D5.7
ENVRI-Hub design study

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

Deliverable abstract

The ENVRI-Hub design study contains the analysis of the technical preconditions for the successful implementation of the design ENVRI-Hub as a centralised access interface to environmental data and services building upon the increase of FAIRness achieved in the ENVRI-FAIR project; it will be based on the validation results from WP7 and Task T5.4, and the achievements in the subdomain implementation work packages WP8-11 of this project.

The design of the ENVRI-Hub started as a multifaceted initiative that brought together engineers, scientists, and managers to create a collaborative platform for environmental research and innovation, aiming at advancing environmental research in Europe. As a web platform, the ENVRI-Hub can provide access to environmental data and services offered by the European environmental Research Infrastructures. The data and services provided by the ENVRIs cover four domains of the Earth system, Atmosphere, Ocean, Solid Earth, and Biodiversity-Ecosystems. The ENVRI-Hub intends to facilitate multi-and-inter-disciplinary environmental research by making the data and services interoperable across the domains and easy to use by anyone in the ENVRI community. From different perspectives, depending on the role and interest of the user, the ENVRI-Hub encompasses various aspects, including technical considerations, scientific content, usability, and strategic management. This document describes the process of designing the ENVRI-Hub during the ENVRI-FAIR project, highlighting the key challenges and achievements as well as the potential of an operational ENVRI-Hub in the future.

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	Angeliki Adamaki Alex Vermeulen	ULUND ICOS ERIC	26 June 2023
Contributing Authors	Members of TF6		
Reviewer(s)	Ulrich Bundke Zhiming Zhao	FZJ UvA	16 June 2023
Approver	Andreas Petzold	FZJ	30 June 2023

DELIVERY LOG

Issue	Date	Comment	Author
V 1.0	2023-06-09	First version of the deliverable sent for review	Angeliki Adamaki
V1.0	2023-06-16	Review received	Ulrich Bundke Zhiming Zhao
V1.1	2023-06-23	Updated version of the deliverable sent for review	Angeliki Adamaki
V2.0	2023-06-26	Second version of the deliverable submitted to the Project Office	Angeliki Adamaki

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

D5.7 – ENVRI-Hub Design Study

The design of the ENVRI-Hub started as a multifaceted initiative that brought together engineers, scientists, and managers to create a collaborative platform for environmental research and innovation, aiming at advancing environmental research in Europe. As a web platform, the ENVRI-Hub can provide access to environmental data and services offered by the European environmental Research Infrastructures (RIs). The data and services provided by the ENVRIs cover four (4) domains of the Earth system: Atmosphere, Ocean, Solid Earth, and Biodiversity-Ecosystems. The ENVRI-Hub aims to facilitate multi-and-inter-disciplinary environmental research by making the data and services interoperable across the domains and easy to use by anyone in the ENVRI community. From different perspectives, depending on the role and interest of the user, the ENVRI-Hub encompasses various aspects, including technical considerations, scientific content, usability, and strategic management. This document describes the process of designing the ENVRI-Hub during ENVRI-FAIR, highlighting the key challenges and achievements as well as the potential of an operational ENVRI-Hub in the future.

Executive Summary

The ENVRI-Hub Design Study outlines the prerequisites for successfully implementing the ENVRI-Hub as a collaborative platform that provides centralized access to environmental data and services from European Research Infrastructures (ENVRIs) across four domains: Atmosphere, Ocean, Solid Earth, and Biodiversity-Ecosystems. It builds upon the achievements of the ENVRI-FAIR project and results from relevant work packages. This document presents the design process, key challenges, achievements, and the future potential of an operational ENVRI-Hub.

The concept of the ENVRI-Hub, the motivation for this development and the expected impact are discussed in section 1, which outlines the purpose of the study and provides an overview of the expected benefits and impact of the ENVRI-Hub concept. ENVRI-FAIR brought together experts from various disciplines to advance the data and services offered by ENVRIs. The ENVRI-Hub concept emerged to integrate FAIR data and services, providing a central gateway and demonstrator for ENVRIs' offerings. To design the ENVRI-Hub, the ENVRIs had to identify the common development targets, harmonize actions for FAIRness enhancement, design validation guidelines, and align with the European Open Science Cloud (EOSC) integration criteria.

The methodology section 2 outlines the approach taken to engineer the ENVRI-Hub and discusses the technical aspects of designing the concept. Stakeholder analyses and user needs assessment are required to ensure that the design aligns with the requirements of the intended users, while the decided architecture needs to allow the integration of diverse ENVRIs. The ENVRI-Hub prototype in its current form can serve as a demonstrator of FAIRness and can provide a testbed for service validation.

The main results that describe the implementation of the delivered ENVRI-Hub prototype as a demonstrator of this concept are given in section 3. User-centred features have been developed to support discovery and access to data, services, documentation, and practical examples. The section highlights the functionalities of the ENVRI-Hub, providing some generic user stories resulting from demonstrations of the prototype accompanied by recommendations for future development.

The strengths and challenges, as well as the role of the ENVRI-Hub in the ENVRI strategy and future considerations about the sustainability and governance of the ENVRI-Hub are discussed in section 4, with the concluding remarks given in section 5. The document underscores the importance of collaboration, communication, and continuous efforts to achieve technical integration and interoperability of data and services. Integrated platforms like the ENVRI-Hub require ongoing investment in infrastructure and harmonization efforts to be able to support and promote interdisciplinary research.

1 Introduction

While in the big-data era, science continues to evolve attempting to address more and more complex natural and social problems within conceptual frameworks that often require computationally expensive advanced methodologies and machine-to-machine (m2m) actionable data and services. At the same time data science and technology rapidly evolve offering more tools and technical solutions to address scientific demands, but also revealing more challenges as the complexity increases at all stages of data and service development lifecycle.

As humans we understand the importance and necessity of using a common language to communicate. In a similar way, humans need to interact with the machines but also machines need to "communicate" between them to exchange information. Scientists and developers need to work together to find a common ground that enhances the m2m interaction and enables the interoperation of data and services to serve the complex and diverse demands of their end-users. This gives room to investigate the requirements for integrating different systems that operate not necessarily using the same technologies but rather adopting technical solutions that facilitate the interoperability among them. Starting at the very low levels of data (e.g. raw, real-time or near real-time) and reaching data collections or advanced data products and services, issues like metadata models, Authentication and Authorisation Infrastructures (AAI), ontologies, machine actionable licenses and PIDs that support the findability, accessibility and sustainable future of data (enabling proper data usage, citation, usage tracking etc.) are still open while data providers try to work on federated solutions using already existing FAIR (Findable, Accessible, Interoperable, Reusable) enabling resources (Schultes et al., 2020; Wilkinson et al., 2016).

The ENVRI-FAIR project brought together not only diverse disciplines but also a variety of experts with interest in advancing the data and services offered by the Environmental RIs (ENVRIs). One of the common goals in the project was to explore and design a concept that would bring together FAIR data and services from the ENVRIs in a common virtual environment. As a result of this effort, an integration platform was designed, aiming at giving access to the ENVRIs, serving as a central gateway but also as demonstrator of their FAIR data and services. The concept of this hub of services, the ENVRI-Hub, the motivation for this development and the expected impact are discussed in this section.

1.1 Objectives and development targets

The ENVRI-FAIR Work Package 5 (WP05) encompasses a range of significant objectives. The primary focus was to identify common development targets for (meta)data services within the ENVRI cluster. This involved conducting a thorough gap analysis at both the Research Infrastructure (RI) and sub-domain levels (Magagna et al., 2020). By understanding the existing gaps, WP05 guided the harmonization of actions necessary for the RIs to enhance their FAIRness, in collaboration with WP07 (Adamaki and Vermeulen, 2020). In addition to the developments, WP05 had to design the guidelines for the validation of these services (Bailo et al., 2022). With the help of the ENVRI-Hub concept a standardized framework for service validation was explored to ensure the overall quality and reliability of the ENVRI services. This process involved active dialogue with the European Open Science Cloud (EOSC) to align with relevant integration criteria and developments. The insights gained from participating in EOSC groups and initiatives contributed to the continuous improvement of the ENVRI services.

