PANOSC AND EXPANDS



Background

The European Photon and Neutron Research Infrastructures (PaN RIs) are used by a large multidisciplinary scientific user community for understanding the structure and functioning of matter. Experimental projects are submitted by research teams, peer reviewed and, if successful, scheduled for beamtime. Typical experiments take between hours to several days of beamtime on the selected experimental setup (also called beamline or instrument). During the experiment, the research infrastructure provides the computational means for data acquisition, pre-processing and quality assessment. Full data analysis is usually taking place in the home

laboratories of the visiting scientists and takes months to years. The time from experiment to publication typically takes years.

The PaNOSC partners are the ESFRI roadmap infrastructures ESRF as coordinator, ILL, Eu-XFEL, ESS, ELI-DC, and CERIC-ERIC, EGI as e-infrastructure and GÉANT as close collaborator for the AAI. The main objectives of PaNOSC are to make FAIR data a reality in the partner RIs and connect the RIs to the EOSC. To do this PaNOSC has developed a FAIR data policy framework together with the national PaN RIs. PaNOSC is developing a common search API to allow searching for FAIR data across the PaN RIs. PaNOSC and ExPaNDS support standardising metadata through the Nexus metadata ontology. Emphasis is being put on promoting the use of Jupyter notebooks as a solution for reproducible data analysis. PaNOSC is extending Jupyter notebooks to display data files in HDF5 format efficiently. The HDF5 viewer is being developed as a web component which can be used by all communities using HDF5 (including a number of the science clusters). A remote data analysis portal (Vista) is being developed and adopted by the PaNOSC partners. Vista provides remote access services for running Jupyter notebooks but also remote desktops as a service. Simulation services for experiments and optical elements used on the beamlines are being extended and made available as a PaNOSC service. A training portal for PaN users has been deployed and is being extended. All these services will be made available via the EOSC services portal. PaNOSC will provide a solution for transferring data to user home institutes and in and out of the EOSC. PaNOSC services support an EOSC ready AAI solution based on eduTEAMS from GÉANT. Last but not least these services need to be sustained after the project and analysis will be made to propose a sustainability roadmap for the future.

All the services and activities of PaNOSC are shared with its sister project ExPaNDS comprising most of the national PaN RIs in Europe. ExPaNDS

is in many respects a sibling of PaNOSC and the project has largely similar goals.

In the last few years, the PaN RIs have experienced a shift to more complex experiments generating more complex and much larger data sets. This has led to the need to extend the access to the IT resources beyond the duration of the experiment and to call on the expertise of the facility staff to help in the data reduction and data analysis process. We also witnessed that some experiments, especially in tomography, generate so much data that carrying the data away from the RI becomes very problematic. These trends put considerable pressure on the facility staff and the IT infrastructure. Many of the visiting scientists do not have easy access to compute facilities in their home laboratory or university. All the above is clearly showing an increased delay between the experiments and the publications and needs to be addressed.

Collaboration Mechanisms

The production of FAIR data is well accepted by the PaN RIs but a lot of work remains to be done to fully implement FAIR for all experimental techniques covered by the RIs. The PaN user community is less active in adopting FAIR data best practices and will need more help, training and documented advantages on why to adopt FAIR data practices. New types of metrics and acknowledgments are required to incite scientists to adopt FAIR faster. LEAPS and LENS (League for European Accelerator-based Photon Sources and League for European Neutron Sources) and later ARIE (Analytical Research Infrastructures in Europe) are ideally suited to provide the mechanisms to sustain and further develop the outputs from PaNOSC and ExPaNDS. The LEAPS and LENS associations are fully functional and have both their collaborative governance structures. Both leagues are also closely working together for discussing and deciding on future common requirements and developments. Discussions have started to develop a cost-sharing

model required to sustain the developments and services of common interest to the PaN community emanating from the PaNOSC/ExPaNDS projects.

