



## D5.3

# Demonstrator of service integration in EOSC and systems

Work Package	WP5
Lead partner	INGV
Status	Final
Deliverable type	Report
Dissemination level	Public
Due date	30 June 2023
Submission date	30 June 2023

### Deliverable abstract

This deliverable showcases the progress and demonstrators of services integration between the ESFRI Cluster of Environmental Research Infrastructures and the European Open Science Cloud. It highlights the successful onboarding of the ENVRI Catalogue of Services onto the EOSC Marketplace, improving visibility and accessibility. Ongoing discussions with OpenAIRE aim to enhance the discoverability of ENVRI digital assets within EOSC. The document emphasises the importance of harmonisation, interoperability, and collaboration in achieving seamless integration and maximising the impact of environmental research infrastructures.



## DELIVERY SLIP

	Name	Partner Organisation	Date
Main Author	Daniele Bailo	INGV	22 June 2023
Contributing Authors	Keith Jeffery Angeliki Adamaki	UKRI/BGS ULUND/ICOS	1 June 2023 5 June 2023
Reviewer(s)	Peter Thijsse Dick Schaap	MARIS	19 June 2023 23 June 2023
Approver	Andreas Petzold	FZJ	30 June 2023

## DELIVERY LOG

Issue	Date	Comment	Author
V 01	25 May 2023	First Draft document	Daniele Bailo
V 02	29 May 2023	Contributions by other authors	Keith Jeffery Angeliki Adamaki
V 03	20 June 2023	Updates and feedback by reviewers	Peter Thijsse Dick Schaap
V 04	24 June 2023	Final Harmonised Version including reviewers' feedback	Daniele Bailo

## DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the Project Manager at [manager@envri-fair.eu](mailto:manager@envri-fair.eu).

## GLOSSARY

A relevant project glossary is included in Appendix A. The latest version of the master list of the glossary is available at <http://doi.org/10.5281/zenodo.4471374>.

## PROJECT SUMMARY

ENVRI-FAIR is the connection of the ESFRI Cluster of Environmental Research Infrastructures (ENVRI) to the European Open Science Cloud (EOSC). Participating research infrastructures (RI) of the environmental domain cover the subdomains Atmosphere, Marine, Solid Earth and Biodiversity / Ecosystems and thus the Earth system in its full complexity.

The overarching goal is that at the end of the proposed project, all participating RIs have built a set of FAIR data services which enhances the efficiency and productivity of researchers, supports innovation, enables data- and knowledge-based decisions, and connects the ENVRI Cluster to the EOSC.

This goal is reached by: (1) well defined community policies and standards on all steps of the data life cycle, aligned with the wider European policies, as well as with international developments; (2) each participating RI will have sustainable, transparent, and auditable data services, for each step of data life cycle, compliant to the FAIR principles. (3) the focus of the proposed work is put on the implementation of prototypes for testing pre-production services at each RI; the catalogue of prepared services is defined for each RI independently, depending on the maturity of the involved RIs; (4) the complete set of thematic data services and tools provided by the ENVRI cluster is exposed under the EOSC catalogue of services.

## TABLE OF CONTENTS

D5.3 – Demonstrator of service integration in EOSC and systems .....	4
1 Introduction.....	4
2 Environment of EOSC and other systems.....	4
3 Integration Plan.....	5
3.1 EOSC Marketplace / Profiles .....	5
3.2 OpenAIRE.....	5
3.3 Other systems (general metadata conversion).....	5
4 Integration progress .....	5
4.1 Portal/catalogue entry to marketplace .....	5
4.2 Ongoing discussions with OpenAIRE.....	6
4.3 Other systems .....	6
5 Individual RIs integration .....	6
6 Discussion.....	8
7 Conclusion .....	8
1. Annex A - Glossary .....	10

## D5.3 – Demonstrator of service integration in EOSC and systems

### 1 Introduction

This deliverable concerns service integration looking outwards from ENVRI towards EOSC and other systems and environments. Following Requirements collection (D5.1) and Implementation Plan (D5.2) the major effort has been in Task Force 1 (TF1) dealing with the catalogue of digital assets documented in D5.4, Task Force 2 (TF2) dealing with Authentication, Authorisation and Accounting Infrastructure (AAAI) and subsequent decisions integrating the catalogue into the ENVRI-Hub (D5.6). The provision of visibility of ENVRI digital assets externally requires some validation process to ensure appropriate quality: this is described in D5.5.