When developing a hub of web services, several important goals can be accomplished as it can offer exploitable results in terms of interoperability, performance, ease of use, and scalability. By prioritizing these objectives, the hub can become a valuable asset that integrates with other systems, delivers high-performance services, provides user-friendly interfaces, and accommodates growing demands. This can be accomplished by

a) adopting protocols and APIs that enable effective communication and information exchange between the integrated systems,
b) making design choices that enable fast response times and reliable service delivery,
c) providing clear documentation, intuitive user interfaces and other user-friendly features to be accessible to users and enhancing the usability of the services,
d) focusing on user-centric design principles, and
e) choosing a scalable architecture that allows the hub to effectively handle increasing demands without compromising performance or availability.

To facilitate the work on the common (for the ENVRI cluster) development targets, cross-cutting Task Forces (TFs) were established, with representatives from various sub-domains to address specific topics critical to the project's success (Adamaki and Vermeulen, 2020). The process presented in this document is based on the achievements of all these cross-RI groups, taking into account their recommendations and challenges. The WP05 tasks, deliverables and cross-WP collaborations that resulted in the current design of the ENVRI-Hub are illustrated in Figure 1.

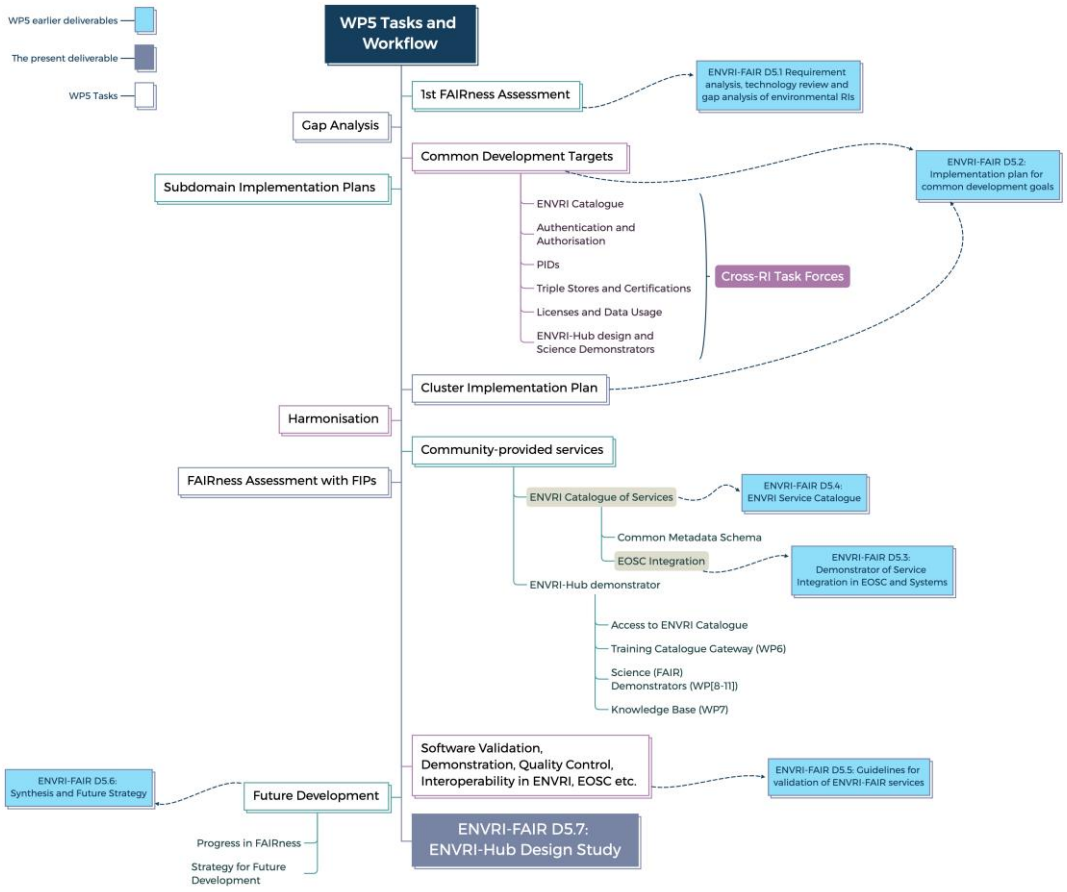


Figure 1. An illustration of the WP05 tasks and the workflow that combined results from cross-RI Task Forces and other ENVRI-FAIR WPs to support the ENVRI-Hub design study.

1.2 Expected benefits and impact

The ENVRI-Hub, as a key output of the ENVRI-FAIR project, represents collaboration, accessibility, and FAIRness. It functions as a demonstration platform, a testbed, and a service validation framework, enabling participating infrastructures and their communities to benefit from enhanced collaboration, improved data accessibility, and a pathway towards achieving FAIR compliance. The participating infrastructures and their communities can thus benefit in many ways:

- **Enhanced Collaboration:** The ENVRI-Hub fosters collaboration among participating infrastructures by providing a common platform for sharing and accessing services. Infrastructures can contribute their services to the Hub, allowing for testing, validation, and integration. This collaborative environment facilitates the exchange of knowledge, expertise, and resources, enabling infrastructures to learn from each other and advance collectively.
- **Improved Data Accessibility:** The ENVRI-Hub greatly enhances the accessibility of data and resources within the participating infrastructures' communities. By centralizing and standardizing services, the Hub ensures that users can easily locate and utilize valuable data products, tools, and services. This increased accessibility accelerates research, supports

multidisciplinary approaches, and empowers researchers to address complex challenges effectively.

- **FAIR Compliance:** The ENVRI-Hub plays a crucial role in ensuring the compliance of services with FAIR principles. It offers a platform for validating services' technical readiness, interoperability, and overall maturity. Through agreed validation processes, services can be assessed on their FAIRness, allowing for the integration of tested and validated services into the Hub. This integration promotes the creation of a robust ecosystem of FAIR-compliant services, benefiting the participating infrastructures and their communities.

Interoperability has a significant impact on environmental sciences by enabling researchers to access and combine data from multiple sources, leading to a better understanding of the environment and more informed decision-making[1]. Interoperable data and big data analytics can be used to monitor the health of ecosystems, detect changes in land use and land cover, track the spread of invasive species, and model the impacts of climate change to answer urgent environmental questions about the environmental challenges the societies face (Runting et al., 2020). Hurricanes, earthquakes, floods, and wildfires have profound effects on ecosystems, human populations, and infrastructure. Climate change has been a topic of scientific discussion for over a century and has now gained significant attention in global politics. Understanding the dynamics and risks associated with natural hazards is essential for effective preparedness and mitigation strategies. Initiatives like the European Green Deal and the United Nations Sustainability Goals call for collective action to address the consequences of climate change. It becomes evident that raising awareness is a crucial catalyst for progress. The environmental research infrastructures bring the responsibility to equip scientific communities with the necessary tools to conduct their research, support policymakers, and raise awareness among citizens. With the provision of data and services the ENVRIs aim to accomplish these goals. The ENVRI-Hub will serve as a platform where the ENVRIs converge, offering a wide range of resources to larger user communities.

The methodology followed in ENVRI-FAIR to investigate the design, architecture and potential of the ENVRI-Hub concept is presented in the following section.

2 Methodology – Engineering the ENVRI-Hub

From a technical point of view, the design of a Hub of services of the type defined by the ENVRIs in ENVRI-FAIR needs a sophisticated technical architecture that enables seamless integration of diverse environmental research infrastructures. To work on the ENVRI-Hub concept the ENVRIs needed to investigate the needs of a complex system that integrates various components and that relies on advanced technologies to handle and support the present and future complex requirements of environmental research. Technical teams across the participating ENVRIs played a crucial role in designing and implementing the underlying technology, ensuring interoperability, data exchange, and compatibility among the different components to the extend this was possible, utilizing the FAIR services they developed or enhanced during the ENVRI-FAIR project. The ENVRIs had to tackle challenges related to the architectural design, the data management, computational resources, and infrastructure scalability among others.

Before developing a hub of scientific services, it is important to identify the needs of the scientific communities the hub will target. Based on the identified needs, the scope of services to be offered by the hub will be better defined.

2.1 Stakeholder analysis - User needs assessment

For several scientists it can be challenging to find the data products they need or discover similar initiatives or communities to collaborate with. While established communities may have their own routines, finding the necessary data products or experts from other fields to address global-scale problems such as the environment can be a difficult task when working in multi/inter-disciplinary fields. With a centralized platform like the ENVRI-Hub such challenges can be overcome and enable collaboration on complex problems. It can be easier for scientists to find the necessary resources and join communities focused on similar research objectives. The scientists will then be able to work together more effectively,

[1] Big Data - Explaining its Uses to Environmental Sciences: https://www.environmentalscience.org/data-science-big-data

particularly when tackling global-scale challenges that require a holistic view of the environment and other complex phenomena.

