D10.5
Final Report on Implementation Activities

Work Package	WP10
Lead partner	UKRI
Status	Final
Deliverable type	Report
Dissemination level	Public
Due date	30 June 2023
Submission date	26 June 2023

Deliverable abstract

D10.5 "Final Report on Implementation Activities" describes the work towards making the RIs in the solid Earth subdomain FAIR in terms of FAIRness of the assets and EOSC-interoperability.
In this deliverable we therefore describe for each WP10 task, the work done to address the gaps identified by Deliverable "D10.2 Roadmap for implementation of FAIR concepts", prioritised and harmonised in D10.3 "D10.3 Report on coordination, prioritisation, and harmonisation in the Solid Earth Subdomain".

ENVRI-FAIR (www.envri-fair.eu) has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824068

DELIVERY SLIP

	Name	Partner Organization	Date
Main Author	Keith G Jeffery	UKRI	15 May 2023
Contributing Authors	Daniele Bailo	INGV	15 May 2023
	Tendry Randriamalala Niantso	BRGM	15 May 2023
	Marc Urvois	BRGM	15 May 2023
	Tor Langeland	NORCE	15 May 2023
	Jan Michalek	UIB	15 May 2023
	Alessandro Spinuso	KNMI	15 May 2023
	Luca Trani	KNMI	15 May 2023
	Michele Manunta	CNR	15 May 2023
	Ivan Rodero	EMSO	15 May 2023
	Wayne Crawford	EMSO	15 May 2023
Reviewer(s)	Peter Thijsse Zhiming Zhao	MARIS UvA	12 June 2023 12 June 2023
Approver	Andreas Petzold	FZJ	26 June 2023

DELIVERY LOG

Issue	Date	Comment	Author
V 0.1	10 January 2023	Initial outline for contributions	Keith G Jeffery
V1.0	17 May 2023	Final	Keith G Jeffery
V2.0	30 May 2023	Final revised	Keith G Jeffery
V3.0	20 June 2023	Final after reviewer comments dealt with	Keith G Jeffery

DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the Project Manager at 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

D10.5 – Final Report on Implementation Activities

1 Introduction

The primary objective of this deliverable, D10.5 “Final Report on Implementation Activities,” is to detail the activities carried out to address the gaps identified in Deliverable D10.2 “Roadmap for implementation of FAIR concepts” [1] and prioritised in D10.3 “Report on coordination, prioritisation, and harmonisation in the Solid Earth Subdomain.” The focus is on the implementation of FAIR principles within the Solid Earth domain community to establish a shared knowledge base, promote standardised data practices, and facilitate interoperability and data sharing [2].
FAIR principles, while essential, do not provide explicit guidelines or methodologies for Research Infrastructures (RIs) to develop and implement the necessary software and technologies to achieve FAIR data. In this deliverable, we outline the technical activities undertaken to address the gaps identified in D10.2 and prioritised in D10.3, thereby moving from FAIR principles to actionable steps that enable RIs to effectively implement FAIR data practices within the Solid Earth domain community.

2 Overview

Research Infrastructure (RIs) implementers typically follow a system development life cycle (SDLC) process when creating or enhancing an existing RI. In the present deliverable, we utilised an SDLC influenced by the waterfall development model with rapid cycling, described in previous work [3]. This SDLC included the following stages: a) analysis, involving use cases and requirements gathering; b) design, which entails architecture design and identification of architectural components that align with requirements; c) implementation, through software development and the adoption of appropriate technologies; d) testing; and e) operation and maintenance. In this report, we concentrate on stages b) to d) and employ a rapid systems development strategy based on pitches or brief sprints of work within the Shape Up method [4]. This was a key element guiding the entire work that produced the outcomes reported in the current deliverable.

Tasks 10.1 and 10.2 provided the FAIRness analysis and implementation roadmap, while D10.3 reported on the harmonisation efforts to achieve a consistent FAIRness level and EOSC interoperability across the subdomain by outlining the work conducted in different EPOS building blocks, represented by the outcomes of Tasks 10.3 to 10.7. In the current deliverable, based on the proposed SDLC and the work presented in deliverables D10.2 and D10.3, we describe the results of Tasks 10.3-10.7, which address various areas of work within EPOS and EMSO, as planned in D10.2 to bridge the identified FAIR gaps. These tasks range from the Integrated Core Services – Central (ICS-C) (T10.3) to the Integrated Core Services – Distributed (ICS-D) (T10.4, where a prototype has been developed in the seismology community) and two of the Thematic Core Services (TCS) domains involved in ENVRI-FAIR – seismology, satellite, and marine, as represented by EMSO (T10.5-10.7). Task 10.8, which focuses on policies, is also discussed since it is integrated with the implementation activities.

3 Task 10.3 ICS-C

Task 10.3 aimed to improve the ICS-C in terms of service metadata and subdomain FAIRness maturity.

3.1 FAIRness

This chapter describes the steps to improve – through implementation - FAIRness in ICS-C. ICS-C is the central hub or EPOS consisting of the rich metadata catalogue of assets stored in CERIF and software to manage user interaction; access to the metadata catalogue; dispatch of requests to asset suppliers; and maintenance of the metadata catalogue by the TCS asset suppliers.

In general ICS-C is FAIR; it complies with the Force-11 Principles as follows:

Table 1: FAIR Principles and EPOS Conformance

FAIR PRINCIPLE	EPOS-ICS
TO BE FINDABLE:	
F1. (meta)data are assigned a globally unique and eternally persistent identifier.	ICS-C catalogue has a unique EPOS persistent resolvable Identifier for each record and also supports federated (role-based, temporal duration limited) IDs related by semantic relationships to the core ID.
F2. data are described with rich metadata.	ICS-C catalogue uses CERIF[1] (a EU recommendation to Member States) as the backend model which is demonstrably a rich superset of other metadata standards, thus encouraging improved searching, contextualisation, accessibility, interoperability and re-use. ICS-C uses EPOS-DCAT-AP[2] based on DCAT v1[3] for the metadata ingestion.
F3. (meta)data are registered or indexed in a searchable resource.	ICS-C catalogue is a registered searchable resource. Certification via CoreTrustSeal[4] is being investigated.
F4. metadata specify the data identifier.	ICS-C metadata records not only specify the asset identifier but may provide a qualified reference (I3) to it with role and temporal validity
TO BE ACCESSIBLE:	
A1 (meta)data are retrievable by their identifier using a standardised communications protocol.	ICS-C provides this facility with both API and GUI
A1.1 the protocol is open, free, and universally implementable.	ICS-C provides this facility through API and GUI
A1.2 the protocol allows for an authentication and authorisation procedure, where necessary.	ICS-C provides authentication aligned with AARC/GEANT blueprint architecture[5]. Various authorisation schemes exist and work on harmonising is ongoing
A2 metadata are accessible, even when the data are no longer available.	ICS-C provides for "tombstone metadata" also linked with curation and provenance information
TO BE INTEROPERABLE:	
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.	ICS-C has the capability by providing JSON API serialisation. The API is an extended version of Open Search. CERIF provides n-tuples of base entities and link entities thus providing a fully connected graph structure. While CERIF can be (and has been) implemented in several language environments, in EPOS we use classical n-tuples to ensure referential and functional integrity in a closed-world environment so permitting deduction and induction.
I2. (meta)data use vocabularies that follow FAIR principles.	CERIF- used in ICS-C - has a semantic layer for vocabularies, imported from various sources and following FAIR principles since the semantic layer of CERIF is a mirror image of the syntactic layer in properties such as integrity.
I3. (meta)data include qualified references to other (meta)data.	As indicated under F4, CERIF provides rich qualified references or minimally references with referential and functional integrity

[1] https://www.eurocris.org/services/main-features-cerif
[2] https://www.epos-eu.org/representing-cross-disciplinary-knowledge-solid-earth-sciences-epos-dcat-ap
[3] https://www.w3.org/TR/2020/SPSD-vocab-dcat-20200204/
[4] https://www.coretrustseal.org/
[5] https://connect.geant.org/2019/04/30/the-final-aarc-blueprint-architecture-a-community-first-approach

FAIR PRINCIPLE	EPOS-ICS
TO BE RE-USABLE:	
R1. meta(data) have a plurality of accurate and relevant attributes.	As indicated under F2, CERIF provides for rich metadata with many attributes, stored in n-tuples connected in a fully connected graph structure using semantic and temporal relationships (qualified references). EPOS-DCAT-AP – as the intermediate format for ingestion - supports progressively the richness of CERIF.
R1.1. (meta)data are released with a clear and accessible data usage license.	CERIF has the capability to store licence documentation (including appropriate links) and also to provide a licence name for searching. However, work is ongoing to translate licence information to authorisation keys for access.
R1.2. (meta)data are associated with their provenance.	CERIF comes with inbuilt provenance because of the temporal attributes in the role-based linking relationships (qualified references) and the EPOS catalogue is populated appropriately
R1.3. (meta)data meet domain-relevant community standards.	In EPOS 17 different community metadata standards have been converted to CERIF to provide a harmonised catalogue, so domain-relevant standards are preserved in terms of content, but in a harmonised form with richer syntax and semantics. The exchange format for ingestion used is EPOS-DCAT-AP which was created – and is being updated incrementally - to match CERIF richness as EPOS communities improve their metadata. The API answers to a request is with JSON which is a good starting point for interoperation.

As indicated from the above table, there are several areas of ongoing work that are being addressed:

- Identifiers
- Repository Certification (CoreTrustSeal)
- Authentication and authorisation
- Vocabularies and vocabulary tool
- Licence translation for autonomic access (links with authentication)
- Extending EPOS-DCAT-AP and CERIF model to accommodate additional metadata needed

It is worth noting that some of this work was coordinated with other RIs in ENVRI through the WP5 Task Forces in an attempt to increase the interoperability across ENVRI.

3.2 Identifiers

In the geoscience domain many identifier systems are in use including DOI, URI, UUID, community specific identifiers, locally unique identifiers. Moreover, different communities have different identifier systems used within their community (commonly internationally) and thus following standards outside of EPOS. Thus, EPOS has to utilise the federated ID mechanism of CERIF to accommodate different IDs for the same digital object – but commonly related to access in different roles or by different communities.

Work is ongoing to identify the different identifier systems in use and their degree of FAIRness. Some are reported in the following chapters.

3.3 Repository Certification (CoreTrustSeal)

there is some advantage (mainly to provide users with some level of confidence in its integrity) in registering the EPOS ICS-C catalogue with CoreTrustSeal. Some EPOS asset suppliers have already such approval. However, the effort involved to fulfil the requirements is not inconsiderable and so a cost-benefit analysis is planned. The major benefit is the assurance for the user to use data coming from a trusted source and respecting a certain degree of curation and data perenniality. The major disbenefit is the resource required to achieve certification.

3.4 Authentication and Authorisation

Within the EPOS IT Board development environment (Shape-Up, which requires development activities to be structured as ‘pitches’ with clear specifications and resources) there is a pitch on authorisation. It has collected information from the TCS and documented the landscape. It has already agreed that we use AARC/GEANT authentication mechanisms with tokens, evolving together to the OpenID standard. Thereafter, the intention is to utilise metadata information concerning the rights and responsibilities of the user and the availability conditions of the asset to provide authorisation keys to be encrypted into the authentication token. This work is ongoing and a proof of concept involving 3 RIs within ENVRI (EPOS, AnaEE and EMSO) has demonstrated that even with different Identity providers (IdPs), federated access and single-sign-on works smoothly for user authentication. Work continues - linked with the policy work of T10.8 - to agree across ENVRI a set of policies to allow authorisation tokens to be federated thus providing open access.

3.5 Vocabularies

In link with other communities in ENVRI and wider, and in order to facilitate cross-domain search, providing a consistent way of managing vocabularies prevents communities redoing the work done by others. This boils down to choosing the right tools and defining the policies of managing vocabularies: namespace, workflow and responsibilities for vocabularies. Being pragmatic, we will reuse the one used in the TCS geology community[6]). A cross-functional category register has been initiated from TCS seismic vocabulary. It will be reused and extended by others, and it can be aligned to existing standard. In this way, interoperability is ensured.
The vocabulary publication tool is deployed and accessible. It is a linked data-based registry tool called UkgovLD[7], already known and used by many authoritative organisations. It provides a web API + a GUI for creating, managing, publishing, and querying (via LD-API[8] or SPARQL[9]) registers. Thus, it allows managing the whole workflow and life cycle of a vocabulary and its entities.

The main vocabulary type managed in this tool are Simple Knowledge Organisation System (SKOS)[10] thesauri (Concept Schemes or Collections). SKOS provides a simple semantic framework for semantics of concepts, i.e., their definitions as well as the hierarchy and relations that organise them. In addition, other registry features are included in the tool: for instance, it allows a nesting structure so vocabularies can be organised under containers. Since it is a linked data-based tool, it provides direct dereferenceable identifiers (HTTP URIs) for all the managed vocabulary and vocabulary concepts. In addition to URI, direct access to vocabularies content, this tool provides a SPARQL endpoint service for accessing the data in another standard way (SPARQL), which allows exploiting the full potential of semantic definitions and relations. The developments related to this tool are still active, allowing it to benefit from new features such as register of register federation for example.
The tool has been installed and will soon be open with an EPOS-ERIC based URI.