Aspects of Collaborative Actions

The PaN user community is intrinsically multi-disciplinary and covers an extremely wide range of scientific domains for which the analytical facilities provide the instruments and tools to investigate the structure and functioning of matter. The scientific challenges in the PaN RIs are increasing with the emergence of 4th generation storage rings of extreme brilliance allowing to go into new and uncharted domains of characterizing nano-materials with extreme precision. ML assisted experiments, data compression, remote access and data analysis services will all require to pool resources and share expertise between the PaN RIs and beyond. PaNOSC and ExPaNDS will have created a sound foundation from which other developments can be undertaken jointly. The PaNOSC software solutions for viewing HDF5 files, remote analysis portal, search API, Jupyter notebooks, data simulation services and e-learning platform are potential collaboration projects with the other clusters. All software is open source and can be shared amongst the communities. Setting up collaborations will furthermore increase software sustainability.

The collaboration and coordination with the other Science Clusters together with the activities pursued within the EIROforum are essential to exchange on best practices, share know-how, and broaden our views of how to advance in the all IT domains required for implementing an Open Science culture in Europe.

SSHOC



Background

The Social Sciences and Humanities Open Cloud Science Cluster (SSHOC) brings together a large fraction of the European research infrastructures in Sociology, Psychology, Economics, Political Sciences, Anthropology, History, Languages, Arts, Cultural Heritage and other SSH disciplines.

Its partners are ESFRI facilities and landmarks: the Common Language Resource and Technology Infrastructure (CLARIN), the Digital Research Infrastructure for the Arts and Humanities (DARIAH), the European Social Survey (ESS), the Survey of Health, Ageing and Retirement in Europe (SHARE), the Consortium of European Social Science Data Archives (CESSDA) and the European Research Infrastructure for Heritage Science (E-RIHS), complemented with platforms like the Association of European Research Libraries (LIBER).

SSHOC represents an impressive diversity of partners that are global data infrastructures (e.g., European Social Survey and Wageindicator.org), research communities and researchers from various universities and research institutions, museums, and finally SSHOC helps solidify the tangible link to industry with specialised and niche SMEs dealing with semantics and knowledge graphs, such as the Semantic Web Company and a marketing and communications SME, Trust-IT, specialised in communications and digital marketing for global data infrastructures.

SSHOC will create the social sciences and humanities area of the European Open Science Cloud (EOSC) thereby facilitating access to flexible, scalable research data and related services streamlined to the precise needs of the SSH community. SSHOC partners cover the full research data cycle - from design and collection up to research data sets & tools to facilitate use.

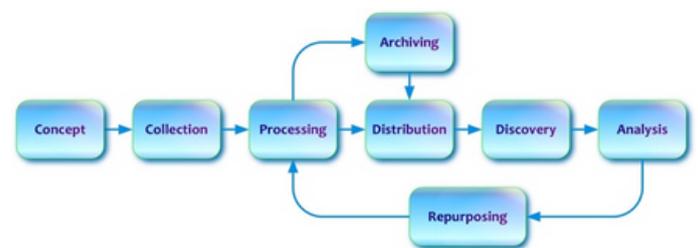


Figure 2 Data Life Cycle (DDI Alliance.org)

The design, collection, curation and processing of SSH data are important elements as they determine the quality of the data and metadata; these steps are time consuming and expensive (e.g., multiple countries, cultures and languages); data size, variety and complexity are keywords for the SSH domain, including sources and artefacts in humanities and cultural heritage. Many SSH domains face additional requirements to ensure privacy of respondents, to protect unauthorized access to data containing sensitive information. The investment in understanding this cross-cultural and cross-national landscape creates rich and powerful data which allow research questions to be answered that could not be answered with only national resources.

The SSHOC cluster stimulated and catalysed cooperation between the partners. This led to the joint development of tools & services, but also created a platform where over 300 researchers, data experts, survey & statistics experts, IT developers can meet.

The success of a project like SSHOC depends on

the quality and viability of its individual partners: in all domains, the international research infrastructures (the ERIC organisations) as well as the national infrastructures and facilities. These standing organisations must be able to do their core business and have the time and funds to develop and innovate. This was also revealed in developing EOSC and during the COVID-19 pandemic, when additional ad hoc projects were successfully launched to bring together data and contribute to targeted research. The clusters served as points of contact and managed to bring together specific partners from all clusters for these projects.