A survey on user requirements was conducted early in the project, collecting anonymously potential user stories that aimed to describe partners' expectations from the ENVRI-Hub. The results of this survey are presented in the White Paper (Petzold et al., 2023) authored by the ENVRIs participating in the cross-RI group that worked on the design of the ENVRI-Hub (TF6, see section 1.1 above). A subset of the described use cases, focusing only on the basic discovery needs of human users, are illustrated in Figure 2.

One more survey was conducted later in the project, once the ENVRI-Hub prototype was released making it easier for the partners to investigate the user needs by reviewing some basic functionalities of the prototype. Based on the second survey, the ENVRI-Hub is expected to be a flexible platform that meets the needs of scientists from diverse backgrounds. It should offer a smooth user experience, allowing undisturbed search, discovery and access to data and services. The ENVRI-Hub services should empower scientists in their academic research projects, policy-making activities, and industrial applications.

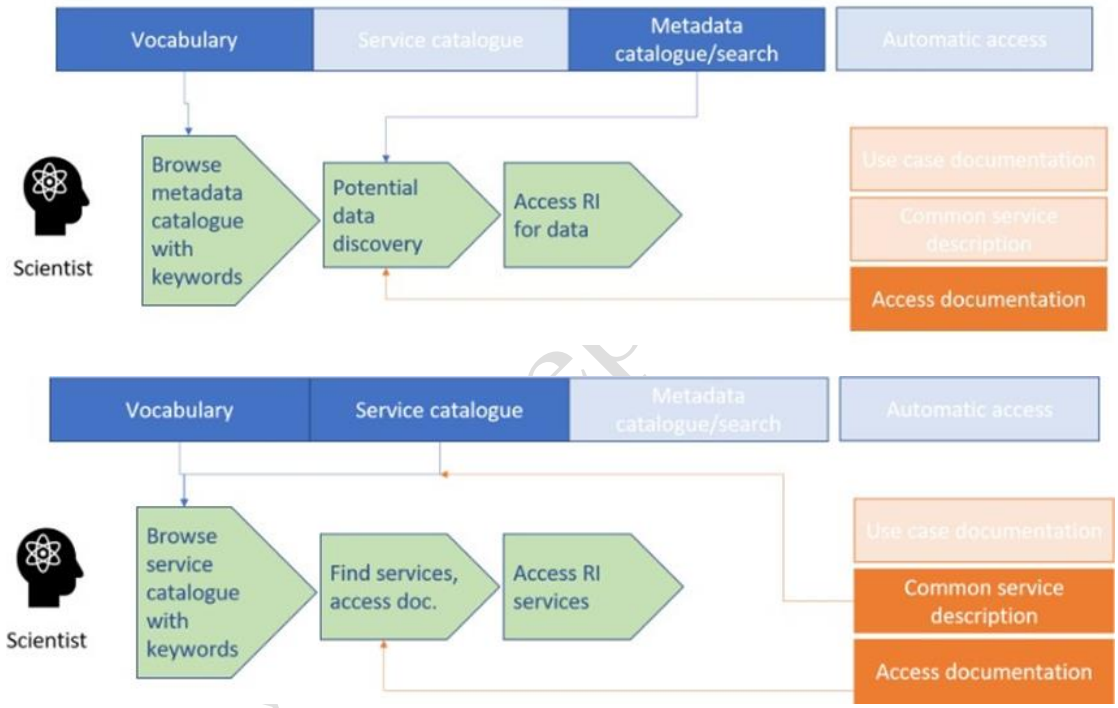


Figure 2. An illustration of the basic user needs identified with the user requirements survey in TF6. Both cases refer to the discovery needs of a human user. More advanced use cases are presented in the TF6 White Paper, where machine action-ability was also required.

The results of the two surveys indicated the key requirements for the design of the ENVRI-Hub, which had to be user-centred and offer discovery and access to data, services and documentation, as well as practical examples and tools that could be used by scientists in their workflows. These requirements are reflected in the conceptual design described below.

2.2 Conceptual design and architecture

Instead of seeking solutions that increase heterogeneity among the ENVRI community, the focus has been on building a platform by scientists for scientists to harmonize the ENVRI developments. The ENVRI community aims to create a platform that addresses the needs of scientists, allowing them to collaborate across different fields and research infrastructures. A community platform of this type should facilitate the discovery and usability of the data and services by providing metadata and descriptions that

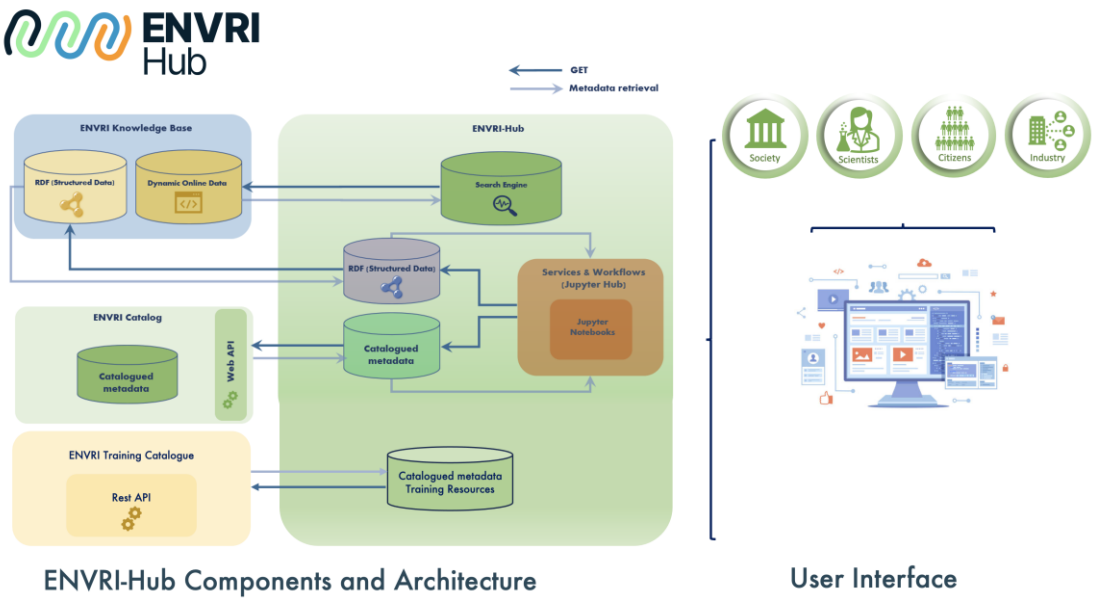
are easy to find, use, and understand. The implementation of the Authentication and Authorization Infrastructure (AAI) across the cluster also needed to be investigated and planned, aiming at providing a single sign-on solution. This would greatly improve the experience for scientific users as they could build workflows that extend beyond their specific field or research infrastructure.

In addition to the above, which were identified as common development targets very early in the project, the ENVRI community worked on developing specific services within their research infrastructures. By integrating these services into the design of the new ENVRI platform, the ENVRI-Hub, the community will have the opportunity to reach broader audiences and attract new communities who can benefit from the platform. The ENVRI-Hub had to be designed to address the advanced needs of users, e.g. those involved in multidisciplinary communities working on interdisciplinary approaches to complex environmental problems.

The ENVRI-Hub prototype consists of four (4) main components:

a) a catalogue of services that provides machine-actionable interfaces to the data and services of the ENVRI partners,
b) a knowledge base that contains technical practices and common data and service requirements for interoperability
c) a training gateway that offers training resources for the ENVRI community, and
d) a series of science demonstrators that showcase interactive data science and scientific computing using mostly Jupyter Notebooks.

Each of these pillars is the result of the work done during the ENVRI-FAIR project to improve the FAIRness of the ENVRIs. The full description of each of these products or services is beyond the scope of this deliverable, as they have all been described in detail in other documents. The ENVRI Catalogue of Services is described in the WP05 Deliverables (e.g. Deliverable D5.4 to be submitted June 2023[2]), the ENVRI Knowledge Base is described in the WP07 Deliverables (Zhao et al., 2020; Deliverable D7.4 to be submitted in 2023[2]), the training gateway is described in the WP06 Deliverables (Vaira et al., 2021; Deliverable D6.3 to be submitted in 2023[2]), and the science demonstrators are described in the sub-domain WP(08-11) Deliverables as their use cases (Carval et al., 2021; Carval and Queric, 2022; Subdomain Deliverables D8.9; D9.8; D10.3; D11.2 to be submitted June 2023[2]). Following the Architecture described in detail in the White Paper of TF6 (Petzold et al., 2023), the main ENVRI-Hub components and the current of the agreed architecture is shown in Figure 3.