### 2 Environment of EOSC and other systems

The European Open Science Cloud (EOSC) is a long-term, large-scale endeavour aiming to enable researchers, industry and the public benefit from digital assets created and maintained by public funding. The ambition of the European Open Science Cloud (EOSC) is to provide European researchers, innovators, companies, and citizens with a federated and open multi-disciplinary environment where they can publish, find and re-use data, tools and services for research, innovation, and educational purposes. Successive projects undertaken during the last decade have produced a working prototype environment. The current EOSC Future project<sup>1</sup> is building on this existing baseline to deliver a platform with a durable set of user-friendly components that are designed for the long haul. They have adopted a system-of-systems approach to the EOSC platform, linking together other research portals, resources, and services to respond to the data needs of a wide range of researchers. ENVRI is involved in the EOSC Future project as one of the five science clusters, to establish a closer cooperation between the research communities and the e-infrastructures, that so far have laid the groundwork for EOSC. A major facet of this cooperation is to achieve more and better tuning between the research communities, which in practice manage the majority of the digital assets, software services, datasets, data products, software, workflows, documents and publications, and the e-infrastructures, aiming for getting more mutual understanding and this way, improving and optimising specifications of EOSC services and interoperability framework.

A good illustration of this process concerns ‘metadata schemas’: the EOSC prototype includes the concept of ‘profiles’ – essentially schemas describing metadata records that EOSC wishes to obtain from the suppliers<sup>2</sup>. These ‘profiles’ should support finding, accessing, interoperating and re-using digital assets, interacting with existing research communities and their infrastructures. Initially, these profiles, designed by the e-infrastructures, were inadequate, but progressively are being adapted under influence of the Research Infrastructures. The metadata Working Group of EOSC-Future has suggested an architecture where the EOSC marketplace (with schemas defined as profiles) should have a ‘zone’ dedicated to portals of Research Infrastructures (RIs) or – in the case of ENVRI which has integrated more closely the RIs within – the cluster Hub (ENVRI-Hub). Then, progressively, from this ‘zone’ distributed federated queries could be executed over the RI portals (or, in time following ENVRI, cluster portals) so allowing access to much richer metadata than that available under the schemas in the EOSC marketplace. Such an architecture evolves naturally into interoperating Science Knowledge Graphs (SKGs) thus providing open and highly linked access to the available digital assets.

ENVRI is well placed for this evolving architecture through the ENVRI catalogue of Services. The catalogue can provide a homogeneous representation of digital assets which are represented as heterogeneous metadata at the portals of the RIs – thus allowing query across environmental digital assets from atmosphere, ocean, solid earth, and biosphere. The catalogue is already a SKG using CERIF (Common European Research Information Format – an EU Recommendation to Member States and utilised for recording information on research for ~25 years)<sup>3</sup>. ENVRI has mapped successive iterations of the schemas proposed by successive EOSC projects to the catalogue metadata schema and so maintained readiness for interoperation. An advantage of the ENVRI-Hub and its catalogue is that the

---

<sup>1</sup> <https://eoscfuture.eu>

<sup>2</sup> <https://eosc-portal.eu/eosc-providers-hub/what-are-eosc-profiles>

<sup>3</sup> <https://eurocris.org/services/main-features-cerif>

ENVRI community needs to maintain only one interface to the evolving EOSC specifications (from the ENVRI-Hub) rather than from every RI in ENVRI.