---

[6] http://data.geoscience.earth/ncl/
[7] https://github.com/UKGovLD/registry-core
[8] https://json-ld.org/spec/latest/json-ld-api-best-practices/
[9] https://www.w3.org/TR/rdf-sparql-query/
[10] https://www.w3.org/2004/02/skos/

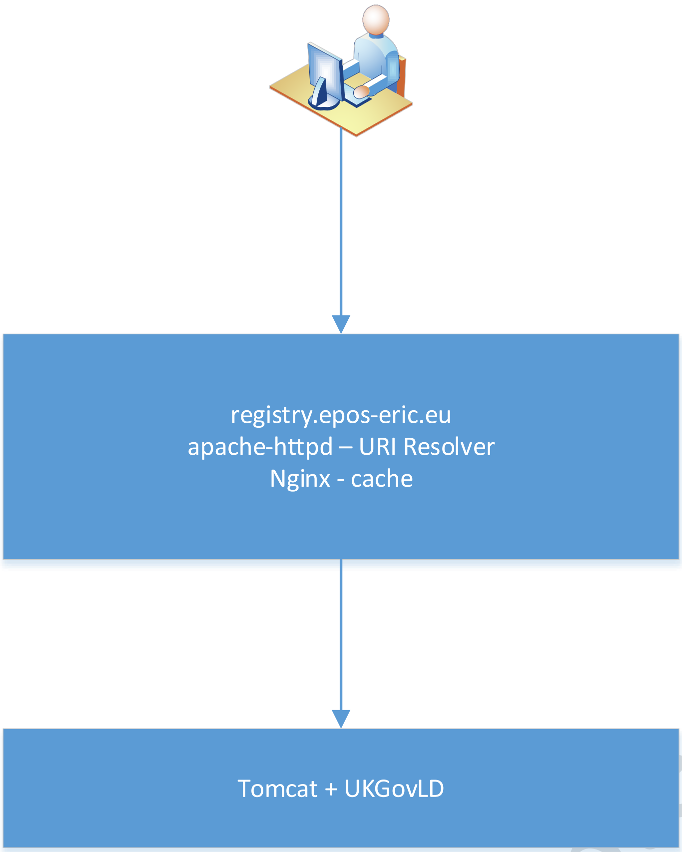


Figure 1 Vocabulary tool technical architecture

The current plan and state is as follows:

- Define the use of a vocabulary tool → done
- Start the implementation within EPOS landscape including translation into and ingestion into CERIF → ongoing: this is a continuous activity
- Identify the common guidelines for vocabulary→ done
- Refine the guidelines according to EPOS landscape → ongoing
- Advertise this tool and guidelines to the different bodies → introduction done in ICS-TCS meeting

3.6 Licence translation for autonomic access (links with authentication)

As indicated above licence information can be stored (a) as a document (with a link from the metadata record); (b) as a licence name. However, the real requirement is for the implied access rules to digital objects of the licence to be encoded such that they can be used within an authorisation framework. Work on licence policy is ongoing in the EPOS Policy Group. Each asset is to be licensed, according to the wishes of the owning organisation. However, this defines the scope of asset use, it does not necessarily indicate the authorisation constraints concerning the asset related to (classes of) users. The authorisation pitch (mentioned above) will investigate how best to represent licence intent in metadata and hence through to authorisation mechanisms.

3.7 Extending EPOS-DCAT-AP towards the CERIF model

EPOS-DCAT-AP is used as an intermediary format for ingestion of metadata records, providing a target for conversion from the many metadata formats used by asset suppliers. Once metadata from many asset suppliers are in EPOS-DCAT-AP it is easier to convert to the richer CERIF standard used for the EPOS metadata catalogue.

Work has been done in extending EPOS-DCAT-AP to align with the CERIF model implementation, first to cover many metadata elements not available in DCAT v1 but available within CERIF and required by the asset suppliers and users, coloured gold in (**Fehler! Verweisquelle konnte nicht gefunden werden.**).

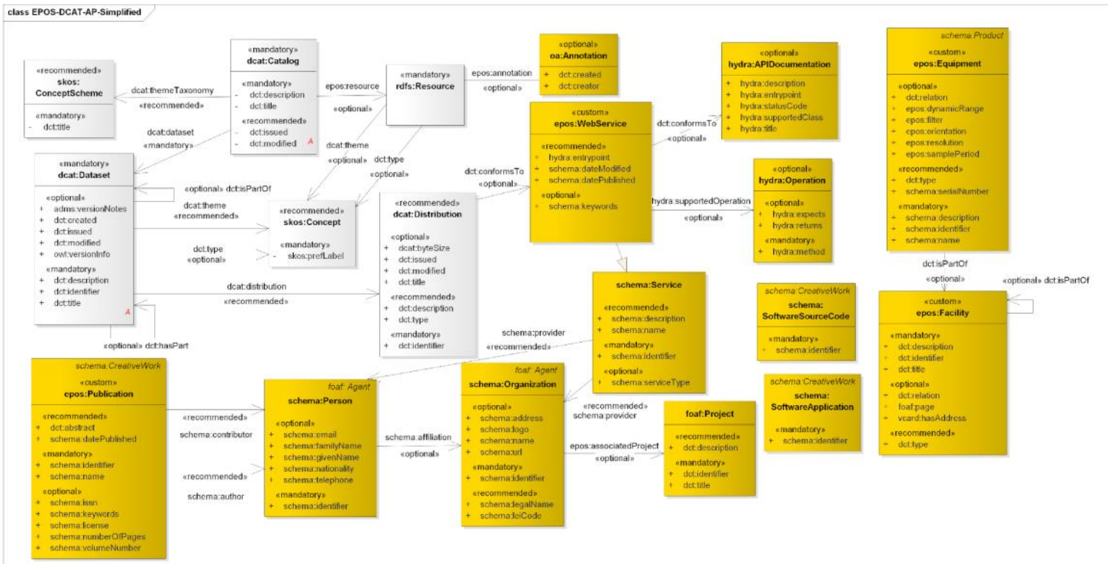


Figure 2 Extending DCAT v1 to EPOS-DCAT-AP

This (**Fehler! Verweisquelle konnte nicht gefunden werden.**) demonstrates clearly how much DCAT had to be extended (to EPOS-DCAT-AP) to approach the richness of CERIF – and then only for a limited set of entities common to the asset suppliers of EPOS. Current work is extending EPOS-DCAT-AP even further to support the CERIF entities Facility and Equipment, required for several purposes including Trans-National Access (TNA). We intend to demonstrate this extension with EPOS communities involved in ENVRI-FAIR. Further extensions are likely to be needed in EPOS-DCAT-AP to manage the metadata required as defined in the areas of work above. There is consequential work on ingestion from EPOS-DCAT-AP into the CERIF metadata catalogue of ICS-C.

As EPOS wants to improve its FAIRness, we intend to study a potential convergence with DCATv2. However, since EPOS is quite mature in its model, this convergence will undergo study to match the EPOS Delivery Framework, as we cannot break the actual features of EPOS.

3.8 Search API

The Search API has been defined based on open search with additional endpoints.
However, there are already existing ISO standards that could cover the needs of EPOS and enhance EPOS FAIRness.
Indeed, by reusing an already standardised API this would greatly facilitate the usage by any user: human or machine.
We aim to find potential better suited API endpoint with the following constraints:

- Have to be standardised: DESIRABLE
- Opensource tools already existing: DESIRABLE
- Output in preference by descending order: DCAT-AP/json, geojson, json: MUST
- Can be extended to support CERIF/XML: DESIRABLE
- Content negotiation for the previous model and serialisation: DESIRABLE

So far, we have found these potential API offerings: STAC, OGC API Records. We have investigated intermediary solutions such as VocPrez[11] (based on SKOS data model), SpacePrez (based OGC API specification and the GeoSPARQL data model) and CatPrez (for DCAT data catalogues). These subsystems are included in Prez[12], which provide a read-only access to Knowledge Graph data according to particular domain profiles. The purpose is to see where and how we could partially or totally reuse the existing API framework.
But also, as EPOS is quite mature in its development, it should be investigated how and to what extent the EPOS experience would benefit these API communities by providing them EPOS use cases in order for them to enhance their API.
As such a presentation of EPOS has been done in OGC meeting for this community to take into account EPOS uses case which comes with the richness of the whole Earth Science community. This provoked interest and some interest in cooperation.

4 Task 10.4 - ICS-D

4.1 Introduction

ICS-D is the term for the distributed ICS services. They are to be managed from the ICS-C portal environment. The overall planned architecture for the life of a workflow in EPOS is indicated by the following diagram (**Fehler! Verweisquelle konnte nicht gefunden werden.**). The idea is to support a Virtual Research Environment (VRE) to allow a researcher or other end-user to compose workflows to satisfy their requirements easily, accurately and optimally. The workflows are likely – after collecting and contextualising assets - to involve analytics, simulations and visualisations.

However, in order to gain relevant experience to ensure the architecture is applicable, pilot projects have been done
(a) external to ENVRI-FAIR in the VRE4EIC project[13] (where ENVRIplus[14] was also involved) and where a general workflow deployment mechanism was provided, tested using TAVERNA[15];
(b) in other EC projects concerning workflow deployment rather than composition where EPOS was represented namely PaaSage[16] concerning optimal deployment based on computing resources) and MELODIC[17] (optimising based also on data location and locality).

This chapter describes a major effort being made in the seismological community to provide an ICS-D based on the use of Jupyter. It outlines performed and planned work on the implementation of FAIR principles for the EPOS ICS-D services. These services will form a VRE for data analysis and visualisation, which can be launched from a data-discovery session in the EPOS ICS-C data portal.

[11] https://rdflib.dev/VocPrez
[12] https://github.com/RDFLib/Prez
[13] https://vre4eic.ercim.eu/
[14] https://www.envriplus.eu/
[15] https://taverna.incubator.apache.org/
[16] https://paasage.ercim.eu/
[17] https://h2020.melodic.cloud/

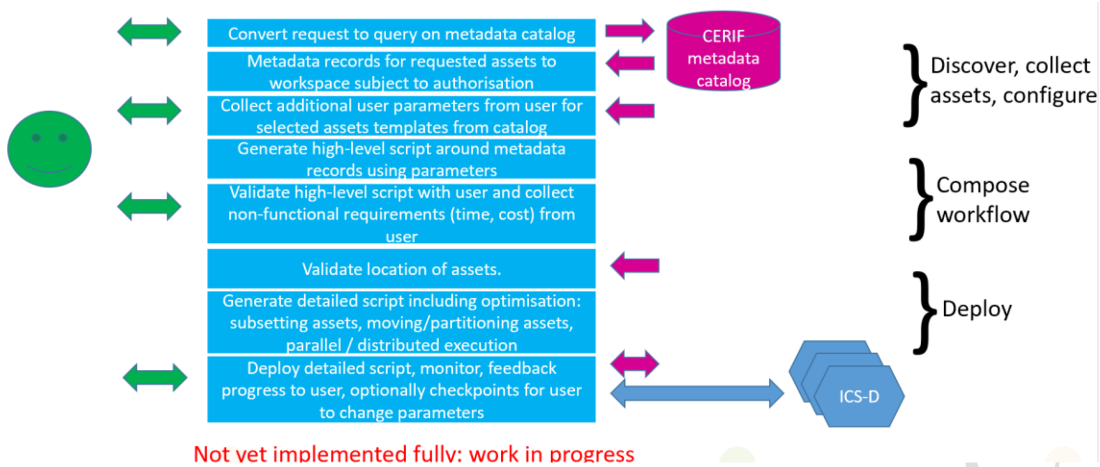


Figure 3 General EPOS Workflow

4.2 Components

The ICS-D to be demonstrated in this project consists of the following components (see Figures below):

- SWIRRL API[18]. This is a framework used for executing workflows for staging required data and for starting the specific ICS-D, in this project specifically Enlighten-web and Jupyter notebook.
- Enlighten-web[19] – This is a web application for interactive visual analysis. In this project we will implement a visualisation workflow that prepares data from the ICS-C workspace and that uses Enlighten-web for visualising the data.
- Jupyter notebook – This is a web programming environment for analysis. It will be made available in the same manner as described for Enlighten-web.
- Hosting facility - Deployment of the ICS-D services at the EPOS e-Infrastructure.

4.2.1 SWIRRL

SWIRRL is a web service API that allows Science Gateways to easily integrate data analysis and visualisation tools in their websites and re-purpose them to their users. A comprehensive overview of the system is shown in (**Fehler! Verweisquelle konnte nicht gefunden werden.**) below. The API deals, on behalf of the clients, with the underlying complexity of requesting and organising resources in a target cloud platform hosting the Kubernetes container-orchestration system. By combining storage and tools, such as Jupyter notebooks[20] and Enlighten-web, implemented as containerised services, the API creates dedicated working sessions on-demand.

Thanks to the API's workflow execution endpoint, which spawns job executing Common Workflow Language (CWL)[21] workflows for data staging and batch processing, SWIRRL sessions can be populated with raw data of interest collected from external data providers. Staged data is considered immutable with identifiers from SWIRRL. In the occurrence of updates, older versions are preserved.
The system is designed to offer customisation and reproducibility. Notebooks in the Jupyter environment can be further customised with additional or updated libraries, and the provenance of such changes is automatically captured. The recording of provenance is performed for each method of the API's affecting the session.

SWIRRL generates and stores provenance data thanks to two dedicated components: the PROV-Template Catalogue and the Neo4j database. Both are integrated as microservices and exposed through the API (**Fehler! Verweisquelle konnte nicht gefunden werden.**).