[2] All ENVRI-FAIR Deliverables can be found in the ENVRI Community Repository in Zenodo: https://zenodo.org/communities/envri/?page=1&size=20

Figure 3. The main ENVRI-Hub components and the architecture that will allow ENVRI end-users access all available ENVRI resources in one platform. The basic principles of this architecture were tested in the implementation of the ENVRI-Hub prototype.

Parts of this architecture have not been implemented as they only offered a proof of concept in the ENVRI-Hub design study (e.g., the Jupyter Hub). Nevertheless, the basic principles of this architecture have been applied in the implementation process of the ENVRI-Hub prototype which is delivered at the end of the ENVRI-FAIR project as a demonstrator of the ENVRI-Hub concept.

2.3 Implementation – The ENVRI-Hub prototype

While exploring the options for the ENVRI-Hub architecture, the cross-RI Task Force investigated whether the architecture of the ENVRI-Hub could be based on either a centralised or a decentralised system. In the first case, the ENVRI-Hub collects and stores the metadata of the research infrastructures (RIs) in a central database. In the second case, the ENVRI-Hub queries the RIs directly and aggregates the responses. Each system has its own strengths and weaknesses regarding speed, reliability, and scalability. A hybrid system was suggested as a solution that adapts to the different levels of maturity and standardisation of the RIs. The ENVRI-Hub prototype is developed follows an n-tier approach as shown in Figure 4.Figure 4 Here the (local) data tier is synced with the decentralised data regularly for performance and robustness reasons. The decision and the framework are presented in detail in the TF6 White Paper (Petzold et al., 2023). Here only a summary is provided.

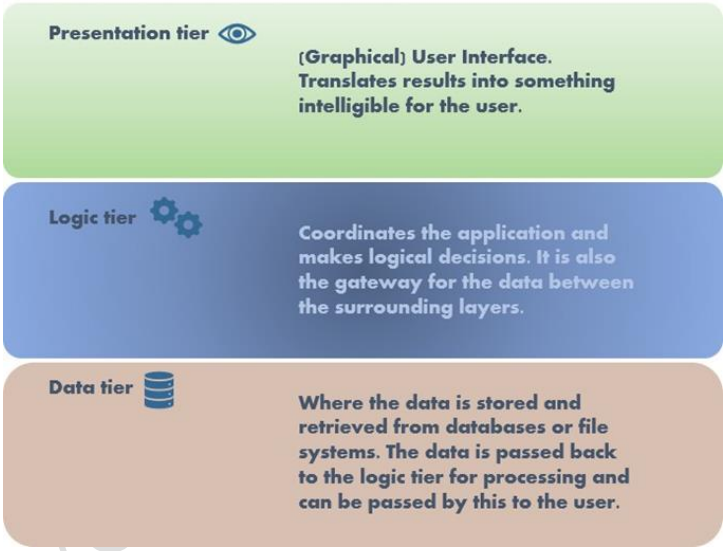


Figure 4. The ENVRI-Hub prototype is developed with a centralised approach following a n-tier pattern with 3 layers as shown here.

Following the suggested architecture, the ENVRI-Hub performs as a harvester entity, collecting metadata from various sources:

ENVRI Catalogue of Services: Metadata in the ENVRI Catalogue is stored in a PostgreSQL relational database, which is also used in the ENVRI-Hub via API. The Hub queries the ENVRI Catalogue's database and persists the retrieved results in the Hub's PostgreSQL database. In addition the primary dcat-AP (RDF) profiles of the services provided by the ENVRIS are stored in Apache Fuseki, a Triple Store that enables resource exposure through SPARQL endpoint and REST service on the Hub.

ENVRI Knowledge Base: The Static Content (RDF) is stored in Apache Jena Fuseki, and exposed through SPARQL endpoint and REST service. The ENVRI-Hub utilizes Jena Fuseki as its persistence system to ensure compliance between the entities. The ENVRI-Hub retrieves metadata from the ENVRI Knowledge Base's Triple Store, which may be in JSON or RDF format, and persists it in the Hub's Triple Store. The Knowledge “Search Engine” is used as the main search engine at the ENVRI-Hub, while the Hub itself has been indexed in the Knowledge Base to become discoverable via the search engine.

ENVRI Training Catalogue: Similarly, the content of the Training Catalogue is queried via the Training Catalogue API, providing at the front-end of the Training Gateway a simplified search functionality.

ENVRI Unified search: A unified search performs a parallel search using the Restful APIs of the ENVRI Catalogue of Services, the ENVRI Knowledge Base and the Training Catalogue, as well as searching through the list of Science Demonstrators. This type of search allows the users to see results from the three sources via one common search function.

Science Demonstrators: Science Demonstrators are developed by various RIs showcasing easy metadata and service access, data discovery, and promoting interoperability at the ENVRI sub-domains. These demonstrators often utilize Jupyter Notebooks, an open-source web application that combines code, visualisations, and narrative text. While no metadata flow exists between the ENVRI-Hub and the Science Demonstrators, the ENVRI Catalogue has the potential to provide metadata to these Demonstrators.

The ENVRI-Hub is built using the Django framework, which is recognized for its ability to facilitate fast development and promote clean, practical design. This powerful high-level Python Web framework has been chosen as the foundation for constructing the ENVRI-Hub prototype. Building on open-source components, Python and PostgreSQL are used at the ENVRI-Hub back-end, which relies on computational resources and infrastructure provided by EGI[3].

2.4 Testing, validation, and evaluation

The ENVRI-Hub is a new concept that serves as a platform for integrating diverse services provided by different infrastructures. To ensure the effectiveness and usability of the ENVRI services and the Hub itself, testing, validation and evaluation processes need to be defined (relevant methods have been reported in ENVRIplus, see e.g., Chen et al., 2017; Hellström et al., 2018).

To ensure successful integration of services in the ENVRI-Hub, those need to be described with rich metadata, be compatible and validated. This involves developing comprehensive guidelines that comply to the FAIR principles (Bailo et al., 2022). The guidelines consider the unique characteristics and requirements of each individual service to guarantee their successful integration. The validation process also includes testing the technical solutions and evaluating their compliance to the minimum requirements and their performance, while the responsibility for the functionality of the asset remains at the level of the provider. The maturity level of the RI services can be assessed using Technology Readiness Level (TRL) criteria[4,5], and the ENVRI-Hub can offer a suitable testbed for experimentation and verification. A simple process considered for the service validation utilising the ENVRI-Hub testbed is illustrated in Figure 5. The testing and validation processes at subdomain, cluster and EOSC level are described in Bailo et al., 2022 and in more detail in WP07 deliverables (Thijsse et al., 2022; Deliverable D7.8 to be submitted June 2023[2]), where several criteria are investigated and a list of tests and their results on RI and ENVRI services are presented. It is important to note that while the minimum requirements for an asset to be ready for ingestion into the ENVRI-Hub are defined by the guidelines agreed within the ENVRI-FAIR project, to build an ENVRI-Hub that is ready for ingestion into the EOSC (or any other higher level platforms) the ENVRI assets need to also comply with the maturity level required for an EOSC service and follow the onboarding criteria defined by the EOSC[6].