Concerning other systems, CERIF has already defined interfaces to multiple metadata formats so allowing for interoperation. Mappings between CERIF and other metadata schemas can be done quickly, since it is a very rich metadata format with base entities and linking entities, the latter relating to each other the base entities (with role and temporal validity) in a fully connected graph. This has been demonstrated within EPOS (one of the RIs in ENVRI) where 17 different metadata formats have been mapped to a CERIF-based catalogue.

## 3 Integration Plan

### 3.1 EOSC Marketplace / Profiles

As indicated above, ENVRI has mapped CERIF to proposed schemas from successive EOSC projects. Within EOSC-Future, representatives of the clusters have criticised (positively but continuously) the schemas provided since they do not allow for the information required – as metadata – for research community users to discover, contextualise, access and re-use the digital assets. The architecture proposed by the EOSC-Future Metadata Working Group overcomes this obstacle, by first providing access to the individual portals of clusters or RIs and then federated distributed access, thereby utilising the richness of the metadata of each community.

### 3.2 OpenAIRE

Since the inception of EOSC, OpenAIRE – sustained also by a series of projects – has been involved in establishing and operating a number of EOSC core services. OpenAIRE provides an SKG (also based on CERIF) linking together outputs of research (such as publications, datasets, and software) with projects, organisations, persons and more. This has a rich metadata format.

During the ENVRI-FAIR project and led by the EPOS team which had previously held discussions with OpenAIRE for the purposes of EPOS, discussions with OpenAIRE were initiated. A succession of technical meetings has refined progressively how ENVRI digital asset metadata collated in the Catalogue of Services could be onboarded to EOSC via OpenAIRE. The discussions are not yet concluded, but – following the architectural roadmap discussed in Section 2 - the logical way forward would be to populate the marketplace with the individual metadata records in the ENVRI Catalogue of Services, while keeping the metadata describing the ENVRI Catalogue of Services in the marketplace profiles. However, as RIs are part of ENVRI but also have their own identities and responsibilities, and are as well as partners in more communities than only ENVRI, both onboarding procedures (individual and centralised) are available.

### 3.3 Other systems (general metadata conversion)

CERIF provides the capability for easy mapping to other metadata schemas, and many converters based on such mappings exist already. This allows the ENVRI-Hub Catalogue of Services (subject to appropriate agreements/governance) to interoperate with other systems utilising different metadata formats. As said previously, individual onboarding is also available.

## 4 Integration progress

### 4.1 Portal/catalogue entry to marketplace

The integration progress involved establishing the portal/catalogue entry of the ENVRI-Hub Catalogue of services into the EOSC marketplace.

After months of meticulous work and collaboration, the ENVRI Catalogue of Services has now been successfully onboarded onto the European Open Science Cloud (EOSC) Marketplace. This significant milestone means that users can now access and explore the catalogue directly on the EOSC Marketplace

using this link: [ENVRI Catalogue of Services on EOSC Marketplace](#)<sup>4</sup>. The integration of the ENVRI Catalogue of Services into the EOSC Marketplace enhances the visibility and accessibility of the environmental research infrastructures data and services, promoting their use in various research projects, also by making all the metadata available through its exposure by means of web APIs. The onboarding process will continue to be refined in the coming months to ensure optimal user experience and utility.

## 4.2 Ongoing discussions with OpenAIRE

OpenAIRE, a key player in the EOSC ecosystem, has been actively engaged in discussions with the ENVRI-FAIR project. Led by the EPOS team, these discussions aim to explore how ENVRI digital asset metadata can be onboarded to EOSC through OpenAIRE. Technical meetings have been held to refine the integration process, ensuring seamless connectivity and metadata linkage between ENVRI and OpenAIRE. While these discussions are ongoing, they hold great potential for leveraging OpenAIRE's rich metadata format and its SKG infrastructure to enhance the discoverability and accessibility of ENVRI digital assets within EOSC.