[18] https://gitlab.com/KNMI-OSS/swirrl/swirrl-api
[19] https://www.norceresearch.no/en/research-theme/enlighten-web
[20] https://gitlab.com/KNMI-OSS/swirrl/jupyterswirrlui
[21] https://www.commonwl.org

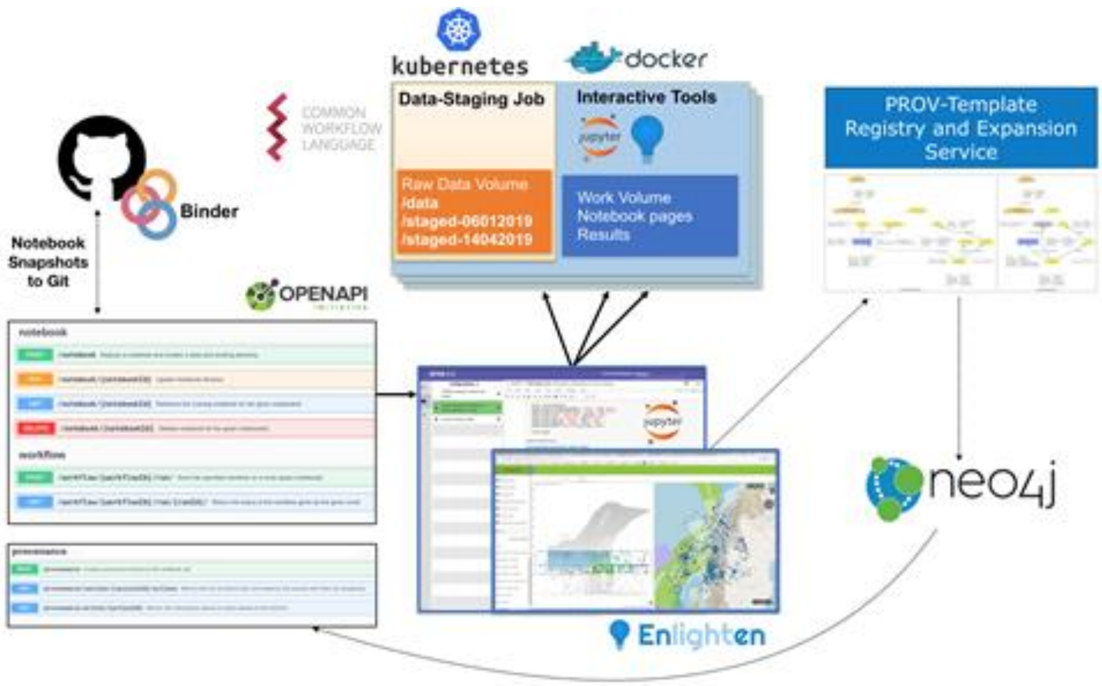


Figure 4 SWIRRL Components overview

4.2.2 Enlighten-Web

Enlighten-web facilitates interactive visual analysis of large multidimensional data sets. Enlighten-web uses the Cesium map engine[22]. The user can use brushing and linking to explore complex data sets to discover correlations and interesting properties hidden in the data. In this project, an Enlighten-web Docker image has been delivered to KNMI for their integration in the workflow. We have done some adaptations of the Enlighten-web code for this purpose.
The Kubernetes cluster for running ICS-D services has been deployed at BGRM, reusing the Infrastructure as a Code approach from EPOS.

4.3 Filling the Gaps

We have performed a FAIR assessment and gap analysis with regards to FAIRness of the ICS-D components. Based on the gap analysis, tasks were defined for filling the identified gaps. These tasks are described in D10.2 (Roadmap for implementation for implementation of FAIR concepts) [1]:

- Define metadata for services (Jupyter Notebook as a service and Enlighten-web visualisation service)
- Define metadata for service instances
- Use case driven integration of the ICS-D with the ICS-C
- Authentication/Authorisation
- Licensing
- Implement SWIRRL API methods offering restoring actions
- Define use case/demonstrator to demonstrate the approach

The following sections summarise the work on these tasks.

4.3.1 Define metadata for the ICS-D services

We have defined metadata for the ICS-D services. The main elements in these metadata are:

[22] https://cesium.com/

- Metadata for three concepts; "SWIRRL Workflow", Literate Programming and VisualAnalytics.
- Metadata for three web services; SWIRRL Data Staging Workflow, JupyterLab and Enlighten-web.
- Metadata for organisations and contact persons.

The metadata was defined using the SHAPEness METADATA EDITOR[23], and stored in a turtle file using the EPOS-AP vocabulary[24]. An EPOS-DCAT-AP SHACL file that can be used for validating EPOS RDF graphs is also provided[25].

INGV deployed a testing environment which contains these metadata in the catalogue and shows the services in the EPOS ICS-C GUI. **Fehler! Verweisquelle konnte nicht gefunden werden.** shows a visualisation of the structure of the metadata for the ICS-D services.

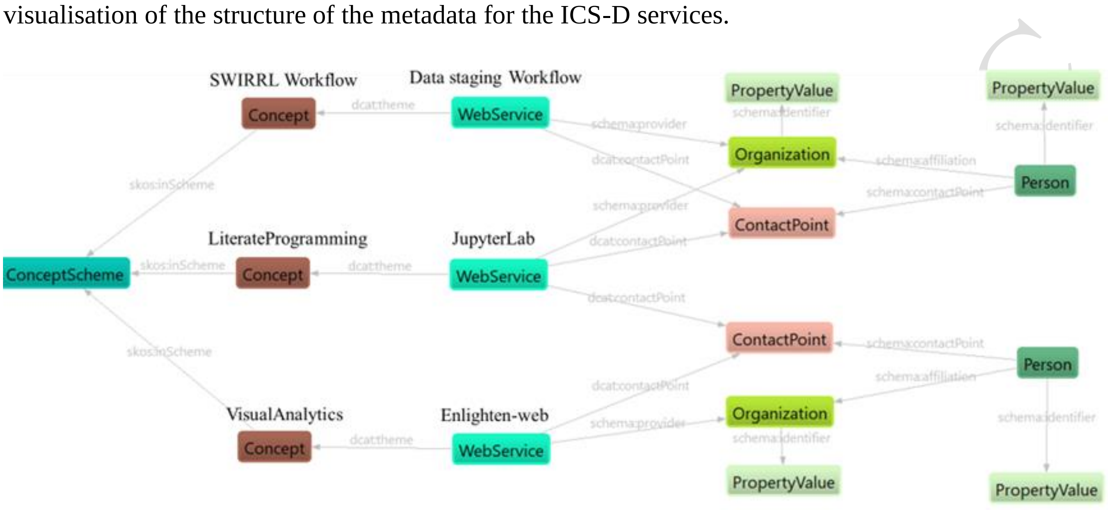


Figure 5 A visualisation of the structure of the metadata for the ICS-D services from the SHAPEness METADATA EDITOR. We added labels on the Concept and WebService nodes to illustrate the contents of the metadata. The other nodes contain information about the involved organisations, persons and contact points.

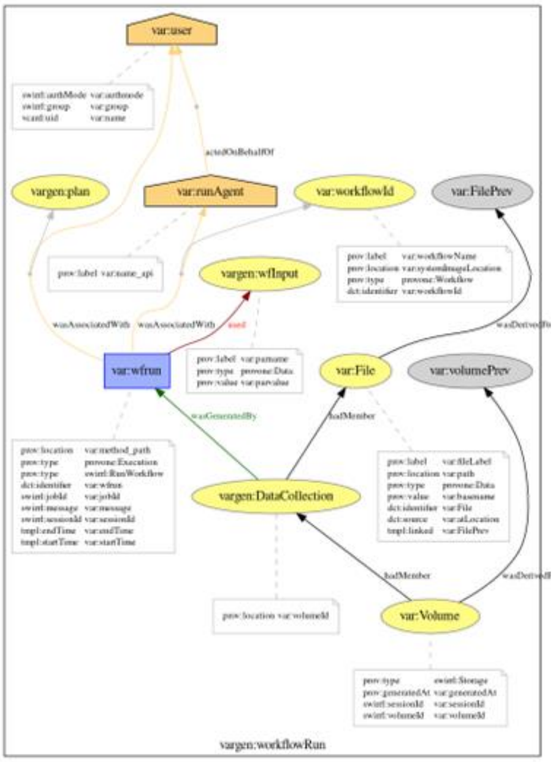
4.3.2 Define metadata for service instances - provenance

SWIRRL supports generation and storage of provenance data thanks to two dedicated components: the PROV-Template Catalogue[26] and the Neo4j database. Both are integrated as microservices and exposed through the API.

To utilise the provenance support in SWIRRL, we defined provenance templates associated with the services offered by the API. For instance, templates describe activities affecting the Jupyter Notebook Service, such as creation, environment update [5] and restore, as well as Workflow executions, see Figure 7. The latter make sure that newly produced data files are thereby traceable throughout their updates and linked to the remote archive, the collecting and preprocessing processes, and environment.

We have also created a template for the instantiation of an Enlighten-web Service, see **Fehler! Verweisquelle konnte nicht gefunden werden.**. We have furthermore configured Enlighten-web to store specifications of visualisations on json files on the server. This will allow SWIRRL to include these files in the provenance for visualisation sessions. Moreover, new templates can be designed aiming at achieving specific FAIR objectives for this tool.

[23] https://epos-eu.github.io/DCAT-Metadata-Editor/
[24] https://www.epos-eu.org/epos-dcat-ap
[25] https://github.com/epos-eu/EPOS-DCAT-AP/blob/EPOS-DCAT-AP-shapes/epos-dcat-ap_shapes.ttl
[26] https://iagos-comm.iek.fz-juelich.de/dmsf/files/5182/view



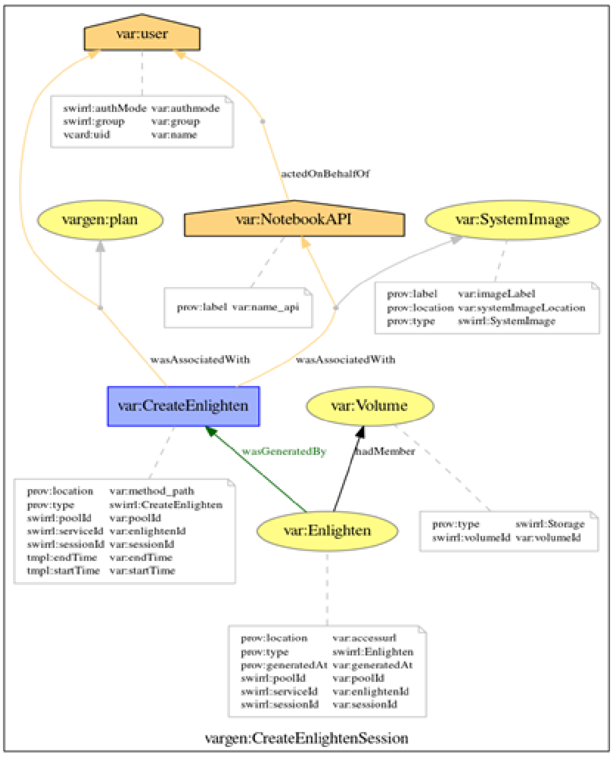


Figure 6 Provenance Templates used to describe the execution of a Workflow (left) and the instantiation of a new Enlighten Service (right) in SWIRRL.

To achieve improved FAIR-ness of the visualisation sessions, we needed to enable SWIRRL to stage immutable data for the visualisations. This means that data files must be created by executing the data search URLs that are retrieved from the ICS-C workspace. We considered two alternatives. The first was to create a separate microservice for this purpose. The second alternative was to create a tool that can be run by SWIRRL as part of the data staging workflow. We decided on the second alternative, and NORCE provided a Docker image with Python code for preparing Feather files from data search URLs from the ICS-C workspace. The Feather files can then be read by Enlighten-web. Together with example data files, the Docker image was made available for testing in SWIRRL by KNMI. Some updates were done after feedback from KNMI. A few adaptations were also required in Enlighten-web. Currently, the data staging process supports seismic event data and WMS map layers.

4.3.3 Use case driven integration of the ICS-D with the ICS-C

The overall idea is that having collected the relevant assets in the workspace, the user through the ICS-C triggers the workflow. However, ICS-C needs to keep track of the workflow execution and deal with both parameterising the workflow before deployment and dealing with the workflow results returned to ICS-C. Basically, the agreed approach is that the ICS-D shall provide the ICS-C with the following information to be used in the metadata:

- URL including search parameters (data source)
- URL to describe the data format returned by the URL
- Link to tools that are compatible with the data format

ICS-C is triggering and controlling the workflow. It is the responsibility of ICS-C to arrange the workflow based on the understanding of the ICS-D API provided by this metadata information from ICS-D to ICS-C.

4.3.4 Authentication/Authorisation

There is ongoing work within the EPOS central (ICS-C) development programme on authentication involving gathering requirements from the TCS communities and designing a mechanism to extract from metadata the authorisation parameters and encode them as keys on authentication tokens (see under Authentication in Section Task 10.3 ICS-C).

SWIRRL does not support Authentication yet. This is a task that will take into account the authentication technology used by the clients adopting it. However, SWIRRL supports delegation of authentication (i.e., confirming user identity) to GitHub. This allows users to store the snapshots generated by SWIRRL to their repository but does not provide full authentication as needed for authorisation in an EPOS context.

4.3.5 Licensing

The intention of EPOS is to make software available as open source with a suitable licence and also to license binaries in a similar way. A Policy Group is working on this currently, alongside management of a liability disclaimer and informed consent to terms and conditions of use, cookies and personal data privacy. Meantime, users can make their analysis performed on SWIRRL which is publicly available on GitHub.

4.3.6 Implement SWIRRL API methods offering reproducibility controls

SWIRRL allows users and software clients to update Jupyter services with additional libraries. These updates are traced in provenance and can be restored on-demand.

The API also supports the production of Snapshots for Jupyter workspaces that are directly stored on GitHub accounts. Snapshots include software and data (or means to access the data) and can be re-deployed onto third parties' infrastructure and tools, such as mybinder.org. We will also implement the possibility to restore full snapshots into SWIRRL itself.

A SWIRRL JupyterLab extension has been implemented with an interactive panel that allows users to control the restoring of updates and the production of snapshots. These are developed as JupyterLab extensions installed automatically in the notebook instances managed by SWIRRL (**Fehler! Verweisquelle konnte nicht gefunden werden.**).

4.3.7 Use case for test/demonstrator

This task is closely linked to task force 6 (TF6) of Work Package 5 (WP5) of ENVRI-FAIR.