Evaluation plays a significant role in assessing the success and impact of the harmonized services provided by the ENVRI-Hub. Target groups of users with varying scientific interests and levels of expertise must be considered. This helps adjusting the services to meet the user needs. For the evaluation process the ENVRIs need to consider some basic parameters which might include the scientific impact,

[3] EGI services used to support the ENVRI-Hub development: https://www.egi.eu/article/the-envri-hub-will-be-powered-by-egi-services/
[4] Technical Readiness Levels (TRL) according to NASA: https://web.archive.org/web/20051206035043/http://as.nasa.gov/aboutus/trl-introduction.html
[5] TRL Assessment Tool: https://ised-isde.canada.ca/site/clean-growth-hub/en/technology-readiness-level-trl-assessment-tool
[6] EOSC Providers Hub: https://eosc-portal.eu/providers-documentation/eosc-provider-portal-inclusion-criteria

the technical impact, the service description and documentation, the user friendliness of the services front-end etc. In the context of evaluation, the ENVRIs also need to consider the granularity of the data and services provided, referring to the level of detail at which data or information is presented to the users and whether it can serve their needs. Regarding the scientific impact the contribution of the services to research and advancements can be evaluated. Proper documentation is crucial for understanding and using the services effectively, while user friendliness ensures good user experience. Technical impact examines the effects of the services on technical processes and systems. An important factor in this case is the evaluation of the FAIRness of the services, which has been a major task during ENVRI-FAIR. the validation of a service for being FAIR is a complex process that depends on the type of the asset that is being assessed for its FAIRness. The ENVRI-FAIR partners have worked on the evaluation of the FAIRness of their repositories by creating FAIR Implementation Profiles (FIPs) which are constructed based on the FAIR enabling resources reported by the under-evaluation RIs. The RI FIPs are then combined into a FAIR convergence matrix which represents the FAIRness level of the ENVRIs and their progress towards becoming more FAIR as a community. The process and the results of the FAIRness assessment using the FIPs is presented in the WP05 deliverables (Magagna et al., 2020; Deliverable D5.6 to be submitted June 2023[2]).

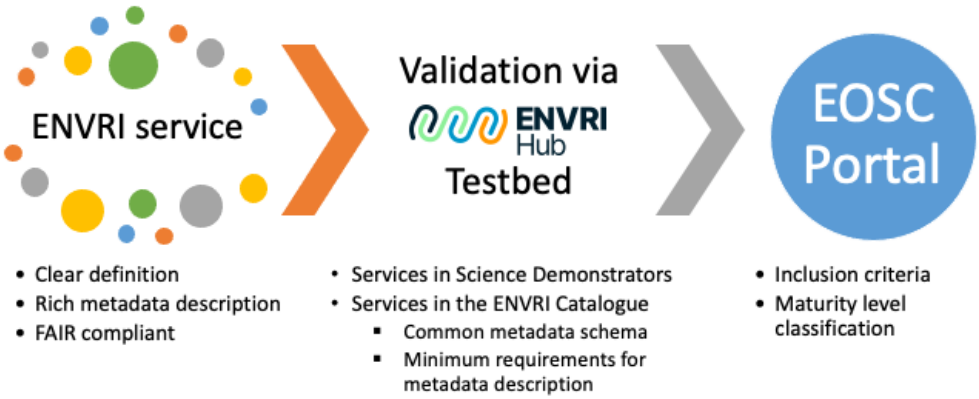


Figure 5. The ENVRI services need to comply with the agreed validation criteria, be described sufficiently with rich metadata and be FAIR compliant. The ENVRI-Hub offers a testbed for the services validation.

3 Results – The ENVRI-Hub Demonstrator

The ENVRI-Hub in its current form serves a double purpose in the ENVRI-FAIR project as it takes the role of:

- **Demonstrator of FAIRness:** As part of the project, each ENVRI has diligently worked on enhancing the FAIRness of their repositories. This progress is showcased through the ENVRI-Hub, which serves as a demonstrator by populating a prototype with several ENVRI services. These services demonstrate the adoption of FAIR solutions and highlight the commitment to making data and resources Findable, Accessible, Interoperable, and Reusable.
- **Testbed for service validation:** The ENVRI-FAIR project has facilitated the development and improvement of services within participating infrastructures to align with FAIR principles. The ENVRI-Hub can serve as a dedicated testbed where these services can be tested and validated. As a prototype the ENVRI-Hub does not provide the required infrastructure for running e.g. science demonstrators, but functions as a platform for evaluating and validating the integration of services in its environment. Contributors have the opportunity to submit their services for evaluation on a shared testbed, promoting collaboration and fostering the interoperability of services. This centralized platform simplifies the onboarding process for ENVRI services and provides users with harmonised access to the validated and interoperable services.

The ENVRI-Hub Demonstrator can be accessed via this link: https://envri-hub.envri.eu The basic features and functionalities of this prototype together with some experience gained by the demonstration of the ENVRI-Hub prototype at larger meetings are described in this section. In Figure 6 a screenshot of the new visual identity of the ENVRI-Hub is provided (last access June 2023).

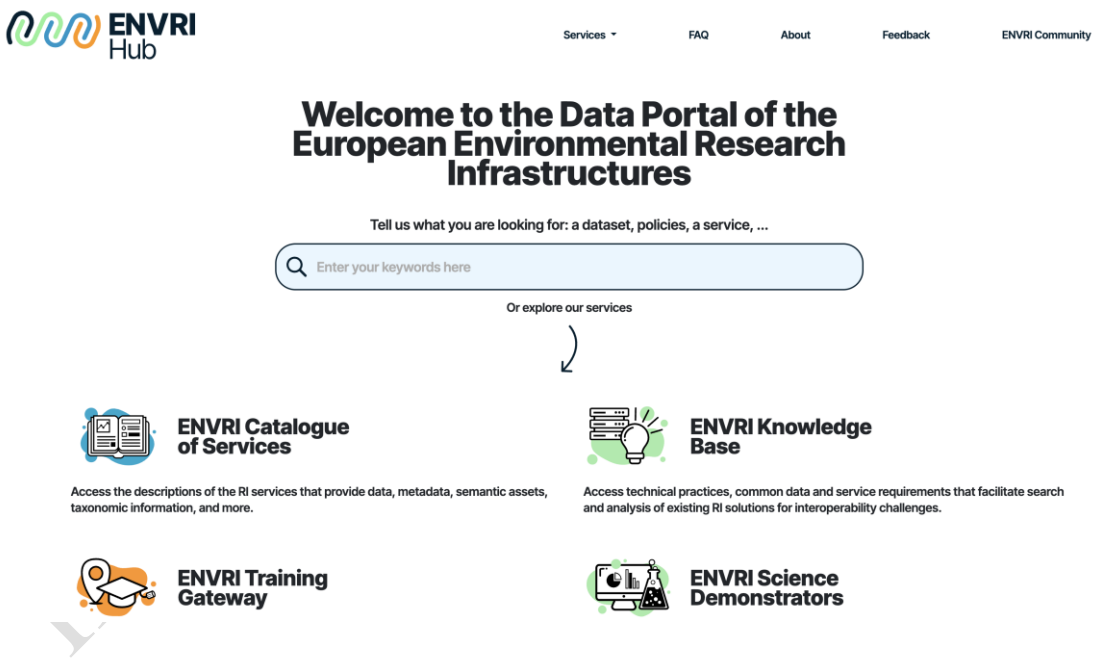


Figure 6. The new visual identity of the ENVRI-Hub, as this was demonstrated at the end of ENVRI-FAIR. The first page gives direct access to all four main components, i.e., the Catalogue of Services, the Knowledge Base, the Training Gateway and the Science Demonstrators, together with the option for a unified search through all components.

3.1 Features and functionalities

The home page of the ENVRI-Hub gives direct access to all four (4) pillars that compose the web application. General information about the project is displayed, following the most recent version of the ENVRI visual identity. A top menu bar allows the users to navigate through the different ENVRI-Hub pages.

Catalogue of Services: This is the User Interface built to enable human users explore, discover and access ENVRI assets which are described following the agreed common metadata schema (DCAT EPOS-AP, see Deliverable D5.4 for a description of the catalogue, the used vocabularies and the metadata schema; to be submitted June 2023[2]). As described in section 2.3, the ENVRI-Hub harvests the service metadata via the catalogue API[7], allowing the users to filter the services based on keywords or standard vocabularies that have been used during the metadata mapping and ingestion process. More than sixty (60) assets are available via this interface, mostly representing web services that give access to the RI data. The user can select the service they wish to explore, read the description given by the provider and display some of the metadata on the Hub interface (e.g., data coverage, service endpoints, documentation where available etc.). The richness of the information the user can find in this metadata catalogue depends on the richness of the chosen metadata profile and the metadata mapped into the common schema by the providers.

SPARQL endpoint and client: A more advanced search function was explored at the final stages of the ENVRI-Hub implementation phase, implementing in one version the RDFlib Python package and in a second version the Apache Jena Fuseki package. In both services, an automatic process harvests the metadata files (in ttl format) from the ENVRI Catalogue GitLab[8], redacts sensitive/personal information, loads the information into the endpoints and exposes the metadata via SPARQL queries in the Hub environment.