## 4.3 Other systems

In addition to the integration with the EOSC marketplace and discussions with OpenAIRE, the ENVRI-FAIR project recognises the importance of interoperating with other systems utilising different metadata formats. CERIF, with its capability for mapping to multiple metadata schemas, provides a solid foundation for such interoperability. By establishing appropriate agreements and governance frameworks, the ENVRI-Hub, through the ENVRI Catalogue of Services, can seamlessly collaborate with other systems, expanding the reach and impact of the environmental research infrastructures beyond the ENVRI Cluster. This is also demonstrated by the ENVRI Science Demonstrators<sup>5</sup>

## 5 Individual RIs integration

In parallel with the onboarding through the ENVRI Catalogue of Services, a number of ENVRI RIs, involved in Science Projects as part of the EOSC Future project, have moved on with an individual RI integration. The ENVRI cluster is one of the five Science Clusters participating in the EOSC Future project, which began in April 2021. As described already in chapter 2, the project aims to unite e-Infrastructures and Science Cluster communities to establish an operational EOSC Platform focused on technology, interoperability, resources, user engagement, and user experience. The participation of science clusters in the project is primarily carried out through Science Projects (SPs). Several ENVRI RIs are involved in the development of two SPs: a) "Climate Change Impact on Biodiversity and Ecosystems in Europe - Assessing the impact of Non-Indigenous Invasive Species (NIS) in European ecosystems" and b) "Dashboard for the state of the Environment." These SPs fall under EOSC Future's WP6: Integration of Community Services and Products into EOSC. The goal of WP6 is to populate, connect, and support the supply and demand sides of the EOSC. The SPs serve to demonstrate the value of EOSC for cross-disciplinary challenges and applications.

To coordinate and support this process, a Service Integration Team was established within the EOSC Future project. The team drafted a Service Integration Roadmap with the aim of providing a robust mechanism for supporting the service integration process, with specific Key Performance Indicators (KPIs). These KPIs were divided into three phases, spanning different months, and involved integrating services from Science Clusters. To facilitate the coordination and progress monitoring a Service Integration Matrix is maintained within EOSC-Future. This matrix lists services from Science Clusters and their status in terms of onboarding and integration with Core services. The matrix serves as a crucial instrument for coordination, as well as for demonstrating planning and progress.

During the ENVRI-FAIR TF6 meeting in Amsterdam, (10-11 October 2022), the participants discussed the situation regarding the EOSC. Specifically, the focus was on how ENVRI-FAIR could contribute to the success of the EOSC within the context of the Service Integration Matrix. It was agreed that ENVRI-FAIR would contribute on three different levels. Firstly, the ENVRI cluster would provide cluster and RI services through its Science Projects, ensuring their onboarding to the EOSC Marketplace as they

<sup>4</sup> [https://marketplace.eosc-portal.eu/services/eosc.epos.envri\\_catalog\\_of\\_services](https://marketplace.eosc-portal.eu/services/eosc.epos.envri_catalog_of_services)

<sup>5</sup> <https://envrihub.vm.fedcloud.eu/sddetail/35>

became ready. Later, decisions would be made regarding the integration of specific components with Core Services. Secondly, ENVRI-FAIR would onboard and integrate the ENVRI Catalogue of Services before the end of the ENVRI-FAIR project. Lastly, RIs would interact with the teams developing the EOSC Core Services to onboard their services in the EOSC Marketplace directly and work on further integrations with the EOSC Core Services (see fig. 1), thus a step further in integration than currently possible via the ENVRI Catalogue. The goal was to conduct an inventory of RI services that fulfilled specific criteria, such as being directed towards data discovery and/or access, possessing an operational readiness level (TRL7 or higher), and considering both services with user interfaces and machine interfaces. The ENVRI Service Integration Matrix was compiled, indicating the options for onboarding RI services and their integration with selected Core Services. To expedite the required momentum and foster collaboration between the e-infrastructures and the RIs, several meetings were organised, aiming at enhancing mutual understanding between the e-infrastructures responsible for developing and upgrading the EOSC Core Services and the RIs providing services to the EOSC. Several individual ENVRI services have been onboarded in the EOSC Marketplace up to this point, with some of them integrated with the upgraded core services, wherever feasible. At the cluster level, the ENVRI Dashboard (for the State of the Environment) has been integrated with the EOSC Marketplace, AAI, monitoring and Helpdesk, while the ENVRI Catalogue of Services has been onboarded in the Marketplace as a single service.