We are currently preparing a demonstrator of the ICS-D VRE involving data staging, processing in Jupyter Notebooks and visualisation in Enlighten-web[27]. We have selected a use case that was used for testing of the EPOS Data Portal in June 2022. It involves various geoscientific datasets concerning the Vesuvius volcano (Italy):

[27] https://hubtest.envri-fair.eu/sddetail/216

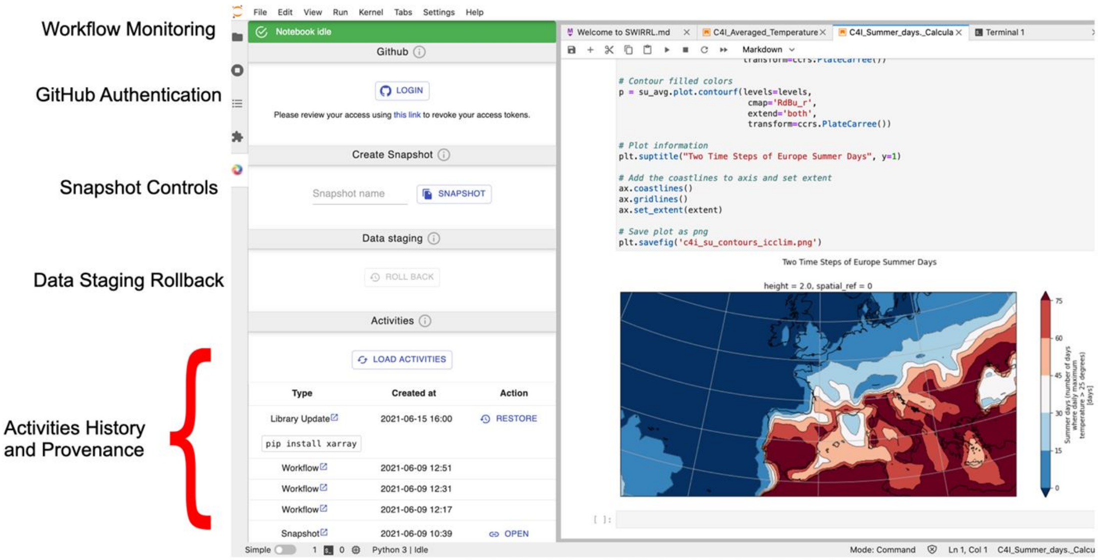


Figure 7 SWIRRL JupyterLab extension: interactive panel with monitoring and user controls to manage reproducibility, traceability and recovery actions.

- Historic and modern earthquakes
- Satellite data
- GNSS data
- European Database of Seismogenic Faults
- Geology maps

The main goal is to demonstrate usage of publicly available data through services (integrated already to EPOS Data Portal) for creating a scientific workflow by combination of various scientific domain datasets. Vesuvius is an active volcano which is observed by many monitoring networks and instruments.

4.3.7.1 Scientific user story

The scientific user story is as follows: “As a researcher, you are investigating Vesuvius volcano (Italy) which is covered by various geoscientific observations. Your main goal is to find available observations/datasets covering the Vesuvius volcano and complete some scientific tasks discovered at the EPOS Data Portal.”

The following tasks are about to be executed:

Task 1: Inspect vertical displacement observed from satellites and compare to trends in GNSS observations from stations near Vesuvius.

Task 2: Explore historical and recent earthquakes in that region and check their relation to known seismogenic faults. Compare it with a geological map.

Task 3: Identify stations for sea level monitoring in the vicinity of the volcano. Explore historical tsunamis in the region and their relation to earthquakes.

1. Satellite observations
Focus on satellite services providing information about vertical displacement.

- Dataset: LOS Displacement Time Series
- Link to data service endpoint: https://catalog.terradue.com/gep-epos/search?format=json&pt=LOS_DISPLACEMENT_TIMESERIES

2. GNSS observations

There are two services to be visited to get the data from surrounding stations and investigate the vertical displacement (Z component).

- Dataset: GNSS Stations with Products
- Link to data service endpoint: https://glass.epos.ubi.pt:8080/GlassFramework/webresources/stations/v2/station/bbox/-25.664/35.60371874069731/27.07/68.0075?with=2
- Dataset: Cleaned GNSS Position Time Series From INGV
- Link to data service endpoint: https://glass.epos.ubi.pt:8080/GlassFramework/webresources/products/timeseries/AQUI/INGV/daily/enu/covjson/?remove_outliers=1&apply_offsets=0&epoch_start=2013-01-01&epoch_end=2021-11-01

3. Earthquakes

Volcanoes are often accompanied by earthquakes. Explore historical earthquakes around the Vesuvius volcano and compare them to modern earthquakes in the same area.

- Dataset: Parameters of historical earthquakes (1000-1899) - FDSN event:
- Link to data service endpoint https://www.emidius.eu/fdsnws/event/1/query?starttime=1000-01-01T00:00:00&endtime=1899-12-31T23:59:59&minlatitude=33&maxlatitude=74&minlongitude=-34&maxlongitude=34&format=xml&orderby=time&includeallorigins=false&includeallmagnitudes=false&limit=300&nodata=204
- Dataset: Parameters of modern earthquakes (1998-present) - FDSN event
- Link to data service endpoint: https://www.seismicportal.eu/fdsnws/event/1/query?minlatitude=30&maxlatitude=90&minlongitude=-40&maxlongitude=50&minmagnitude=5&includeallorigins=false&includearrivals=false&limit=250&format=xml&nodata=204

4. Seismogenic faults

The task is to find and visualise faults in the map.

- Dataset: **European Database of Seismogenic Faults** (OGC WMS)
- Link to data service endpoint: https://services.seismofaults.eu/geoserver/EDSF/ows?service=WMS&version=1.3.0&request=GetMap&layers=EDSF:EDSF&width=768&height=330&crs=EPSG:4326&format=image/png&bbox=30.303092956543,-12.3925075531006,52,45.2800407409668

5. Geological map

Search for a geological map. Navigate to the geological unit that Vesuvius is located in.

- Dataset: Geological Feature View Service (EGDI Geological Map 1:1,000,000)
- Link to data service endpoint: https://data.geoscience.earth/api/wmsGeologicUnit?service=WMS&version=1.3.0&request=GetMap&layers=GeologicUnitView_Lithology&crs=EPSG:4326&format=image/png&width=1536&height=660&bbox=-90,-180,90,180

6. Tsunami observations

Explore sea level monitoring stations close to the Vesuvius volcano.

- Dataset**: IOC Sealevel Monitoring Station List** (as OGC WMS layer)
- Link to data service endpoint: https://geo.vliz.be/geoserver/Sealevels/ows?service=WMS&version=1.3.0&request=GetMap&layers=Sealevels:sealevels&width=768&height=330&srs=EPSG:4326&format=image/png&bbox=-69.00777938015557,-178.1602285632045,82.49000164979999,179.195203583904
- Dataset: **IOC Sealevel Monitoring Station List** (as OGC WFS layer)
- Link to data service endpoint: https://geo.vliz.be/geoserver/Sealevels/ows?service=WFS&version=2.0.0&request=GetFeature&typeName=Sealevels:sealevels&outputFormat=application/json&srsName=EPSG:4326&b

[box=-69.00777938015557,-178.1602285632045,82.49000164979999,179.195203583904](box=-69.00777938015557,-178.1602285632045,82.49000164979999,179.195203583904)
Dataset: **Italian Tsunami Effects Database** - Tsunami History WFS (ITED V1)

- Link to data service endpoint: https://tsunamiarchiveservices.ingv.it/services/tsunami/wfs/wfs?service=WFS&version=2.0.0&request=getFeature&typeNames=tsunami:ITED_v1_localities&outputFormat=json

The seismic data (earthquakes and seismogenic faults) and WMS map layers were processed using the developed data conversion tool and visualised in Enlighten-web tool (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Satellite data and GNSS observations were processed using Jupyter Notebooks and then also visualised together (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Tsunami observations (WMS) together with modern earthquakes are shown in **Fehler! Verweisquelle konnte nicht gefunden werden.**.

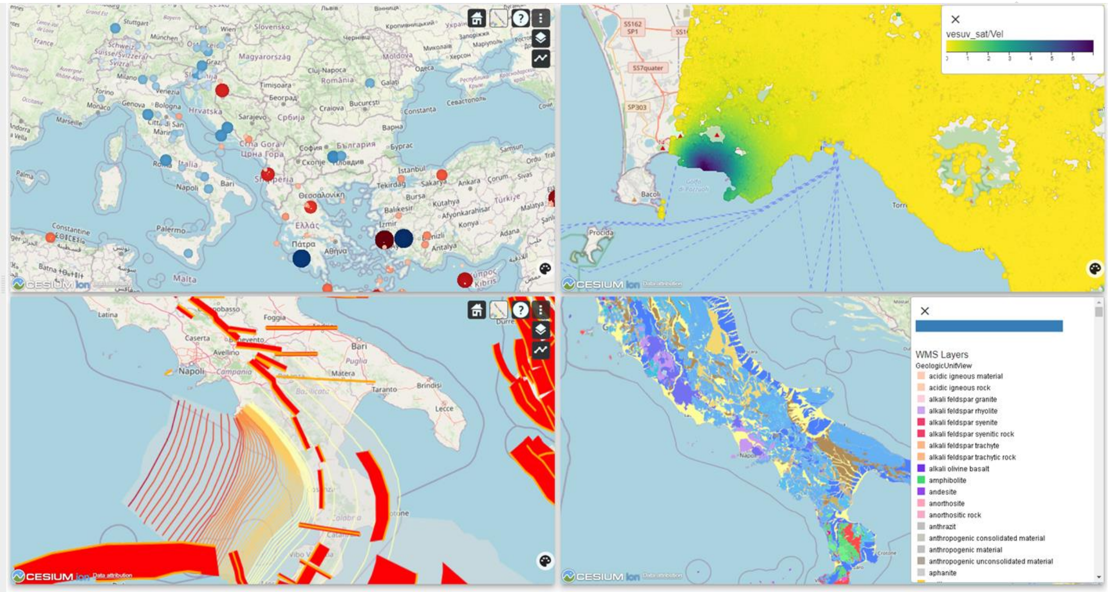


Figure 8 Enlighten-web visualisations of various geoscientific data sets concerning the Vesuvius volcano (Italy). Upper left are historical and modern earthquakes, upper right is average LOS velocity observed from satellite, lower left is seismogenic faults and lower right is geological map with legend.

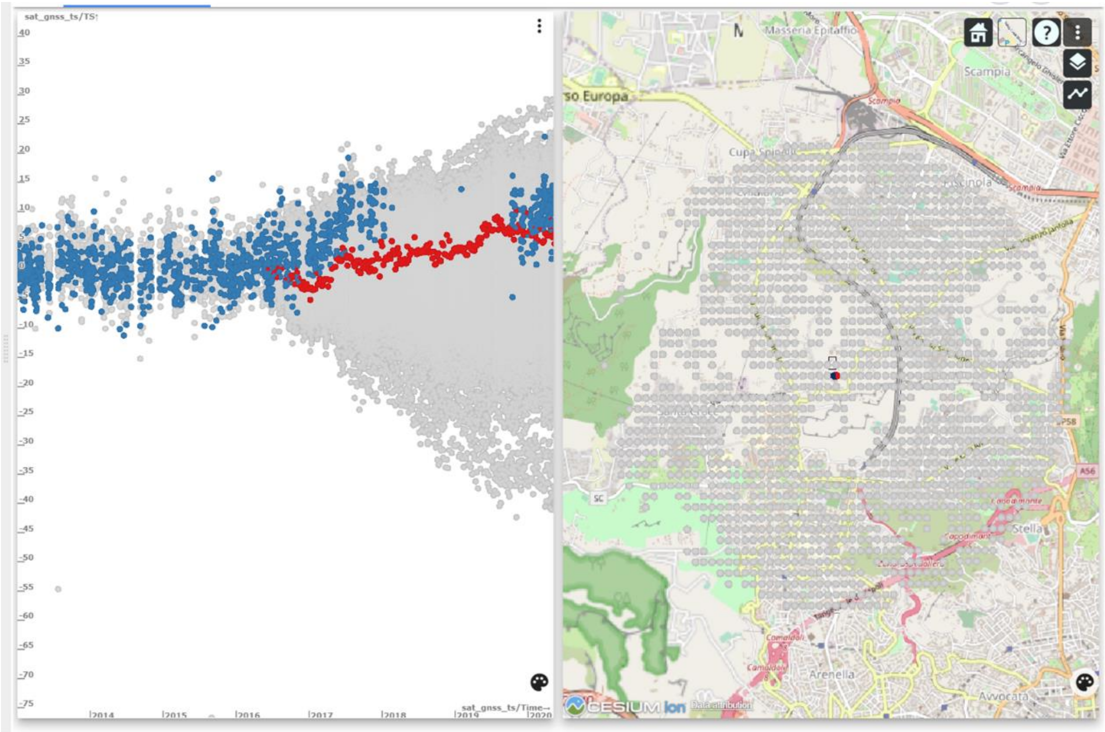


Figure 9 Comparison of vertical displacement from GNSS observations to data from satellite observations. The left scatter plot shows vertical GNSS observations (blue points) together with vertical displacement from satellite observations (red points). The right map plot shows a blue GNSS station and selected spatially closest satellite pixel providing a time series of observations. This spatial filtering allows a quick selection of points of interest. Grey points are satellite data currently not selected.

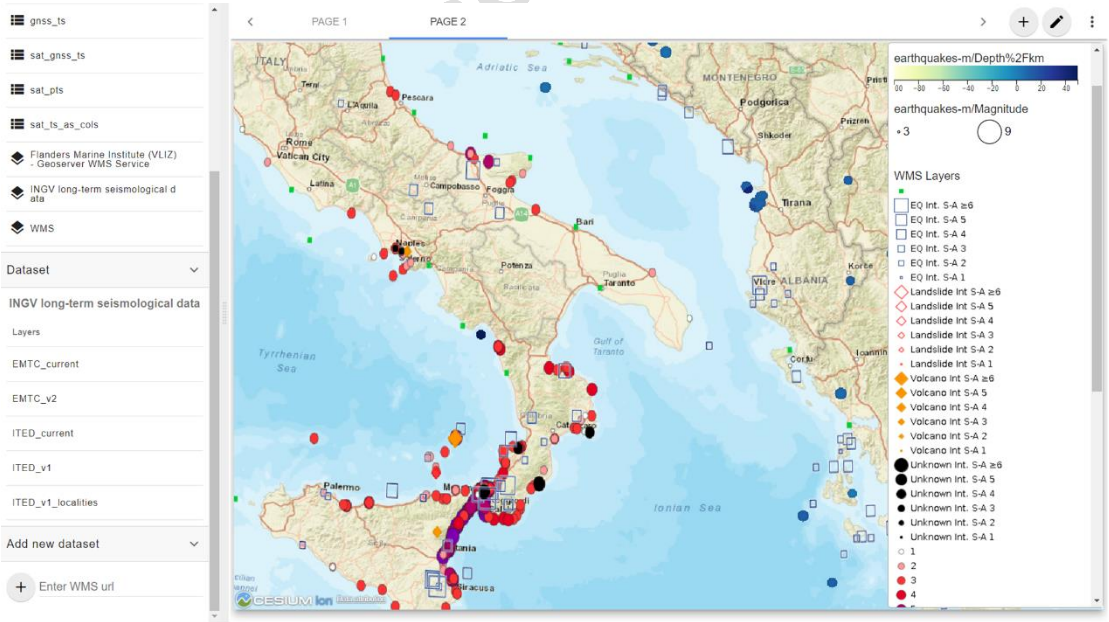


Figure 10 Comparison of modern earthquakes (blue points) with dataset provided by tsunami community (as WMS) in a map view.