Service availability check: This is a quick automatic test which validates frequently (every 15min for this prototype) the services in the ENVRI Catalogue. By querying the field "serviceEndpoint" in the metadata template, this test can provide information about the status of a service (if it is live or not) in near-real time. The results of this test are displayed on the Hub, allowing the user to see if the service they want to access is available at the time they perform their search. It is important to note that this test cannot provide accurate results if e.g. the metadata profile of a service is incomplete due to lack of specific parameters which might be required to query the service endpoint. Nevertheless, the service providers bring the responsibility of the status of their services, and they can provide instructions about how to query their endpoints by linking the service documentation to their metadata profiles.

ENVRI Search engine and Knowledge Base: The search engine of the Knowledge Base (KB) was at first exposed at several pages of the ENVRI-Hub Demonstrator, sometimes adjusting the search field to the content of the page (e.g., at the section of the Science Demonstrators the search engine would search the KB for Notebooks). The user could in that case search on the ENVRI-Hub interface, but for the results they would have been redirected to the KB interface. At a later step, a unified search option was implemented in the Hub, allowing the user to search the keywords of the Catalogue of Services, the Knowledge Base and the keywords of the Training Catalogue simultaneously using their APIs, and also search through the list of Science Demonstrators. When the users choose to enter the service "Knowledge Base" exposed on the ENVRI-Hub landing page or listed in the top menu bar, the users will be directed to the KB environment and get access to the KB search engine.

ENVRI Training Gateway: As in the case of the ENVRI Catalogue of Services, the ENVRI-Hub harvests the metadata of the training resources that are available at the training catalogue (hosted and managed externally by LifeWatch ERIC). On the User Interface of the Training Gateway the user gets information about the training catalogue (which includes all metadata of the available training resources), the training platform (which offers a virtual learning environment to users that wish to follow the available courses) and the learning paths suggested by WP6 in ENVRI-FAIR. The users can search for training resources and receive the results on the Hub environment (resulting from querying the training catalogue), from where they will be redirected to the training platform if they wish to access the training material.

ENVRI Science Demonstrators: The ENVRI-Hub prototype at this point provides static web pages with the descriptions of the ENVRI-FAIR use cases and URLs to locate the Science Demonstrators' resources. The resources include repositories that host their source codes, links to services where those are available, relevant publications etc. A very promising feature that is currently tested only in one scientific demonstrator is the "Hands-On" option, where the implementation creates an offline version of the notebook output (in the available example an interactive map of RI stations where the user can

[7] ENVRI Catalogue of Services API Gateway: https://ics-c.epos-ip.org/demo/k8s-epos-deploy/envri-fair-catalogue/api/v1/ui/

[8] ENVRI Catalogue GitLab: https://epos-ci.brgm.fr/ics-tcs/epos-dcat-ap/envri

choose the area of their interest and display basic metadata from the available measuring stations or sites in the area) and displays this on the Hub interface, giving also the option to the user to download the notebook that creates this output and run it in their own environment. A Jupyter Hub was considered to offer a virtual environment to the users to run the ENVRI notebooks (as seen in the architecture design in Figure 3) but this is not offered at the moment by the ENVRI-Hub demonstrator for security reasons.

FAQ, About, Policy, ENVRI Community link: The ENVRI-Hub Demonstrator offers a series of static pages which provide additional information to the user. The FAQ aims at listing the most common questions related to the structure or the functionalities of the Hub, the About page gives more general information about the project, while the Policy section clearly states the privacy policy and terms of use as those have been decided by the ENVRI-FAIR project. The ENVRI Community is linked to the ENVRI-Hub at the moment via the ENVRI website[9].

Feedback: A feedback form has been included in the top menu bar of the ENVRI-Hub, asking the users to provide their valuable feedback on the services they have used. The feedback is anonymous, and the form was implemented when the project partners started offering live demonstrations of the ENVRI-Hub prototype at larger meetings (e.g., the General Assemblies of EGU). The user stories explored in these demonstrations are summarized in the following section.

3.2 Demonstration of usability – User stories

For scientists working at different domains and sectors, the ENVRI-Hub can serve as a comprehensive scientific platform that brings together environmental data, models, and tools. It can offer a centralized access point to explore, analyse, and share data across different domains, fostering interdisciplinary collaborations. Researchers would value an ENVRI-Hub that can provide rich content, covering a wide range of environmental parameters, high-quality and diverse observation data and simulations, and analytical tools. A key aspect from the scientific perspective is the usability and ease of understanding the platform, enabling researchers to efficiently find and access open and free relevant data, tools, and services based on their needs.

Scientists interested in environmental research: A generic user journey would start with the end-user accessing the ENVRI-Hub platform via its website and start with the search and discovery of data and services relevant to their research project. As a next step the user can integrate the data and data products into their research or use the Jupyter Notebooks provided with the science demonstrators for data science tasks. The user might choose to customize and adapt the available software, accesses the knowledge base to explore more resources or use the training gateway for further learning and skill development. The ENVRI-Hub empowers the end-user as a scientific user, providing a comprehensive platform for efficient data exploration, analysis, and collaboration within the environmental sciences domain. An example of such a user journey, for a user that has scientific background and is interested in environmental data and services is illustrated in Figure 7. An example of a user's search is shown in Figure 8 where the ENVRI-Hub results from a search regarding "jupyter" are listed. In this example the user gets access to several pages and notebooks from the Knowledge Page, the Jupyter Hub offered as a service from ICOS, and the training resources from the Training Catalogue.

Scientists working in policy making organisations: A more specific example of user that can be considered when demonstrating or evaluating the usability of the ENVRI-Hub corresponds to the stakeholder group of policy makers. In a coastal region for example, policy makers might work on an integrated coastal management plan to address the challenges of climate change, marine ecosystem conservation, and sustainable tourism. They will need data on water quality, species distribution, marine biodiversity, aerosols, and trace gases to understand the air quality impacts on the ecosystems and humans. Or in an agricultural region of the EU policy makers might be interested in promoting sustainable farming practices and building climate resilience, needing data on greenhouse gas emissions and soil conditions to assess the situation and make decisions related to land use planning. Although their needs might change, their journey as ENVRI-Hub end-users still needs to cover data and service discovery, accessibility and use, following a similar path as the one described in Figure 7.

[9] ENVRI Community website: https://envri.eu

Scientists working in the industry: Here the example of the end-user can come from the energy industry, where stakeholders are increasingly focused on optimizing energy efficiency and transitioning to renewable resources. They recognize the importance of data-driven solutions to drive sustainable energy practices. To achieve this, they would need to rely on data from multiple infrastructures, e.g. air pollutants and aerosols to assess the environmental impacts of the energy production, carbon footprints and geological profiles of the areas they investigate, possibly combined with their own data from their energy projects and monitoring. The user journey once again will start with the discovery of relevant datasets, the services and software that will allow them to use the ENVRI data and combine them with their own resources.

