Periodic report #2
Summary

ENVRI
FAIR

Periodic report #2
Summary

Project Coordinator
Andreas Petzold

Project Co-Coordinator
Anca Hienola

Project Manager
Katrin Seemeyer

1.

BACKGROUND

Project overview, goals & objectives

Background: project overview, goals & objectives

European Environmental Research Infrastructures (ENVRIs) on the ESFRI level are core facilities for providing data, research products and services from the four subdomains of Earth system science – Atmosphere, Marine, Solid Earth, and Biodiversity/Terrestrial Ecosystems. 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 topics covered by the ENVRIs span the entire range of scientific objectives relevant for Earth system monitoring.

ENVRI-FAIR is the cluster project of the ESFRI Environment (ENV) Domain, aiming towards the provision of services compliant with the FAIR principles. The overarching goal of ENVRI-FAIR is to advance the findability, accessibility, interoperability, and reusability (FAIRness) of digital assets, in particular research data, products, and services, provided by the ENVRI Cluster, and to connect them to the emerging service ecosystem of the European Open Science Cloud (EOSC).

ENVRI-FAIR targets the development and implementation of both technical frameworks and policy solutions that make subdomain boundaries permeable for environmental scientists and prepare Earth system science for the new Open Science paradigm.

Cross-discipline harmonisation and standardisation activities, together with the implementation of joint data management and access structures at RI level, facilitate the strategic coordination of observation systems required for interdisciplinary science. In addition, common policies, open standards, interoperability solutions, operational services, and stewardship of data based on the FAIR principles will essentially help maturing RIs to solve those common problems.

Background: project overview, goals & objectives

ENVRI
FAIR

One of the highest priorities of ENVRI-FAIR, and obviously among the biggest challenges, is the design of a machine-actionable ENVRI Catalogue of Services that also supports the integration of ENVRI digital assets into EOSC.

Through this catalogue the service providers will be able to make their assets findable and accessible by mapping their resources into common and rich metadata standards. The human interaction with the FAIR services will be demonstrated by means of a web application, named the ENVRI-Hub.

"One of the highest priorities of ENVRI-FAIR, and obviously among the biggest challenges, is the design of a machine-actionable ENVRI Catalogue of Services that also supports the integration of ENVRI digital assets into EOSC."

The ENVRI objectives at a glance

ENVRI FAIR

1 To further develop common standards, protocols and policies for the data life cycle, including cataloguing, curation, provenance and service provision within the ENVRI cluster, with specific consideration of the FAIR principles including interoperability, and of the tools and methods created during the preceding EU-projects ENVRI and ENVRIplus.

2 To align these policies and standards with more extensive European policies (e.g. ISO 19115 INSPIRE) as well as with relevant international developments.

3 To develop and implement the necessary tools for reaching Objective 1 in each research infrastructure, thereby adopting an open approach for sharing data and software.

4 To improve the skills of research infrastructure personnel to develop and sustain knowledge on Research Data Management and FAIRness, including both cross-cutting and subdomain-specific knowledge, and on the FAIR infrastructures resulting from Objectives 1 and 2 through an extensive training programme.

5 To increase the potential for innovation of each research infrastructure by establishing a specific ENVRI-FAIR section in the EOSC service catalogue, with the aim of stimulating common pre-commercial procurement processes and dissemination of outcomes and thus enhancing the uptake of research infrastructure services by private partners.

6 To establish cohesion with the global research infrastructure landscape, including clusters and regional/ international initiatives in the environmental sector; maintain ENVRI community knowledge with particular consideration of developing integrated activities.

7 To expose thematic data services and tools from the research infrastructure catalogues to the EOSC catalogue of services, COPERNICUS, GEO, and other end-users.

2.

SUMMARY

**An overview
of the most important
accomplishments**

Summary: An overview of the most important accomplishments

To achieve the high-level ENVRI-FAIR goals, the work conducted in the second project phase was mostly dedicated to the implementation of the tasks defined in the various implementation plans, with the ENVRIs working towards not only improving the FAIRness of their own data and services, but also reflecting their efforts at a higher level by becoming FAIR as a cluster. To build on a common foundation, the key technical gaps were prioritised and the ENVRIs identified common requirements and design patterns and work together to make meaningful use of existing technical solutions that enhance their FAIRness. The implementation of FAIR-enabling resources at RI level and the development of solutions for interoperability is progressing and will support the ingestion of the ENV Domain into EOSC with their specific user requirements and experiences.

A highly coordinated joint effort

The common developments in the ENVRI cluster were guided by both a subdomain driven approach to make use of the synergies in the respective communities, and an overarching coordination through the six technically focussed task forces established in the first reporting period.

These task forces concentrate on the most important aspects of FAIRness as identified in the analysis of the first FAIR Implementation profile (FIP) survey, and target the requirements for the implementation of a federated cross-RI AAI, the definition and introduction of the ENVRI Catalogue of Services, Persistent Identification throughout the whole data lifecycle, the use of ontologies, linked open data, vocabularies, and certification of repositories, licenses, data attribution and GDPR issues, and finally, the ENVRI-Hub

Summary: An overview of the most important accomplishments

visualizing the architectural design of the services by science demonstrators and giving practical hands on examples..
The task forces have developed into the main drivers of technology-related developments, and serve now as the powerhouse for the ENVRI-FAIR project.