5 Task 10.5 – Seismology

This chapter reports the final status of the activities performed in Task 10.5. It provides readers with updates on the results described in D10.4 and introduces additional developments and their corresponding outcomes.

5.1 Objectives

Task 10.5 aims at the definition and implementation of a FAIR roadmap addressing seismological data and products. Building on the results of previous initiative and projects it delivers guidelines and concrete solutions to enable the FAIR principles as common practices in the EPOS seismology community.

5.2 Status of activities and results

In D10.2 [1] we indicated an implementation plan resulted from a FAIR gap analysis. In the project timeframe it was possible to complete only some of the planned activities mainly due to external factors such as the agreement forming process and the heterogeneous level of engagement in the community. For these reasons some envisaged outcomes could not be finalised as operational products. Therefore, we decided to prioritise additional activities beyond the initial plan.

The wide adoption and implementation of FAIR principles in a distributed community with a federated infrastructure is by nature a remarkable endeavour. It requires time and resources that go beyond the scope of a project. In the case of the EPOS seismology community such a task is facilitated by a long-term commitment and established collaborations. However, aligning priorities and roadmaps across different institutions, planning the required capacity, and involving the user community remain challenging tasks.

For these reasons we focused on the definition of a set of recommendations that was produced to help the community identify the issues of FAIR implementation and provide participants with concrete guidance.
Such recommendations are based on the results of the FAIR gap analysis and the knowledge acquired in technical feasibility experiments and proof-of-concepts performed at KNMI. They can be exploited independently from the project to trigger discussions and define development plans.

5.3 Recommendations for FAIR enhancements

The proposed recommendations are based on the analysis of the ORFEUS-EIDA infrastructure and focus primarily on seismic waveform data. The work in the Task Forces of ENVRI-FAIR WP5 has been considered. However, they can be extended to other types of datasets such as seismic station information.

Findability
Key features to enable findability of assets are Persistent Identifiers (PIDs), rich metadata and indexed resources such as catalogues.
The ORFEUS EIDA community offers a portfolio of standard services which includes WFCatalog -- a rich metadata catalogue which contains descriptions of the available seismic waveform data. WFCatalog can be accessed via a HTTP API, its resources are defined as a JSON-schema, and it supports PIDs.

```
"properties": {
        "wfmetadata_id": {
                "description": "Unique identifier of the metadata document. This can
be a DOI, a Handle or any other type of PID",
                "$ref": "#/definitions/stringLiteral",
                "format": "uri"
        },
```

Figure 11 Example of schema definition for PID support

Current issues are that the PID is not populated in all the ORFEUS-EIDA data centres and there is not a shared agreement on the type of PID to use for waveform data and metadata.
The findability recommendations are summarized below:

- Define PID policies and a shared provisioning strategy
- Populate PIDs in WFCatalog
- Consider PIDs for both metadata and data
- Review WFCatalog data model

The work on the WFCatalog data model has been addressed in the project and will be continued.
The assignment of PID to seismic waveform data is necessary to comply with the FAIR principles. Rather than discussing the granularity issue it is important to agree on a common strategy to assign PIDs. Experiments have demonstrated the feasibility with B2Handle (which implements ePIC handles) from the EOSC service catalogue. A possible alternative could be to set up a community specific PID service e.g., based on cool URIs.

Accessibility
The use of standard protocols and mechanisms for user authentication and authorisation are relevant aspects concerning accessibility. The ORFEUS-EIDA infrastructure is quite advanced in this respect and provides users with a broad range of standardised data services together with an integrated AAI system.

To improve the accessibility the following activities are recommended:

- Add support for PID-based queries in current APIs
- Define harmonised policies for long-term metadata management
- Retain metadata upon data decommissioning

Interoperability
A broad adoption of shared knowledge representation methods, the use of vocabularies and qualified metadata are required features to enable interoperability. As an important partner of EPOS, the ORFEUS-EIDA community can benefit from the effort and the resources available in the RI. EPOS has invested a large effort in order to harmonise and represent cross-disciplinary knowledge in the solid Earth sciences. EPOS-DCAT- AP is a rich data model that enables such representation (**Fehler! Verweisquelle konnte nicht gefunden werden.**). A tool was developed to facilitate the creation and management of EPOS-DCAT-AP compliant RDF graphs: the SHAPEness [6] metadata editor. ORFEUS-EIDA resources including seismic waveforms have been described in EPOS-DCAT-AP thus fulfilling most of the interoperability requirements.

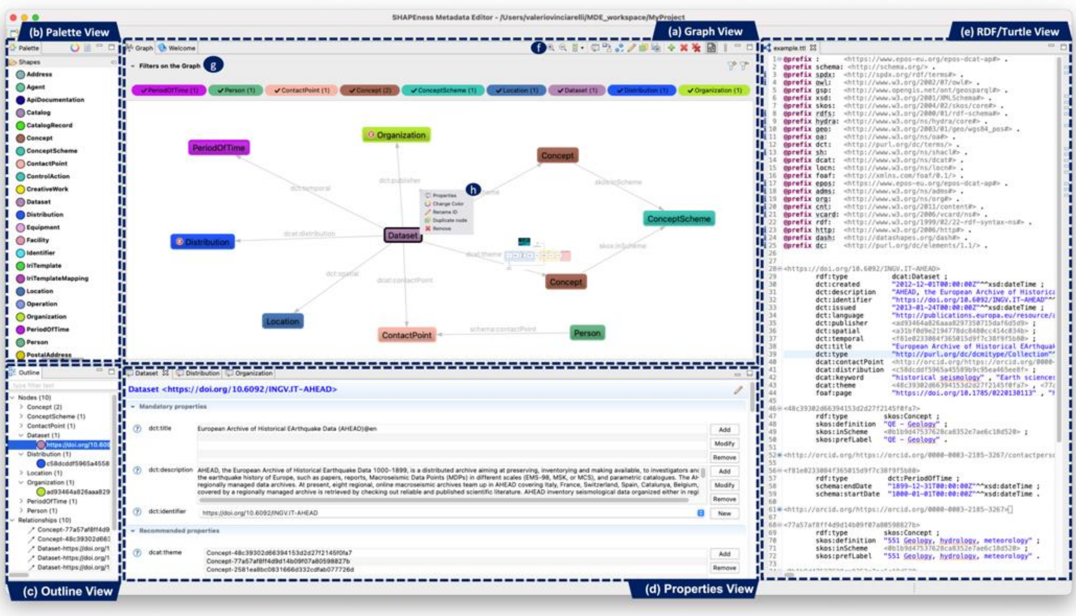
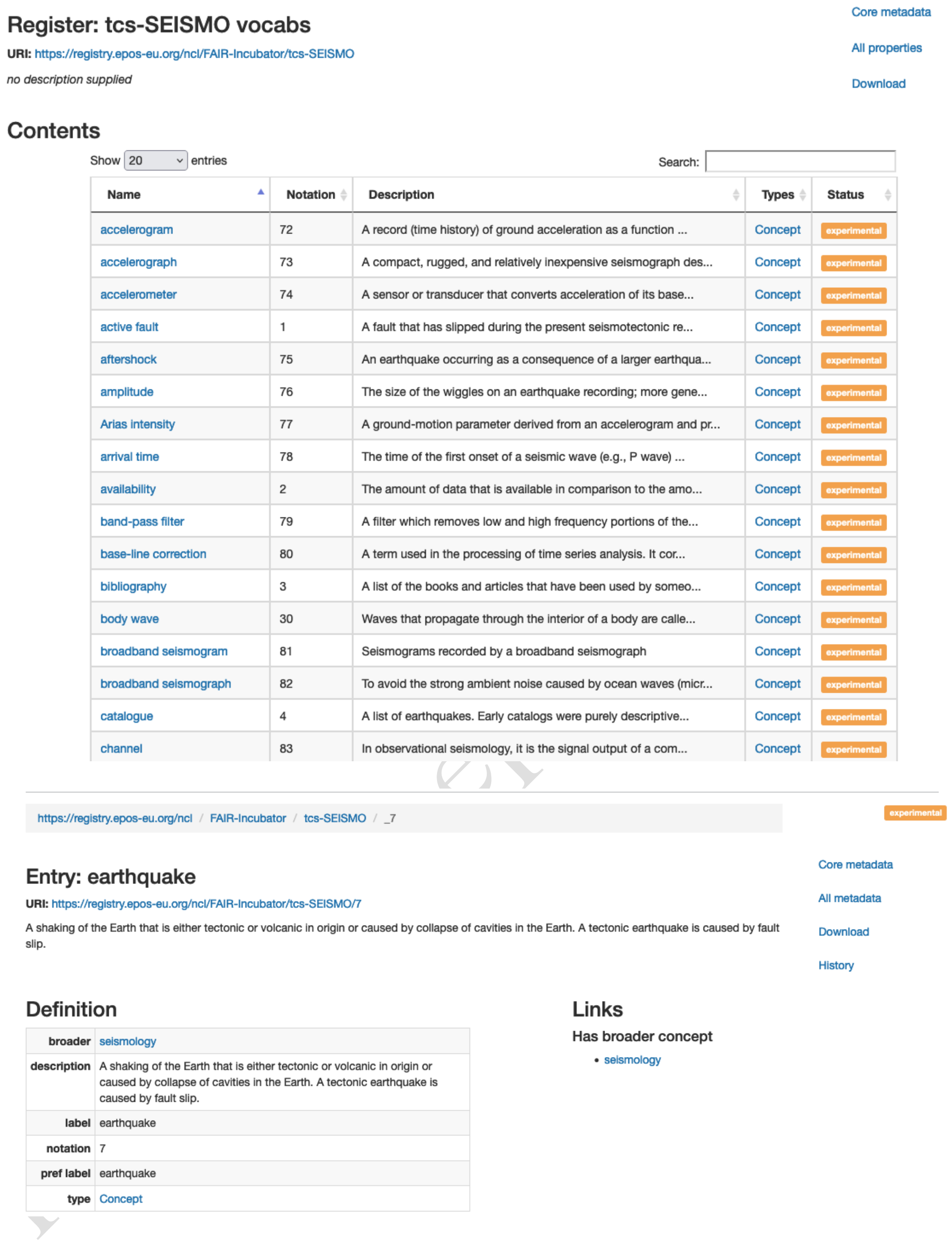


Figure 12 Overview of the SHAPEness user interface



Register: tcs-SEISMO vocabs

URI: https://registry.epos-eu.org/ncl/FAIR-Incubator/tcs-SEISMO

no description supplied

Core metadata
All properties
Download

Contents

Show 20 entries Search:

Name	Notation	Description	Types	Status
accelerogram	72	A record (time history) of ground acceleration as a function ...	Concept	experimental
accelerograph	73	A compact, rugged, and relatively inexpensive seismograph des...	Concept	experimental
accelerometer	74	A sensor or transducer that converts acceleration of its base...	Concept	experimental
active fault	1	A fault that has slipped during the present seismotectonic re...	Concept	experimental
aftershock	75	An earthquake occurring as a consequence of a larger earthqua...	Concept	experimental
amplitude	76	The size of the wiggles on an earthquake recording; more gene...	Concept	experimental
Arias intensity	77	A ground-motion parameter derived from an accelerogram and pr...	Concept	experimental
arrival time	78	The time of the first onset of a seismic wave (e.g., P wave) ...	Concept	experimental
availability	2	The amount of data that is available in comparison to the amo...	Concept	experimental
band-pass filter	79	A filter which removes low and high frequency portions of the...	Concept	experimental
base-line correction	80	A term used in the processing of time series analysis. It cor...	Concept	experimental
bibliography	3	A list of the books and articles that have been used by someo...	Concept	experimental
body wave	30	Waves that propagate through the interior of a body are calle...	Concept	experimental
broadband seismogram	81	Seismograms recorded by a broadband seismograph	Concept	experimental
broadband seismograph	82	To avoid the strong ambient noise caused by ocean waves (micr...	Concept	experimental
catalogue	4	A list of earthquakes. Early catalogs were purely descriptive...	Concept	experimental
channel	83	In observational seismology, it is the signal output of a com...	Concept	experimental

https://registry.epos-eu.org/ncl / FAIR-Incubator / tcs-SEISMO / _7 experimental

Entry: earthquake

URI: https://registry.epos-eu.org/ncl/FAIR-Incubator/tcs-SEISMO/7

A shaking of the Earth that is either tectonic or volcanic in origin or caused by collapse of cavities in the Earth. A tectonic earthquake is caused by fault slip.

Core metadata
All metadata
Download
History

Definition

broader	seismology
description	A shaking of the Earth that is either tectonic or volcanic in origin or caused by collapse of cavities in the Earth. A tectonic earthquake is caused by fault slip.
label	earthquake
notation	7
pref label	earthquake
type	Concept

Links

Has broader concept

- seismology

Figure 13 Overview of the EPOS Vocabulary Registry – https://registry.epos-eu.org

The recommendations for interoperability are the following:

- Enable access to EPOS-DCAT-AP RDF serialisations
- Define a shared seismology vocabulary
- Define governance with roles and responsibilities
- Rollout vocabulary registry

Although resources are encoded in EPOS-DCAT-AP it is still not possible to access RDF serialisations via open endpoints. As previously mentioned, the community vocabulary has been realised and will be made available by EPOS-ERIC.

A harmonised adoption of PIDs can benefit the definition of EPOS-DCAT-AP resources.

Reusability

The reusability requirements focus on the use of rich descriptions, data usage policies, provenance and community standards. Several resources are available in the ORFEUS-EIDA community to fulfil those requirements.

EPOS-DCAT-AP with its rich data model coupled with domain specific standards;

the policies established in EPOS and ORFEUS about licensing, use and reuse of resources; and a recently defined provenance management system used within the SWIRRL web API (M7.8) [5]. In that context, the conceptual work pursued for the representation of the ICS-D in the ICS-C catalogue, besides the underlying technical component of the API, could be used and further refined to address the following.