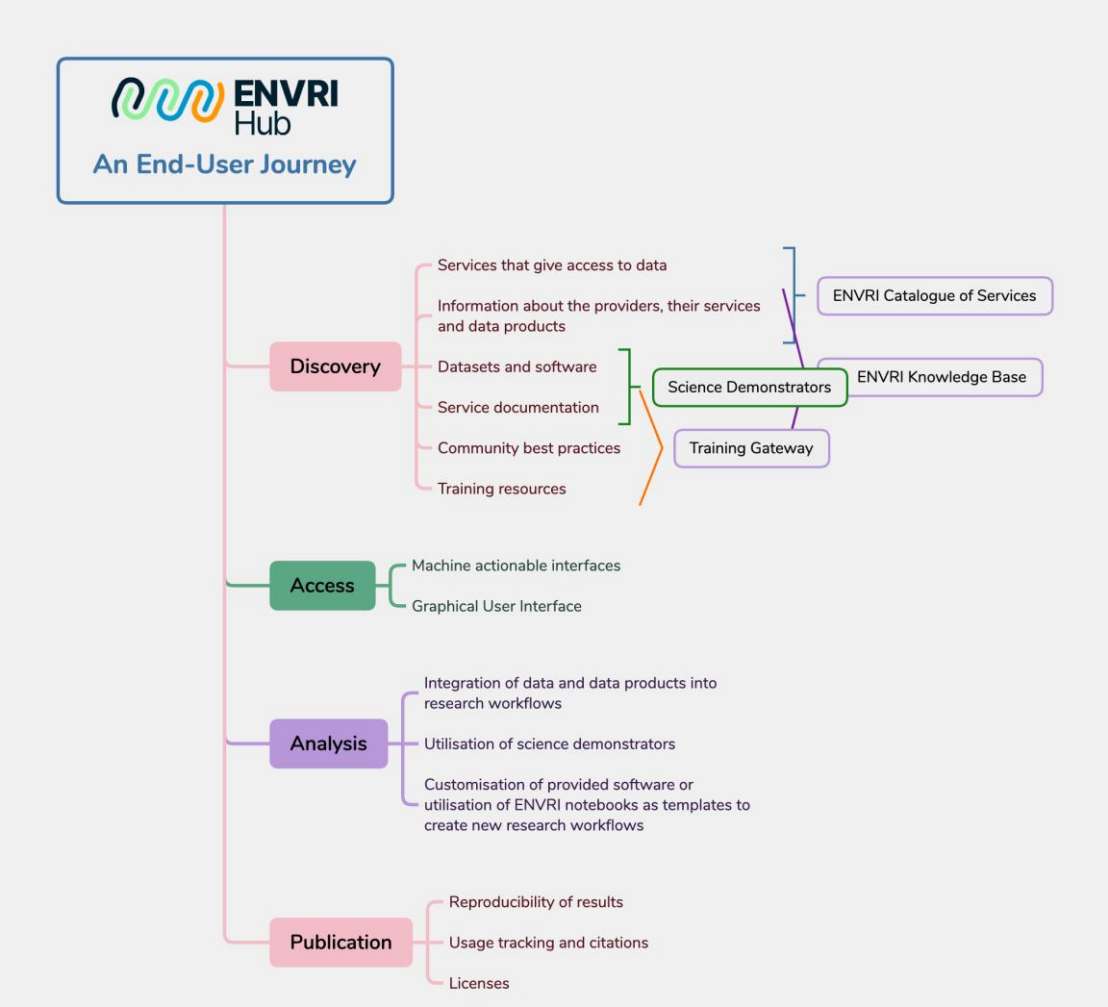


Figure 7. An example of an end-user journey, assuming that the end-user has a scientific background and is interested in environmental data and services for their project.

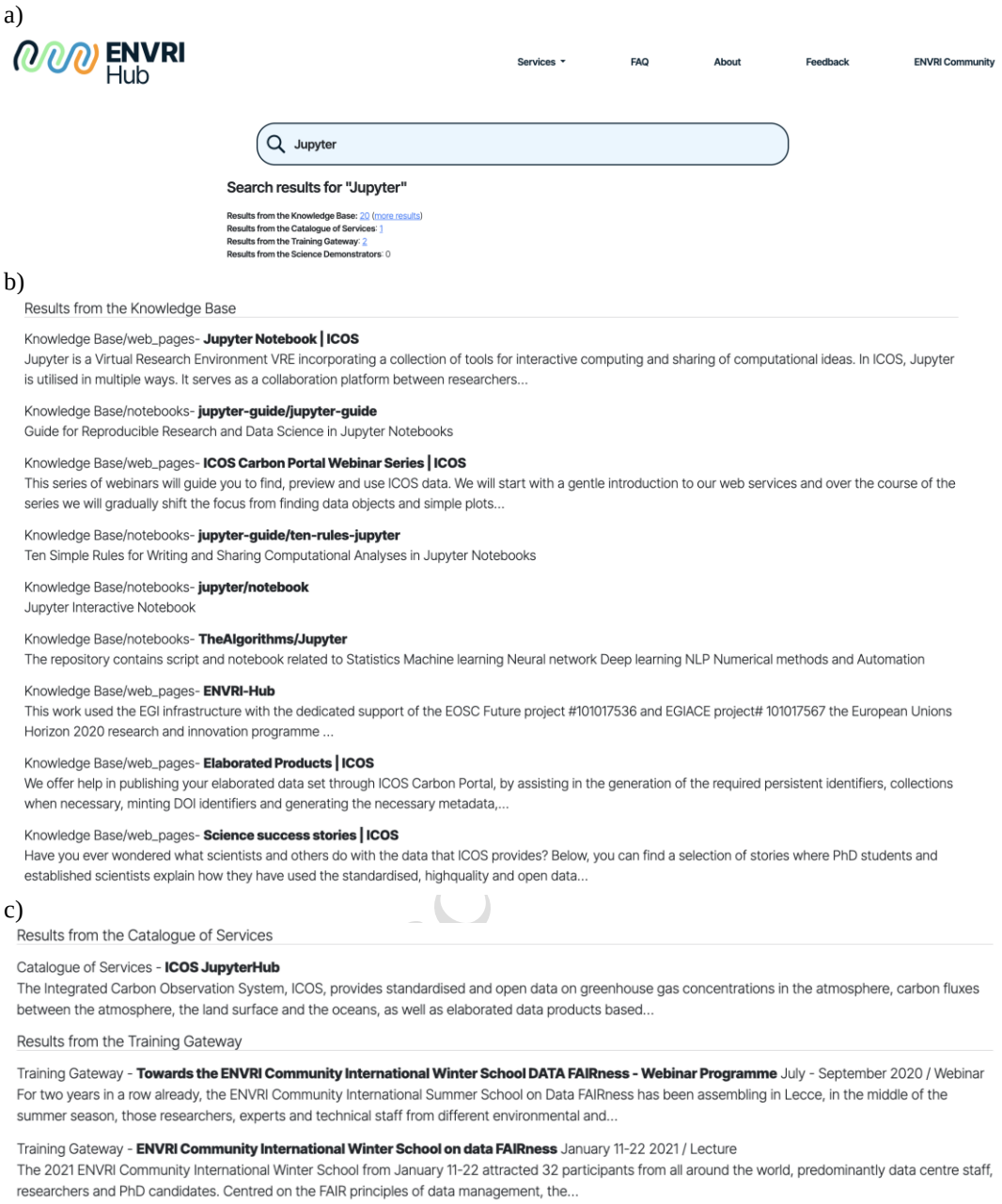


Figure 8. a) The ENVRI-Hub unified search has been used to search for “Jupyter”, giving results from 3 sources. b) Some of the results from the Knowledge Base displayed in the ENVRI-Hub environment include webpages and notebooks. c) The Catalogue of Services gives access to the ICOS Jupyter Hub, while the Training Gateway gives 2 training resources to the users.

3.3 Recommendations for future development

The ENVRI-Hub has been designed to facilitate growth and collaboration by enabling the participation of additional infrastructures that meet the necessary standards. The ultimate vision is to establish a dynamic system within the ENVRI cluster, where scientists can actively engage as users, providers, and co-creators. Although the current version of the ENVRI-Hub represents the "minimum viable product" within the scope of the ENVRI-FAIR project, the concept holds significant potential. To fully realize this potential, a more sophisticated interoperation layer is required, which would ensure better

harmonization of the ENVRI services, data, and products. This enhanced integration will provide ENVRI users with an advanced virtual environment, offering a seamless research experience.

The current architecture of the ENVRI-Hub is modular and scalable, allowing for the implementation of new functionalities. One desired feature is the establishment of access to data, enabling researchers to easily discover and utilize relevant datasets. A community metadata schema, that will cover the basic needs of all service providers, needs to be developed. To enhance user experience and streamline access, the consideration of a single sign-on solution is also important. This would enable users to access multiple services within the ENVRI-Hub using a single set of credentials, simplifying the authentication process. Furthermore, to optimize the services provided by the ENVRI-Hub at both the individual Research Infrastructure level and the cluster level, higher level computing resources and e-services will be required. The future developments at the EOSC environment and the provision of EOSC services to the Science Cluster will play an important role in the further development of the ENVRI-Hub.

Looking towards the future, the ENVRI-Hub should aim to provide virtual research environments (VREs) that empower environmental scientists to perform their analysis and computations directly within the ENVRI-Hub. By integrating tools like Jupyter Notebooks, researchers can conduct their scientific tasks efficiently and collaboratively. Additionally, to foster engagement and facilitate knowledge sharing the ENVRI-Hub should also consider establishing a helpdesk or a forum platform. These platforms would serve as spaces for users to seek assistance, exchange ideas, and contribute to the broader ENVRI community, creating a supportive and interactive environment.

4 Discussion

Designing the ENVRI-Hub and engineering its very first prototype gave the opportunity to the ENVRIs to explore the key areas of such a project, gain deeper understanding of the requirements, the implications, and the potential impact of the ENVRI-Hub within the ENVRI community. This section discusses some of the lessons learnt during this process, the sustainability and governance considerations when focusing on the long-term viability of such a project and its strategic role for the ENVRI community.