Through the ENVRI-Hub, the ENVRI community shares their FAIRness experience, technologies, and training as well as research products and services. The architecture and functionalities of the ENVRI-Hub are driven by the applications, use cases and user needs. Its four main pillars are the ENVRI Knowledge Base as the human interface to the ENVRI ecosystem,

the ENVRI Catalogue of Services as the machine-actionable interface to the ENVRI ecosystem, the ENVRI Training Catalogue giving access to training resources by the ENVRI community and finally subdomain and cross-domain scientific use cases as demonstrators for the capabilities of service provision among ENVRIs and across Science Clusters. The architecture is designed in anticipation of interoperation with EOSC and is intended to act as a key platform for users and developers planning to include ENVRI services in their workflows.

"The task forces have developed into the main drivers of technology-related developments, and serve now as the powerhouse for the ENVRI-FAIR project."

The most important accomplishments of the second project phase include:

1

Increasing convergence between the subdomains in terms of implemented FAIR enabling resources

2

Implementation plan updates

3

Joint development of the FIP Wizard together with the GO FAIR Foundation to generate FIPs, and its application in three consecutive assessments of ENVRIs for monitoring the progress in FAIRness

4

Metadata mapping to DCAT-AP and services registered in the ENVRI Catalogue of Services

5

Implementation of subdomain scientific demonstrators as Jupyter notebooks

6

Progress in repository certification

7

Implementation of the ENVRI-Hub demonstrator

3.

CONCRETE RESULTS

Six major achievements implemented

ENVRI-Hub Demonstrator

The ENVRI-Hub Demonstrator is designed and implemented as a web application, which acts as an integrator by bringing together existing ENVRI services and interoperable services across research infrastructure boundaries.

Exploring the potentials of the ENVRI-Hub already from the design phase, the ingestion of metadata from ENVRI assets such as the ENVRI Knowledge Base, the ENVRI Catalogue of Services and the ENVRI

Training Catalogue is realised, aiming to provide the users with functionalities that are relevant to e.g. the discovery of environmental observations, services, tutorials and other available resources.

The chosen architectural pattern for the development of the ENVRI-Hub can be compared to a classical n-tier architecture, comprising 1) a data tier, 2) a logic tier and 3) a presentation tier. To integrate the different ENVRI tools while preserving

the independence of the applications, the ENVRI-Hub demonstrator opens the machine to machine world of the ENVRI Catalogue of Services for human users.

Following a centralised architectural approach, the ENVRI-Hub serves as a harvester entity, collecting data and metadata and integrates the ENVRI Knowledge Base and the ENVRI Catalogue of Services. It bridges these ENVRI platforms into one single portal.

Scientific Use Cases

Starting from domain-specific or cross-domain use cases environmental science stories crossing disciplinary boundaries are designed. A set of Jupyter Notebooks developed by the contributing RIs and accessible from the ENVRI-Hub are used to demonstrate and validate the capabilities of service provision among ENVRIs and across Science Clusters. These examples show what a user can achieve through the ENVRI-Hub. In one of the examples, a user-friendly well-structured Jupyter Notebook that makes use of research infrastructures'

application programming interfaces (APIs) jointly plots in a map the geographical locations of several Marine and Atmospheric stations. Clicking on a station the meta data of the station are shown including links to service endpoints or landing pages of the stations. That's what a user needs as start-point for further investigation.

The FAIR principles provide a firm foundation defining the layer that supports the ENVRI-Hub structure and the preliminary results are promising.

Considering that the APIs can become discoverable via a common ENVRI catalogue of services, the ENVRI-Hub aims to make full use of the machine-actionability of such a catalogue. In the future to this kind of use case might be solved by SPARQL federated search functionality within ENVRI-Hub itself.

FAIR Implementation Profiles and Convergence

In the second project phase, the RIs of the ENV Domain regularly conducted FAIRness assessments using FIPs generated by the FIP wizard to monitor their progress in FAIRness. The process was facilitated via domain-specific workshops and training events on the FIP approach.

The FIP wizard will become publicly available as one of the services provided by the ENVRI-Hub and is already recognised by RDA and other FAIR supporting initiatives as a valuable tool.

Training in FAIRness

The further improvement of personnel skills is handled through continuous FAIRness training activities. By the end of the second project phase, 15 training events like webinars, workshops and schools were organised, with an average of 40 participants. Topics covered include FAIRness of data and services, provenance documentation, virtual research environments, FAIR data and metadata description, and semantic interoperability of RI data. The ENVRI Training activities contribute substantially to EOSC Future and to various RDA Interest and Working Groups.

Policy Framework

The analysis of the Policy Landscape in the ENVRI domain has been revised and updated to support other relevant tasks including the training workshops held in ENVRI-FAIR. The policies to be implemented in the ENVRI cluster are aligned with EOSC Rules of Participation. The developed policy framework is set up as RI agnostic to allow later the adaption of this methodology also outside of the ENVRI Science Cluster.

Cross-Domain Coordination through Science Clusters

Stimulated by EOSC activities, and by the participation of the coordinators of the five ESFRI Science Cluster projects in EOSC Future, the cooperation among the five ESFRI clusters deepened towards cohesion and alignment of work and resulted in an implemented collaboration framework of regular meetings and consultations. Driven by these activities their role in EOSC is growing continuously in relevance with the ENVRI-Hub being considered one of the major community portals in EOSC.

"The cooperation among the five ESFRI clusters deepened towards cohesion and alignment of work and resulted in an implemented collaboration framework of regular meetings and consultations."

THE EARTH IS OUR LAB.

ENVRI
FAIR

Contact & more info:
envri.eu