- Description of computational services in EPOS-DCAT-AP.
    - Computational Workflows and interactive analysis tools have been defined in the EPOS ICS-C catalogue. These could be used to evaluate relationships and metadata which may further refine the EPOS-DCAT-AP representation of the remote computational services, for instance by exploiting and extending the current classification (Concepts) used to represent the services offered by SWIRRL (Workflows, LiterateProgramming, VisualAnalytics), see (**Fehler! Verweisquelle konnte nicht gefunden werden.**)
- Refinements of the provenance templates for seismic workflows and management of the provenance.
    - SWIRRL already generates provenance information associated with workflows. These are very coarse grain in order to be reusable and generic, see (**Fehler! Verweisquelle konnte nicht gefunden werden.**). However, further refinement to the template could capture particular workflow scenarios. The underlying technology based on PROV Template and PROV Template Catalogue [5] and dissemination API, would provide the workflows developers with the needed technical framework to represent, store and expand the templates. Finally, the SWIRRL API can be reused to ingest and disseminate the provenance traces produced at runtime.

Finally, reproducibility is also part of the design of SWIRRL. It implements use cases such as the recovery of the software environment of a running instance of a notebook (e.g. to recover to a previous state of the libraires), and the reproducibility of the analysis, thereby including the data and the methods. The former uses the provenance information to drive the recovery actions, while reproducibility is achieved with the integration of an internal data-versioning system, and direct connection to Git. This is used to generate snapshots that can be shared and reproduced on a target computational platform based on MyBinder[28]. SWIRRL API's services and functionalities are discussed in the context of a particular seismological use case, to assist scientists with answering the question "Do Earthquakes Occur Only in the Crust or Also Below Moho?"

5.4 Application of the recommendations to other datasets

As previously mentioned, the recommendations are based on a specific dataset, but they can be of broader application. In the seismology community another dataset of primary importance is the information about seismic sensor networks. It includes details such as naming of network, seismic station, channel, geographical coordinates of the deployment, technical characteristics of the sensor, information about the provider and more. This is typically known in the community as *station metadata*.

PIDs, in particular DOIs are currently applied to this type of dataset. However, their granularity is coarse as they identify a full network which might include hundreds of instruments. In the light of current developments which look at the inclusion of new types of instruments such as Distributed Acoustic Sensing (DAS), an interesting extension could be the application of PIDs to single instruments. To achieve this goal EOSC services could be exploited such as the new FAIR instrument service based on B2INST (DOI:10.5281/zenodo.6247512). Such an extension would require development effort to adjust the data model, databases and current APIs serving station metadata i.e., FDSN StationXML.

The recommendations can help identify required changes and guide developments.

[28] https://mybinder.org/

5.5 Additional developments (beyond implementation plan)

Management of Seismic Station Metadata (Seismic Inventory)

To improve the FAIRness of seismological metadata a system was designed which enables the management of metadata describing seismic networks of sensors. The application enables operator to register new sensors, modify existing information and publish changes in seismic metadata. The application was developed, tested and released at KNMI. It was also presented to the ORFEUS-EIDA community and received positive feedback and interest by other partners. Future work will include the integration of additional features such as full provenance support and adaptations to facilitate deployment and configuration at other EIDA nodes.

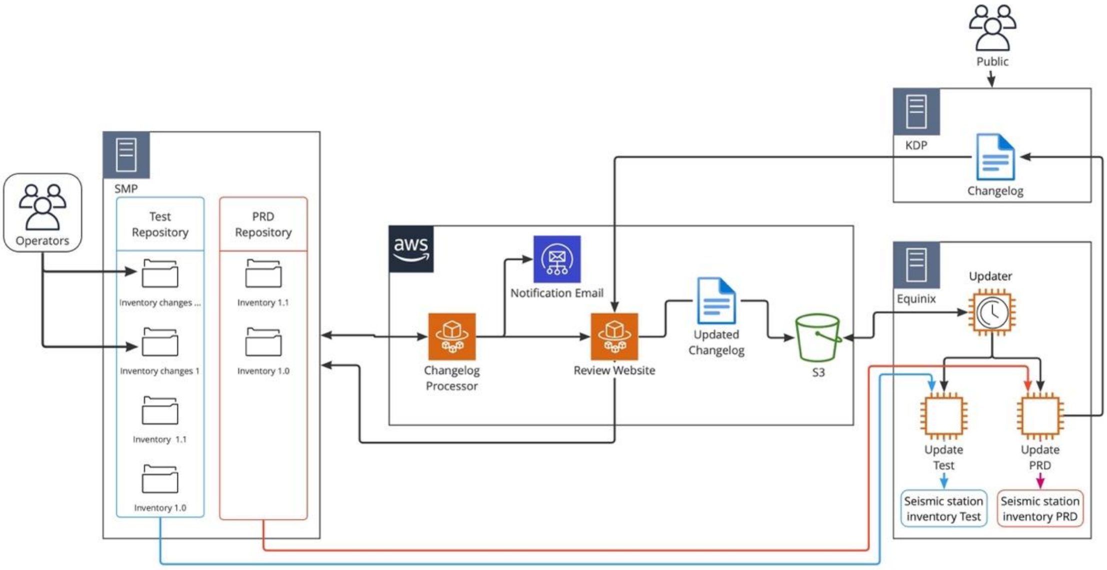


Figure 14 High Level Design of the ChangeLog Application – Data centre operators can modify seismic inventory information in the Station Metadata Portal (SMP). Inventory changes can be marked as release and sent for review. Approved releases can be automatically deployed in the operational environment. Inventory changes stored and made available for public consultation on the KNMI Data Platform (KDP).

```
▼ 2:
  ▼ Summary:
    ▼ Description:        "Added response to infrasound arrays DBN and EXL, and state-of-health channels to BAPP, G010, G710, VNDM0"
    ▼ Motivation:         "Sensor update in infrasound arrays and additional SOH channels"
      Releasedate:        "2023-04-11T14:02:18"
  ▶ Inventory-fileformat: "http://geofon.gfz-potsda…ns/seiscomp3-schema/0.10"
  ▼ Changed:
    ▼ NL:
      ▼ BAPP:
        ▼ New Channel:
            0:            "NL.BAPP..CPU"
            1:            "NL.BAPP..DEG"
            2:            "NL.BAPP..HUM"
            3:            "NL.BAPP..LCE"
            4:            "NL.BAPP..LCQ"
            5:            "NL.BAPP..VEC"
            6:            "NL.BAPP..VEP"
            7:            "NL.BAPP..VVB"
      ▼ DBN01:
        ▼ channel-2.sensor.description:
          ▼ Old:
              channel-name:   ".01.HDF"
              value:          "KNMI microbarometer"
          ▼ New:
              channel-name:   ".01.HDF"
              value:          "KNMI 500s differential microbarometer"
        ▼ New:
          ▼ 0:
              0:              "_pz_transfer_function_type"
          ▼ 1:
              0:              "_cf_transfer_function_type"
        ▼ channel-2.extra.format.value:
          ▼ Old:
              channel-name:   ".01.HDF"
              value:          "Steim2"
          ▼ New:
              channel-name:   ".01.HDF"
              value:          ""
```


Figure 15 Example of ChangeLog output in JSON format. A simple text format is also available

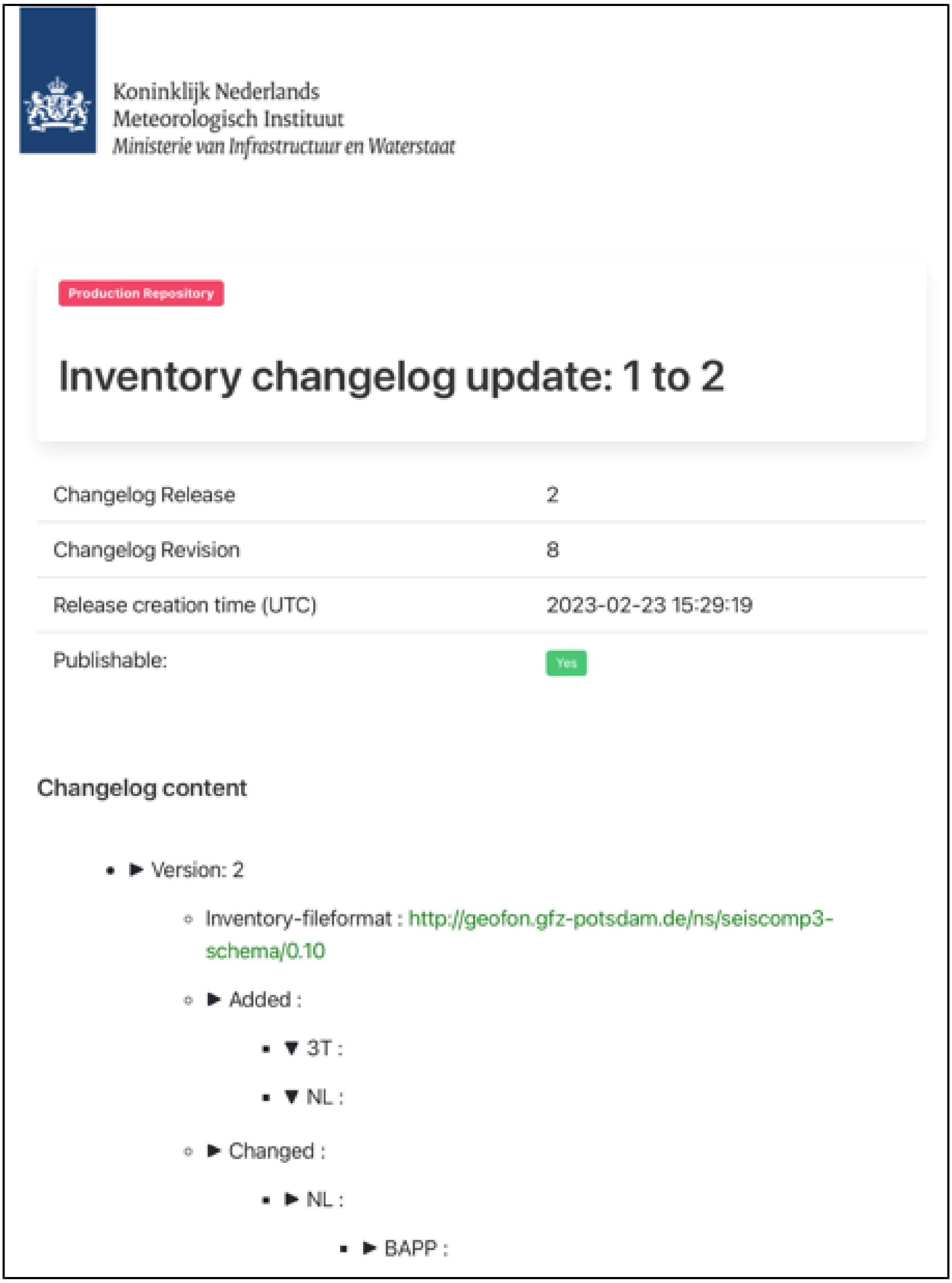


Figure 16 Screenshot of the review page. An operator can check the content of the ChangeLog and approve it for publication

6 Task 10.6 - Satellite Earth Observation community

6.1 Introduction

This chapter describes the activities developed in Task 10.6 relevant to the implementation of the FAIR principles for the Earth Observation products of the EPOS TCS Satellite Data. The goal of the task was to develop awareness and share knowledge about FAIRness in the satellite Earth Observation domain to foster the application of the FAIR principles into existing practices and methods. This task has an effective link with the results of previous and ongoing activities carried out in the context of other projects such as EOSC-hub, EPOS-IP, EPOS-SP and OpenAIRE-Advance[29]. Last but not least, the task had to tackle the integration with the Copernicus DIAS[30] environments.

To produce an efficient framework to address the FAIR aspects in the different phases of the product lifecycle such as generation, curation, computation, dissemination and publication, the task addressed the following implementation activities:

- FAIRness gap analysis and implementation;
- Integration and exploitation of DIAS computational services;
- Enhancement of AAAI systems for Earth Observation RIs.

6.2 Gap analysis of FAIR principles in the EPOS TCS Satellite Data

The FAIRness gap analysis was developed by considering the Data, Data Products, Software and Services (DDSS) which are provided by the EPOS TCS Satellite Data (SATD) community and that were validated and included in the EPOS products portfolio. In particular, we focused on the DDSS referred to as “Line of sight displacement time series”. This DDSS is the most complex and complete among all the services deployed by the TCS SATD and it contains all the main FAIR issues of the TCS. Accordingly, this DDSS represents the touchstone of the TCS SATD FAIRness maturity.

The analysed DDSS is released with two components:

- The data, embedded in a csv (ASCII) file, represented a sparse matrix of points located on the Earth surface. Each point is represented by a coordinate triplet (latitude, longitude and altitude), a predetermined number of parameters coming from the data processing, and a time-series of displacement values. The time-series does not have a fixed size, i.e., it can change dataset by dataset, because the number of samples depends on the number of satellite acquisitions used (processed) to generate the final product.
- The metadata, formatted in a xml file, represented according to the ISO19115[31] standard.

The gap analysis was carried out by critically analysing the several components of the TCS SATD with respect to the requirements of the FAIR principles. Such an analysis identified the following gaps that impact on different FAIR principles (Table 2).

Table 2: Gaps Related to FAIR Principles

Issue	Description	Impact on
PID	No PID system is used and applied	A1, F1, F3
Data and Metadata Life	No policy has been agreed on the lifecycle of data and metadata. Metadata preservation is not guaranteed.	A2
Vocabulary	A preliminary vocabulary has been drafted but the activity is not complete	I2

[29] https://www.openaire.eu/advance/

[30] https://www.copernicus.eu/en/access-data/dias

[31] https://www.iso.org/standard/53798.html

On the basis of the gap analysis, a list of activities to be carried out has been drawn up and jointly agreed to form a roadmap for their implementation.