## EOSC-FUTURE development of EOSC Core Services

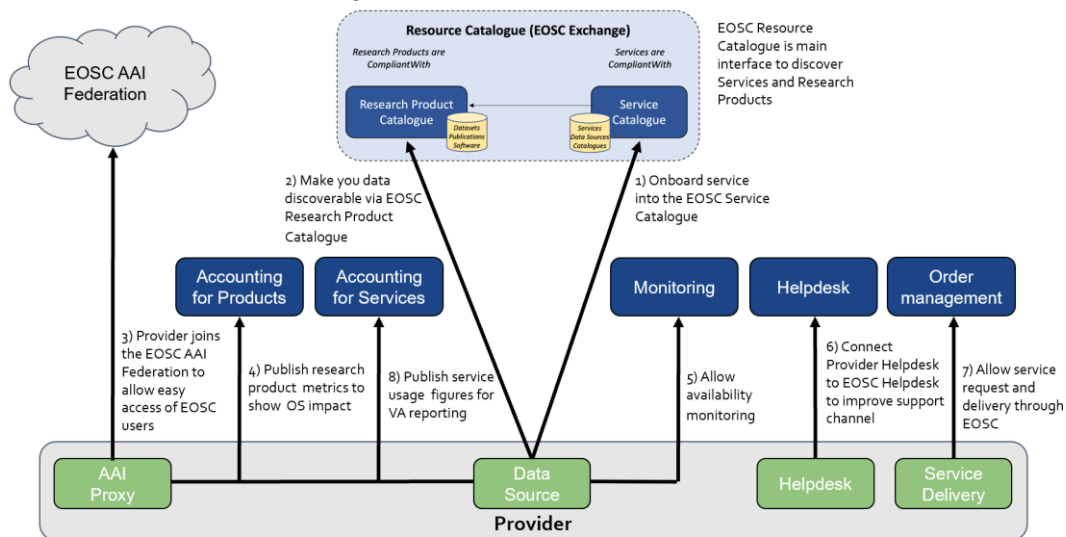


Figure 1: Overview of the Core Services which have been released by M18 (end September 2022).



## 6 Discussion

ENVRI-FAIR (and the ENVRI community) has achieved - with considerable success - one of the objectives of the project by increasing FAIRness in the metadata records of individual RIs allowing onward access to digital assets such as services, datasets and others. The second objective – integration with EOSC and other systems - has been somewhat hindered by (a) the changing policies and governance (especially conditions for onboarding of metadata describing digital assets) of EOSC; (b) the changing technology of EOSC supporting (a); (c) the request in other EOSC related projects of onboarding also individual RIs from ENVRI (and other clusters). Fulfilling these requirements, has required (and requires) considerable efforts within the ENVRI cluster for ensuring and coordinating that all RIs apply a harmonised and consistent approach.

The major thrust in WP5 of ENVRI-FAIR towards both objectives was the provision of a homogeneous metadata catalogue describing digital assets derived from heterogeneous metadata records of the RIs. Early in the project it was agreed that this was managed by TF1 led by INGV and UKRI/BGS - both basing architecture and technologies on EPOS-ERIC RI. TF1 within WP5 recommended and implemented a mechanism to integrate services from the RIs into a homogeneous metadata catalogue. For reasons of security, privacy, acknowledgement, and citation TF1 recommended that only the portal accessing the catalogue (and the corresponding API) should be onboarded to EOSC.