Collaboration Mechanisms

The SSHOC Science Cluster will continue its cooperation, and continue to represent the SSH community as a whole through this effort of collaboration, aiming to:

- Encourage and facilitate discussions on issues of interest common to the SSH disciplines.
- Optimise the use of resources and facilities.
- Coordinate the outreach activities and provide representation to the outside world.
- Take an active part in EOSC and other European or global initiatives.
- Simplify interactions (single point of contact) with the European Commission and other organs of the European Union (e.g., Eurostat).

We aim to remain flexible and agile, to be able to quickly shift points of interest on policy and research issues.

Aspects of Collaborative Actions

The 40+ tools and services that are developed in the SSHOC project will have an uptake by their partners. This ensures the continuity and viability of the tools and services.

Some tools can be upgraded to applications for other disciplines:

- CLARIN Switchboard (support to connect data and tools) and the Virtual Collection (users can select and save tools & data into an individual

‘shopping bag’).

- The cultural heritage Aioli-platform for 3D-annotation of artefacts.
- The sample management system for cross-national surveys and panels.
- The SSH Open Marketplace may serve as a best practice for other (EOSC-related) marketplaces.

The following tools & services will serve the SSH community directly:

- SSH Conversion Hub: Metadata conversion services between the most relevant metadata formats.
- The Automatic Verification Tool to verify translations of survey questions.
- The SSH Open Ontology for organising knowledge and information found distributed across various primary sources of information in the Social Sciences and Humanities Open Cloud, that is, a semantic interoperability framework for the description of the SSH data lifecycle.
- ARIADNE-plus for improved archaeological data management.
- SSH Open Training Toolkit: a curated registry of training resources (links to existing SSH training hubs, materials, events, reports) for trainers.
- Knowledge Graphs in Electoral Studies to bring together structured and unstructured data.
- The Ethnic and Migrant Minority Survey Registry (800+ surveys from 30+ (EU) countries) to easily search for and learn about existing quantitative surveys in this domain.
- Similarly, other research communities will also ensure the uptake of their services that have been developed in SSHOC.

SSH Research Infrastructures are at the eve of major developments in Artificial Intelligence, machine supported Text & Data Mining (e.g., digitizing medieval handwriting; automated extraction of keywords from PDF and spoken language; machine-assisted translations, discovery of fake news), and data interoperability with other domains (e.g., health and environmental data), and

across cultures and languages. New data types (social media, streaming data from municipalities and governments) contribute to the data deluge in SSH and may require new techniques for analysing and presenting outcomes.

For this, it is important that the community that was set up during the SSHOC project stays connected and continues to have a platform for communication and cooperation. Moreover, the emerging model for collaboration can be used to connect with the other Science Clusters, working on interoperability of data, combining metadata standards, and encouraging and facilitating research communities to work together on complex scientific and societal challenges (e.g., understand the long-run implications of the current COVID crisis), sustainable development goals, and – when needed – to quickly and efficiently join forces in case there is a new global crisis.

ESFRI SCIENCE CLUSTERS

POSITION STATEMENT

ON EXPECTATIONS AND LONG-TERM
COMMITMENT IN OPEN SCIENCE

DOI: [10.5281/ZENODO.4889503](https://doi.org/10.5281/ZENODO.4889503)

MORE INFORMATION

ENVRI-FAIR - <https://envri.eu/>

EOSC-Life - <https://www.eosc-life.eu/>

ESCAPE - <https://projectescape.eu/>

PaNOSC - <https://www.panosc.eu/>

SSHOC - <https://sshopencloud.eu/>



ENVRI-FAIR, EOSC-LIFE, ESCAPE, PANOSC and SSHOC have received funding from the European Union's Horizon 2020 research and innovation programme under the following grant agreement numbers: 824064, 823852, 824068, 824087, 823782.