6.2.1 Application of a PID system

This is a long-term activity since it does not have a trivial solution. The adoption of a PID system has a strong impact on the financial and technical sustainability of the RIs. Indeed, a large number of satellite DDSS are live products, because they are regularly updated each time a new satellite image is available (in same cases several acquisitions per week are available). The preservation of all generated products (instead to have a single product regularly updated) could become unmanageable because each product is quite large (up to several gigabytes) and the disk space would exponentially increase. However, the metadata of previous versions could be preserved to provide appropriate provenance. Moreover, even if several PID systems are available (e.g., DOI and Handle) the services related to PID management have a cost that has to be supported by the RIs. A PID system suitable for the satellite community has yet to be found, also investigating its impact on the financial sustainability.
The activities to adopt a PID system within the satellite EO community started by analysing guidelines, best practices, and solutions developed by Space Agencies and other research groups. In this context, particularly useful has been the work done by Data Stewardship Interest Group (DSIG)[32] (formerly known as the WGISS Archive Task Team) of the CEOS-WGISS working group[33]. This team deals with data archiving as well as data and associated knowledge consolidation and valorisation aspects.
The DSIG recently released a document entitled Persistent Identifiers Best Practice[34] focused on providing recommendations and best practices on the use of Persistent Identifiers to Earth Observation mission data, allowing globally unique, unambiguous, and permanent identification of a digital object. In particular, the report addresses several use case scenarios, tailored for the Earth Observation community, some of which perfectly match with the EPOS TCS SATD context. The implementation activities are going ahead by building some effective pilots for the PIDs in the satellite community.

6.2.2 Data and Metadata Lifecycles

This is a short- to medium-term activity; the metadata preservation has no strong impact on sustainability and is technically manageable. It needs an agreement at RIs level to be implemented in the Data Management Plan.

6.2.3 Vocabulary

This is a dynamic action. The building and updating of the vocabulary is an ongoing activity. Once the vocabulary is consolidated, it will be documented and resolvable using globally unique and persistent identifiers.

6.3 Implementation activities and outcomes

6.3.1 PID system

The main purpose of a persistent identifier is to help the users to cite and find specific datasets. The main goals that the labelling process should achieve are:

- Globally unique, unambiguous and permanent identification of a digital object for locating and accessing over time.
- Improve discoverability and accessibility.
- Enable users to retrieve objects without knowing their location.
- Enable repositories to change the location of objects internally.
- Enable repositories to share objects with other services where appropriate.

[32] https://ceos.org/ourwork/workinggroups/wgiss/preservation/
[33] https://ceos.org/ourwork/workinggroups/wgiss/
[34] http://ceos.org/document_management/Working_Groups/WGISS/Interest_Groups/Data_Stewardship/White_Papers/WGISS_DSIG_Data%20Stewardship%20Reference%20Model%20White%20Paper_v1.0.docx

- Enable researchers to cite digital objects consistently over time, which also benefits data holders.
- Increase data visibility and use.
- Increase credibility and value of data holdings.

The EPOS Thematic Core Service Satellite Data (SATD) is developing a process to univocal identify datasets provided through the TCS platform by using a persistent identifier for each data resource. As a persistent identifier, we have chosen the DOI (Digital Object Identifier) because of its structure that guarantees persistence and interoperability at the software level. To obtain a DOI service the user has to contact a DOI Registration Agency, an authority recognised by the International DOI Foundation (IDF). A Registration Agency allocates DOI prefixes, registers DOIs and provides the necessary infrastructure to allow registrants to declare and maintain metadata and state data (Figure 17) for a PID functionality scheme description).

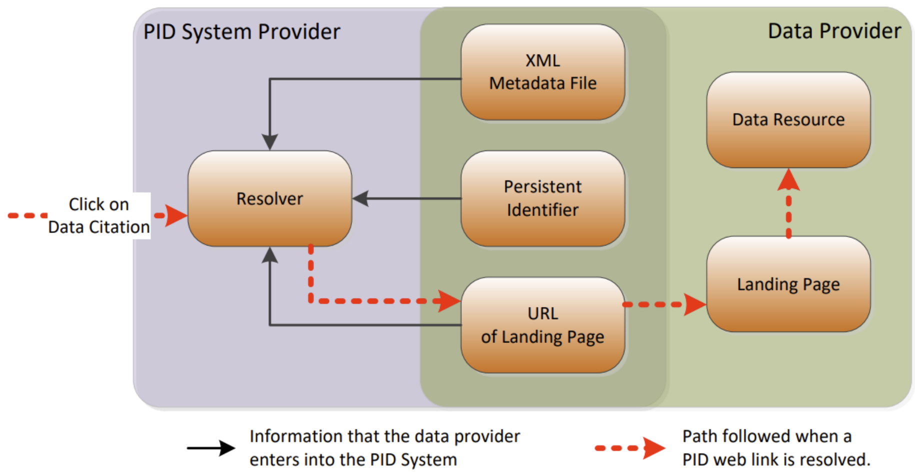


Figure 17 Schematic description of PID functionality process.

After a careful analysis of the available Registration Agencies, our choice fell on DataCite since the agency's mission is mainly to offer a service for the generation of the persistent identified for datasets. In addition, DataCite provides a metadata structure suitable for our dataset and REST API for retrieving, querying and browsing DataCite DOI metadata records.
DataCite has created an account for CNR in the DataCite DOI Fabrica test system (https://doi.test.datacite.org). As "Member", we needed to set up a "Repository". Repositories use Fabrica to create and manage DOIs and their associated metadata. Prefix to create DOI has to be requested to support@datacite.org. DOIs can be registered using DataCite's web interface Fabrica or one of DataCite APIs.
Test accounts have the same capabilities as production accounts, but none of the DOIs created during the test period is permanent or resolvable. The same account credentials work for the test instance of Fabrica and the test instances of DataCite APIs:
* REST API (https://api.test.datacite.org)
* MDS API (https://mds.test.datacite.org)
CNR has implemented a set of python codes to automatically generate a DOI name to be associated with the dataset, before uploading the data in the TCS data catalogue. To complete the PID creation process, CNR is building a landing page generation service. A landing page is a web page with information about the data and a download link. For the SATD community, this service will be hosted by the SATD website (under construction). When a DOI is registered by DataCite, XML metadata file is generated. Clicking on a DOI citation, the resolver (https://doi.org/) redirects the user to the landing page (Figure 17) for the whole DOI assignment process).
The main issue faced in developing the automatic process for DOI application is related to DataCite documentation that is not always adequate. In particular, the generation of JSON files format which manages the information exchange between the data provider (CNR in this case) and DataCite server is poorly documented. On the DataCite website (https://support.datacite.org/docs/api-create-dois) there is a note that states possible problems creating/updating DOI using DataCite JSON and DataCite informs users that they are going to solve the problems within this year. The DataCite support has always been punctual in replying to the emails, but unfortunately not conclusive.

However, we were able to obtain the correct JSON format to update DOI metadata information, by using the JSON format generated from the following webpage: https://support.datacite.org/reference/get_dois Once retrieved the correct JSON format, we used that scheme to update the DataCite server with the relevant metadata information for the dataset.
Finally, in order to evaluate the sustainability of the DataCite solution, its cost assessment has been carried out. The cost for DataCite membership is 2,000 euros per year, that allows users to create and manage DOIs. There is an additional fee for DOI services. More details can be found at the link https://datacite.org/feemodel.html.

6.3.2 AAAI system of Earth Observation RIs

The enhancement of the AAAI system of Earth Observation RIs has been successfully tackled. The TCS SATD has a unique interface with the EPOS ICS-C represented by the Geohazards Exploitation Platform (GEP)[35], a cloud-based platform developed with the support of European Space Agency (ESA)[36]. GEP is an interoperable platform that is queried by EPOS ICS to retrieve data and metadata; the RIs of the TCS SATD are integrated within GEP and their products and services can be retrieved through GEP. The AAAI system of GEP has been successfully integrated with ICS and users registered within the EPOS management system can easily access the TCS resources. It is worth noting that the TCS AAAI is integrated with other systems largely used in the satellite EO community, such as ESA and EOSC AAAI systems (Figure 18)

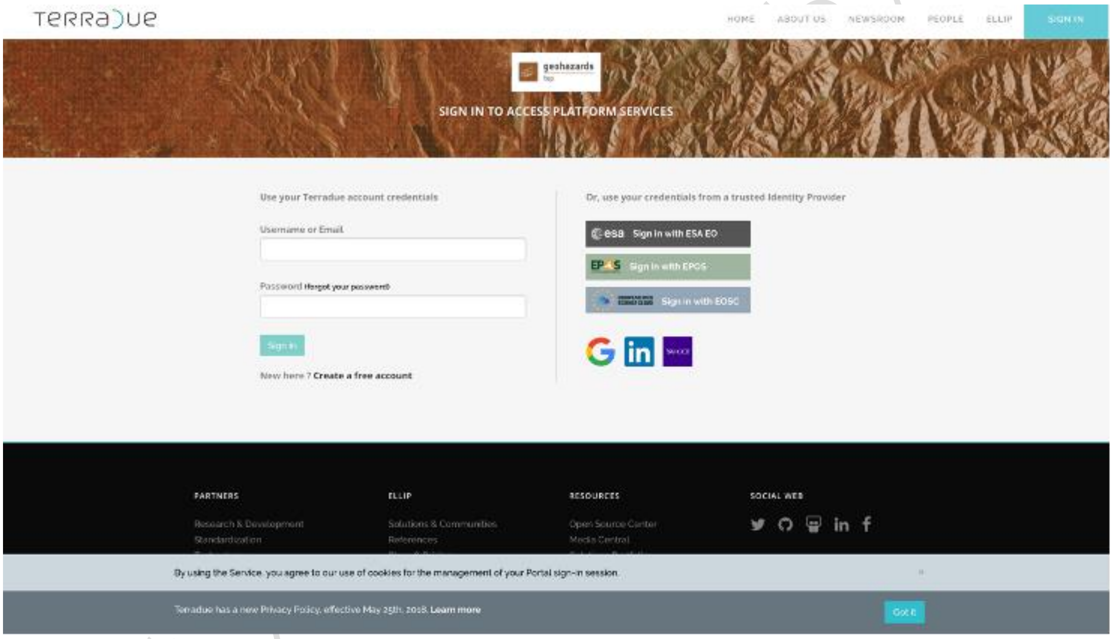


Figure 18 Login Page of Geohazards Exploration Platform

6.3.3 DIAS computational services

In addition to the above-mentioned activities, the task 10.6 had to address the integration with and exploitation of DIAS and EOSC computational services.
At the moment, the EOSC initiative cannot be considered an effective solution for the TCS SATD needs. Indeed, EOSC does not provide a direct access to the Sentinel-1 data archives; this limit represents an insurmountable obstacle considering the Sentinel-1 data amount to be transferred for an advanced interferometric analysis. Anyway, EOSC represents the most important initiative in the cloud computing field currently developed in Europe and in the last years, the TCS SATD partners have contributed to the EOSC development through several H2020 projects. Accordingly, in the next years the TCS SATD

35 https://geohazards-tep.eu/#!
36 https://www.esa.int/

intends to fully benefit from the potentialities offered by the EOSC galaxy when it will represent a competitive alternative with respect to other services providers.

The integration of satellite RIs with Copernicus DIAS computational services represents a challenging activity. In particular, DIAS, the cloud providers specifically selected and developed with the support of ESA and DG-GROW for Copernicus users, seems to be the more suitable environment where satellite processing services can be effectively implemented. Unfortunately, in order to efficiently exploit such an environment, RIs needs to thoroughly investigate the technical and financial sustainability, i.e., the suitability and robustness of the offered services, as well as their medium- to long-term competitiveness with respect to similar solutions deployed by the commercial sector, have to be accurately evaluated. In this framework, the analysis of the DIAS solutions has been carried out. CNR integrated its processing chains on the virtual environment provided by Onda[37] and CreoDIAS (two of the DIAS providers). Various experiments were carried out to evaluate cloud providers technical performance and financial effort, by performing use cases based on the exploitation of these environments for the creation of operational services in the Earth Observation field.

CNR, at the moment, has identified ONDA platform as the best DIAS solution for the EPOSAR service purposes. ONDA, indeed, provides a fast access to the Sentinel-1 data archives over the whole European territory and HPC computing facilities able to perform interferometric processing in a limited time frame. ONDA does not provide a portfolio of cloud services and computing resources comparable with other private competitors, even if it is able to offer a high number of possible configurations. On the other hand, ONDA offers the opportunity to purchase CC resources with a reserved configuration, with competitive costs and payment configurations (e.g., on-demand, monthly, yearly).

Within ONDA environment, CNR have developed a national ground service for the generation of the ground displacement time series over the entire Italian territory.

Even if ONDA has demonstrated to be suitable for the TCS scopes, there are several shortcomings that have to be considered. In particular, ONDA-DIAS does not guarantee a world-wide data availability. Moreover, European Commission and ESA have a medium-term plan to support the DIAS initiative, thus implying a crucial criticality for the long-term sustainability. On the other hand, we have to consider that the European Commission is continuing to invest in the launch of new SAR satellite constellations (e.g., Sentinel-1C/1D and ROSE-L), and new investments in the Satellite Ground Segment sector are very likely.

As evident from the description above, in this complex and unpredictable scenario, the TCS SATD has decided to diversify the investment on different CC solutions. Indeed, with this approach, we believe that the sustainability plan of the TCS SATD can be more robust and reactive with respect to the market and institutional changes.

[37] https://www.onda-dias.eu/cms/

7 Task 10.7 - Marine

This chapter describes the state for T10.7 covering the solid-earth aspects of the marine domain. The prime objective of task 10.7 is to increase the still limited interoperability of geophysical data/metadata of EMSO ERIC[38] and EPOS. The adoption of common standards for data and metadata will support scientists in the joint use of land and marine data. The task gives a particular focus to the enrichment of metadata of data and sensors of Ocean Bottom Seismometers (velocity-meter and accelerometers) / Hydrophones / Magnetometers (OBS/H/M), portable modules used to set-up networks in marine areas unreachable by land networks or to extend land networks out into marine areas.