Catalogue records for individual RI digital assets are more complex to onboard, and there has been a cross-ENVRI group working with OpenAIRE to develop an appropriate mechanism to ensure homogeneity also at this level of detailed metadata. This involves rich metadata in SKGs and – as OpenAIRE is also based on CERIF – integration from the ENVRI catalogue should be easy.

The architecture developed by TF1 within ENVRI-FAIR allows integration of services both for facing EOSC (with onboarding first of just the portal metadata and subsequently - through OpenAIRE - the rich detailed metadata) and also for interoperating with other systems (i.e. other metadata standards). The work in TF1 used a formal systems engineering approach based on consensus across many RIs of ENVRI. Potential evolution of this approach for the future involves Research Infrastructures (RI) and other service providers publishing interoperable and standardised service descriptions alongside their offerings, using RDF format for universal accessibility. This method eliminates the need for manual publication.

In the parallel EOSC-Future project WP3 (architecture) the five science clusters through the Metadata Working Group have proposed a strategy for all clusters as outlined above, namely a zone of EOSC profiles catalogue for portals, and detailed metadata suitable for research purposes in SKGs compliant with OpenAIRE (but probably even richer) and federated for interoperation. The proposal is that ENVRI should align with this architecture to be active in future participation.

## 7 Conclusion

It is clear that the architectural design – based on that of EPOS – provides the necessary capability for integration with EOSC and other systems, as well as integrating the heterogeneous metadata records describing the digital assets of the RIs within ENVRI. Integration and a first validation have been achieved at ENVRI level by creating a homogeneous catalogue using CERIF and integrating this with the ENVRI-Hub so providing a single interface to EOSC. This insulates the RIs in ENVRI from changes in EOSC while preserving the appropriate acknowledgement/citation as required by the ENVRI RIs.

The successful integration of ENVRI to EOSC – by onboarding to EOSC a metadata record describing the ENVRI portal - clearly impacts the project since this was one of the objectives (the other being to improve FAIRness of the ENVRI RI digital assets). Furthermore, the decision to use CERIF for the metadata catalogue of the ENVRI-Hub provides opportunities – as demonstrated elsewhere – for integration with many other systems in the environmental domain and even wider. Ongoing work with OpenAIRE opens the possibility of federated SKGs among the science clusters thus providing wider and deeper interoperation.

The exercise of harmonisation of heterogeneous metadata records from the RIs in ENVRI to a homogeneous metadata standard to provide a comfortable end-user or API interface has been a learning experience for the RIs in ENVRI. And this will be continued in the future.



The ENVRI RI organisations – whether ERICs or not – benefit from the developed catalogue/ENVRI-Hub environment which provides homogeneous access over heterogeneous digital assets, and a first level of validation. The research community of ENVRI feels this benefit most, but ongoing efforts related to the EOSC-Future project aim to allow interoperation across the science clusters thus allowing integration of environmental assets with – for example – social science assets. Beyond the research community, the consultancy sector of industry, as well as all industries either protecting the environmental domain and/or extracting value from it will benefit. Finally, government agencies benefit from improved and integrated environmental information to aid strategic planning decisions.

not yet accepted by EC

## 1. Annex A - Glossary

AAAI	Authentication, Authorisation, and Accounting Infrastructure
CERIF	Common European Research Information Format
DCAT	Data catalogue Vocabulary
EPOS-DCAT-AP	DCAT Application Profile for EPOS
DDSS	Data, Data products, Software and Services
DIAS	Data and Information Access Services
EOSC	European Open Science Cloud
EPOS Strategic Plan	Defines the activities of EPOS-ERIC
GEP	Geohazards Exploitation Platform
ICS-C	Integrated Core Services - Central Hub
ICS-D	Distributed Integrated Core Services Distributed
PID	Persistent Identifier
SDLC	Software Development Life Cycle
SKG	Science Knowledge Graph
TCS	Thematic Core Services