Implementation activities include engagements with stakeholders involved in the data pipeline, as well as enriching metadata to improve data documentation and ensure their re-use. A shared workflow (data curation, long-term preservation) that was not ensured across EPOS and EMSO ERIC, is being analysed and agreed for OBS/H/M data to make them findable and accessible over the long term. This task is expected to improve the range of data products and the adoption of FAIR principles over the data pipeline.

7.1 State of FAIR implementation

EMSO ERIC regional facilities do have an essential role in delivering seismic data to broad seismological and geophysical communities, to national agencies and other stakeholders. In fact, because of the still uneven geographical distribution of EMSO marine facilities[39], regional scale seismological and geophysical studies of the Mediterranean and North-East Atlantic shall become possible from the regular joint use of EPOS and EMSO data. Seismological data at several EMSO regional facilities have been utilised where standardised seismological data flow to national data centres with links to civil protection and to international seismological agencies exist. This integration helps to improve the reliability of the localisation of the seismicity, especially those events occurring in marine coastal and open sea areas.

The data workflow includes interactions with EPOS through ORFEUS-EIDA. However, the data workflow implemented by the EMSO across regional facilities is not standardised yet. Further, the level of adoption of FAIR principles varies depending on the regional facility. These interactions with EPOS require an extensive standardisation of the acquisition and validation process from the sensor level to the recorded metadata, including data format, data transmission protocols, data archiving platforms in order to be effective in interoperation. Additional service components are based on dedicated methods and software for integrating data from the different observation systems of the facilities and retrieving the basic standard earthquake parameters.

In order to improve the interoperability between EMSO and EPOS, it is crucial for EMSO to develop a harmonisation abstraction layer (Figure 19). We investigated the process and key features for such a harmonisation based on the FAIR principles. A summarised assessment is provided below.

[38] http://emso.eu/
[39] http://emso.eu/observatories/#map

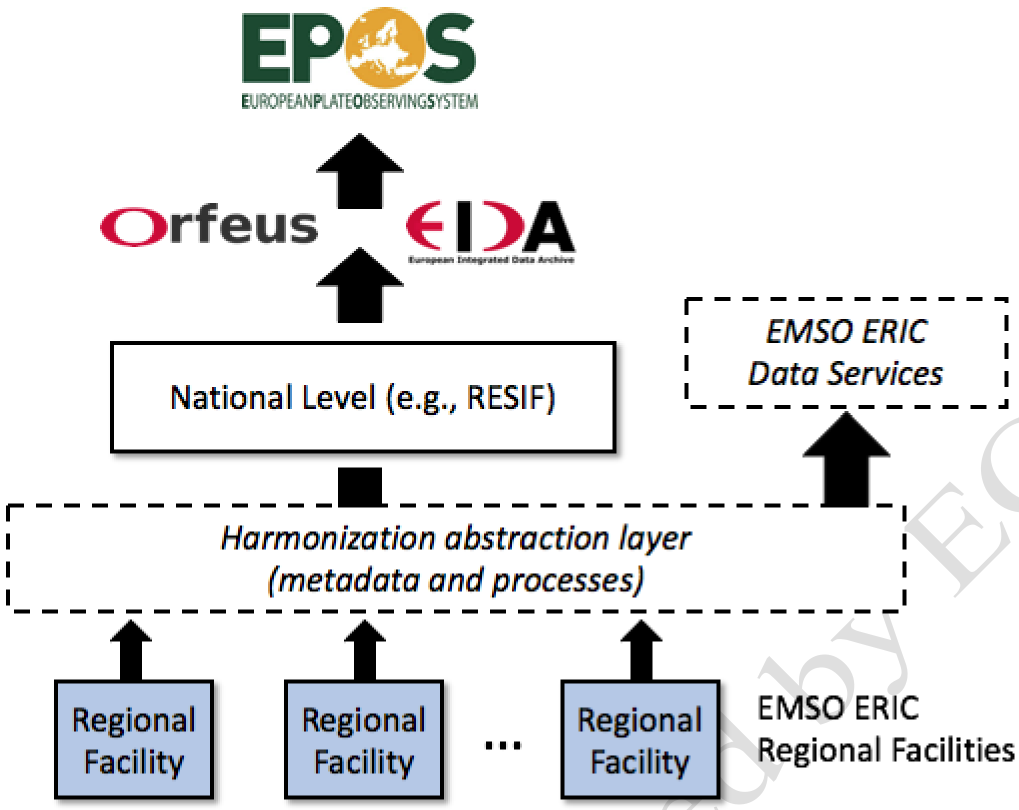


Figure 19 Seismic Data Flows from EMSO to EPOS with Proposed Harmonisation Layer

7.1.1 Findable

While in EMSO some data is already findable through standardised mechanisms (e.g., national agencies and ORFEUS-EIDA), not all regional facilities use the same data discovery mechanism. Our goal is to establish the most effective tools to harmonise and enrich metadata to enable discovery and integration into EPOS while providing visibility of EMSO contributions to EPOS. To this end, we have developed tools to create metadata and data compatible with EIDA data centres. These tools are designed to also apply to other marine continuous-sampling instrumentation, such as tide gauges and magnetotelluric instruments. In May 2023, members of 10 European marine seismology and acoustics facilities participated in an ENVRI-FAIR training course on this tool. We also developed a data/metadata preparation node for marine data (EPOS France A-node), necessary for quality control before delivery to an EIDA data centre. The node software will be made available online through a GitLab site: finished components already are there.

7.1.2 Accessible

Current access to EMSO seismological data includes different mechanisms such as those from RESIF (French seismological network)[40] and EIDA to distribute data and metadata according to the standards of the International Federation of Digital Seismograph Networks (FDSN), central control systems or specialised applications. The harmonisation of mechanisms is essential for accessibility. Current efforts are focused on better integration with ORFEUS-EIDA and harmonised access interfaces across EMSO regional facilities.

[40] http://seismology.resif.fr/

7.1.3 Interoperable

While EPOS represents the front-end in the primary sub-domain, an essential activity within the EMSO back-end is the standardisation of processes to ensure interoperability between regional facilities and with other key stakeholders such as EPOS. EMSO will investigate the adoption of standardised seismological vocabularies at the harmonisation abstraction layer as they are established.

7.1.4 Reusable

Harmonised and enriched metadata are expected to improve current documentation processes and ensure re-use. A significant challenge is establishing an agreed workflow between EMSO and EPOS (ORFEUS-EIDA), which is necessary to enhance the data curation process. The A-node software includes specific tools for including marine seismology-specific information (with respect to issues such as timing and location) in the EIDA-standard metadata and the group works with EIDA and FDSN to suggest modifications to this metadata standard to integrate these and other information natively.

The implementation of technical activities for the adoption of FAIR principles is preceded by an analysis of available EMSO sources of seismic-related data. In this analysis, we identified:

- regional facilities delivering seismic data;
- types of data produced at each of the regional facilities;
- current processes for providing data and metadata, including ongoing interactions with ORFEUS-EIDA

Additionally, regional facilities currently not interacting with ORFEUS-EIDA have explored the requirements and interfaces for engaging with ORFEUS-EIDA.

Based on the FAIR assessment and gap analysis, our roadmap for technical activities and implementation has focussed on enriching metadata and establishing an agreed workflow with ORFEUS-EIDA to enhance the integration and improve the visibility of EMSO contributions through its regional facilities.

8 Task 10.8 - Policies

UKRI has been particularly active in this task, relating it to the work of the policy group within EPOS and to the work of WP4 in ENVRI-FAIR. A comprehensive set of policies has been developed, presented to SCC (Services Coordination Committee of EPOS) and modified following suggestions. The v2 policies are now accepted and consolidated into a policy document. In parallel, the EPOS portal has user consents for data privacy, terms and conditions and cookies with policies attached for ease of reference.

Outside of WP10 activity, but closely linked with it, in WP4 UKRI has produced deliverables (D4.4, D4.6, D4.7) and - after participating as presenter in workshops 1 and 2 run by UHEL - has presented workshops 3 and 4. UKRI also contributed to the other WP4 deliverables, maintaining the link between WP10 and the generic ENVRI policy work. The EPOS policy experience has been used as examples and templates to push forward the WP4 work.

9 Conclusion

EPOS started in ENVRI-FAIR with an architecture designed already for FAIR since EPOS had been represented in the Force-11 discussions on principles and subsequently in the RDA FAIR Data Maturity Model Working Group[41] discussing criteria for assessing FAIRness. EPOS has also been in communication with projects such as GO-FAIR and FAIR's FAIR and more recently FAIR-IMPACT.

Furthermore, EPOS initiated the metadata catalogue using CERIF with services as the content in order to intercept the developing strategy for catalogues in EOSC. OF course, the catalogue may also contain metadata on datasets, data products, software and other assets.

This deliverable describes the outcomes from the detailed implementation pathways (defined in D10.2, prioritised in D10.3 and with progress reported in D10.4) followed for the various parts of the EPOS – and EMSO – environments to achieve improved FAIRness for end-users. Starting from a base that was already FAIR, the improvements are to allow for more sophisticated FAIR implementation, moving towards autonomic mechanisms and not relying solely on human interventions.

10 Impact on the Project

Having identified gaps in FAIRness in T10.1 (D10.1) [2] and constructed an implementation plan in T10.2 (D10.2) [1] the state of implementation following the more detailed plan for implementation is provided in this deliverable.

It provides information for other project members on how EPOS and EMSO (solid earth subdomain) have improved FAIRness and provides a basis for discussions on best practice. Moreover, if all subdomains harmonise to a rich level of FAIRness – such as that already achieved and being improved within the solid-earth domain - then discovery, contextualisation, access, interoperability and re-use across the RIs of ENVRI becomes much easier. Furthermore, such rich metadata provision improves the catalogue of ENVRI-Hub, fed from the catalogues of the different sub-domains/RIs of ENVRI.

The ENVRI-FAIR Work Package 5 (WP5) Task force 1 (TF1) (catalogue) has already endorsed the approach of building a rich metadata catalogue for ENVRI-Hub using the EPOS technology. The WP5 TF2 (AAAI) is evolving towards harmonisation in parallel with the EPOS approach. Nonetheless, other TFs are proposing an approach to metadata based on triple stores, Resource Description Framework (RDF)[42] and SPARQL for the catalogue, or even eschewing a central catalogue of metadata. The EPOS point of view is that – for the metadata catalogue - n-tuples are better than triples for expressivity (including temporal and – potentially - modal aspects), formality (referential and functional integrity) and performance. However, (a) in parts of EPOS e.g., the prototype work on ICS-D in section (Task 10.4 - ICS-D) work on triple stores is ongoing – in this case related to provenance; (b) in the VRE4EIC project the EPOS team demonstrated conversion of CERIF to RDF for the purposes of offering interoperability to systems choosing to use triple stores instead of n-tuples.

FAIRness is a journey, and development work continues even after ENVRI-FAIR.

11 Impact on Stakeholders

Stakeholders of the solid earth subdomain – researchers, government agencies, commercial organisations, educators, interested citizens - already enjoy findability, accessibility, interoperability, and re-use through a portal (EPOS ICS-C) with a rich metadata catalogue structured with rich syntax and declared semantics and supporting referential and functional integrity. Thus, the stage is set for reliable interoperation (in the widest sense) among RIs in the subdomain. This is already an advantage to stakeholders in the solid earth subdomain and in the EPOS-SP project additional stakeholders are encouraged to participate in EPOS to improve sustainability. This includes commercial organisations and requires not only FAIRness but also relevant quality services with resilience. The solid earth subdomain offering, being integrated into the ENVRI-Hub will provide a much wider group of stakeholders across environmental science with appropriate services and asset access. Populating the

41 https://www.rd-alliance.org/groups/fair-data-maturity-model-wg
42 https://www.w3.org/RDF/

EOSC asset catalogue from the ENVRI-Hub (ongoing work) will reach an even wider group of stakeholders providing them with the services and assets to approach a wide range of environmental problems.

12 References

[1] Bailo,D et al. D10.2: Roadmap for Implementation of FAIR Concepts. June 2020. https://envri.eu/wp-content/uploads/2020/08/ENVRI-FAIR_D_10-2.pdf retrieved 2020-12-18

[2] Rabissoni,R et.al. D10.1: TECHNICAL ANALYSIS AND DEFINITION OF IMPLEMENTATION COMPONENTS FOR FAIR IMPLEMENTATION OF RIS IN THE SOLID EARTH SUBDOMAIN June 2020. https://envri.eu/wp-content/uploads/2020/06/ENVRI-FAIR_D_10-1_Technical-analysis-and-definition-of-implementation-components-in-the-solid-Earth-subdomain.pdf retrieved 2020-12-18

[3] Bailo,D et.al. Perspectives on the Implementation of FAIR Principles in Solid Earth Research Infrastructures Front. Earth Sci., 31 January 2020 https://doi.org/10.3389/feart.2020.00003

[4] Singer, R. (2020). Shape Up. Stop running in circles. Basecamp, 176pp. https://basecamp.com/shapeup/shape-up.pdf . Accessed on 2020-12-xx.

[5] Alessandro Spinuso, Mats Veldhuizen, Daniele Bailo, Valerio Vinciarelli, Tor Langeland; SWIRRL. Managing Provenance-aware and Reproducible Workspaces. Data Intelligence 2022; 4 (2): 243–258. DOI: https://doi.org/10.1162/dint_a_00129

[6] Paciello R., Trani L., Bailo D., Vinciarelli V., Sbarra M. (2023) SHAPEness: a SHACL-driven Metadata Editor. In: Garoufallou, Emmanouel and Vlachidis, Andreas (eds), Metadata and Semantic Research (MTSR 2022): 16th International Conference on Metadata and Semantics Research, London, UK, November 7-11, 2022. Proceedings. Communications in Computer and Information Science (CCIS), Vol. 1789, pp. XXX. Cham (Switzerland): Springer Nature. (Under publication).

13 Annex A – Glossary

AAAI	Authentication, Authorisation, and Accounting Infrastructure
AARC/GEANT	Authentication and Authorisation for Research and Collaboration / Gigabit European Academic Network
API	Application Programming Interface
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
GUI	Graphical User Interface
ICS-C	Integrated Core Services - Central Hub
ICS-D	Distributed Integrated Core Services Distributed
PID	Persistent Identifier
SDL	Software Development Life Cycle
TCS	Thematic Core Services