4.1 Strengths, challenges, and limitations

As explained in the beginning of this document (see section 1.2) the ENVRI-Hub can be seen as a dynamic and collaborative platform. Working on this concept enhanced the collaboration among the participating RIs and improved significantly service discoverability and accessibility, improving at the same time the FAIRness of the RIs, their services and the ENVRI cluster as a whole. When approached from different perspectives, successful implementation and operation of the ENVRI-Hub depend on close collaboration and synergy among engineers, researchers, and managers. Engineers provide the technical expertise to design and develop the platform, addressing the computational requirements and infrastructure. Researchers contribute their scientific knowledge and domain-specific requirements, shaping the content and usability of the ENVRI-Hub. Managers provide the strategic vision, securing funding, and establishing governance frameworks to support the long-term sustainability and growth of the initiative. The success of the ENVRI-Hub relies on the collective efforts and collaboration among these diverse perspectives, facilitating ground-breaking environmental research and innovation.

During the ENVRI-FAIR project, the ENVRIs encountered several challenges while working on the design of the ENVRI-Hub. There were several technical challenges associated with the heterogeneity of the existing technologies in the ENVRI landscape. These challenges encompassed various areas such as metadata schemata, data management routines, data discoverability, and even different definitions and understanding of the services and datasets at various levels, including a variety of types, formats etc. Inconsistent or ambiguous definitions can lead to confusion and interoperability issues when attempting the integration of different data sources, services, or systems at higher level integration platforms. To address these technical challenges and improve the overall functionality of the ENVRI-Hub, it is important to focus on achieving interoperability, ensuring that the technologies used within each infrastructure can work together. Aligning these technologies with the FAIR principles is crucial to facilitate the higher-level integration.

At cluster-level services, there were technical challenges in devising architectures for cluster-level products that meet the diverse requirements of providers, users, and strategies. Developing such architectures requires careful consideration and collaborative planning, but also clearly defined governance and sustainability plans. Identifying the common drivers of service development was challenging as well. Should they be primarily user-driven, technology-driven, science-driven, or data-driven? Agile or similar iterative processes are proven to be useful for service development to facilitate continuous improvements within the project.

Another significant challenge was the communication among the different teams, technical or scientific, all coming from different backgrounds and fields. Academic institutions participating in such projects may face resource constraints when it comes to technical expertise required for the tasks. Aligning the advancements in technologies, often driven by industry, with the evolving needs of science, which is primarily driven by academic research, is not easy. Such tasks need to respect the rich traditions of each scientific field while accommodating the modern requirements of current and future generations of scientific users. Additional guidance is required in understanding service development and effectively communicating the user needs to the developers. Developing a common scientific language that could later be translated into standard vocabularies understandable by machines was important and it will remain work in progress.

4.2 Sustainability and governance considerations

As the ENVRI-FAIR project is about to conclude, it is necessary to carefully consider the sustainability of the outcomes of the project. Even though the ENVRI-Hub demonstrator has been developed as a proof of concept, the key factors to sustain this platform were often discussed among the project partners.

Regarding the technical considerations the ENVRIs had to investigate and assess the available technical solutions and identify technical developments to meet evolving requirements. Formulating recommendations to improve the ENVRI-Hub's technical aspects and tackle known or expected challenges are necessary steps. Technical upgrades and maintenance are essential for the ENVRI-Hub's sustainability. Providing ongoing technical support, regular upgrades, and bug fixes are necessary to maintain a reliable user experience. Collaboration with technical experts and leveraging open-source solutions can streamline these activities.

Financial considerations play a significant role when thinking about the sustainability of the ENVRI-Hub. Developing sustainable funding mechanisms is necessary to support ongoing maintenance, enhancement, and expansion. Allocating resources and planning for further development at RI level are important aspects to ensure the long-term viability of the ENVRI-Hub.

Another aspect to consider is governance. Exploring federated solutions can foster collaboration among participating infrastructures, while defining policies governing the ENVRI-Hub's operation, maintenance, and development can ensure effective management. It is crucial to establish a governance framework that remains adaptable and robust to meet present and future requirements.

User engagement is also vital for sustainability. Evaluating the usefulness of the ENVRI-Hub and its services helps ensure continued relevance to the scientific communities and societal challenges. Gathering feedback from stakeholders, establishing communication channels for user inquiries and concerns, and promoting integration with external initiatives can contribute to active user engagement.

Fostering collaborations and partnerships with relevant organizations and initiatives is another important factor. Establishing connections with research infrastructures, scientific communities, and international initiatives working towards similar goals enables resource pooling, sharing of best practices, and knowledge exchange. Communicating the value proposition of the ENVRI-Hub to participating infrastructures and communities will be crucial. Highlighting the benefits, such as access to diverse services and improved research outcomes, reinforces the significance of the ENVRI-Hub and encourages continued engagement and support.

The ENVRI-Hub can thrive as a valuable resource for the environmental research community, facilitating collaboration, data sharing, and innovation beyond the ENVRI-FAIR project. Continual collaboration, funding stability, effective governance, user-centric approaches, technical excellence, and strategic partnerships will contribute to its longevity and continued impact.

4.3 The ENVRI-Hub in the ENVRI strategy

From a strategical point of view, the ENVRI-Hub is a key outcome of the ENVRI-FAIR project as it can serve as a central gateway that showcases the achievements and benefits of the ENVRI-FAIR project for various stakeholders, such as policy makers, the industry, and the public. The ENVRIs, considering aspects such as stakeholder engagement, funding, sustainability, and governance, recognize the importance of fostering collaborations among research institutions, industry partners, and policymakers. While the long-term viability of the ENVRI-Hub depends on the future funding for the development of the concept into an operational platform, the RIs needed to address already at the design phase of the ENVRI-Hub the governance models that will ensure transparency and efficient decision-making processes, considering the needs and expectations of the diverse stakeholders involved.

As a key outcome of the ENVRI-FAIR project, the ENVRI-Hub also plays an important role in the ENVRI strategy towards the EOSC. By implementing FAIR principles and demonstrating FAIR practices through its prototype, the ENVRI-Hub sets an example of prioritising Open Science practices, aligning the ENVRI community with the broader objectives of the EOSC.

5 Conclusion

The development of the ENVRI-Hub prototype within the ENVRI-FAIR project has shed light on both best practices and challenges when aiming to achieve technical integration and interoperability of (meta)data and services across different levels of granularity. Integration at this level requires careful consideration of various factors, including technical standards, metadata standards, ontologies etc. To achieve a common understanding of semantics and bridge the gaps between them, the ENVRIs need to foster collaboration and communication among experts from diverse disciplines. This collaboration can involve establishing common vocabularies, developing ontologies, and harmonizing metadata standards.

The development of integrated platforms that cater to all types of end-users, whether within a single discipline or across multiple disciplines, is both a challenging endeavour and an ongoing process. Such platforms can facilitate collaboration and foster interdisciplinary research while harmonising access to data and services. However, achieving this vision requires continuous effort, collaboration, and investment in technical infrastructure and harmonisation efforts. Is it ambitious to envision a fully integrated platform that serves all end-users, or is it is a goal worth pursuing?

The ENVRI-Hub concept represents a significant step towards the future of the ENVRI community. It showcases the potential of bringing together diverse data and services in a common virtual environment. By addressing the challenges of integration, semantics, and FAIRness, the ENVRI-FAIR project has demonstrated the value of harmonizing efforts across RIs to advance open science and benefit various sectors. Improving the FAIRness of data and services is a vital aspect of advancing research in academia, industry and policy making. As the ENVRIs continue their journey towards improved integration and FAIRness they get the chance contribute to the evolution of integrated platforms, promote the openness and exchange of data and services, ultimately enabling more impactful research outcomes and empowering the wider scientific community.

6 References

Adamaki, Angeliki, & Vermeulen, Alex. (2020). ENVRI-FAIR D5.2: Implementation plan for common development goals (Version 1). Zenodo. https://doi.org/10.5281/zenodo.4061702

Bailo, Daniele, Tarabusi, Gabriele, Jeffery, Keith, Adamaki, Angeliki, & Thijsse, Peter. (2022). ENVRI-FAIR D5.5: Guidelines for validation of ENVRI-FAIR services (1.1). Zenodo. https://doi.org/10.5281/zenodo.6758244

Carval, Thierry, Thijsse, Peter, Portier, Marc, Exter, Katrina, Pfeil, Benjamin, Pouliquen, Sylvie, Rodero, Ivan, Kokkinaki, Alexandra, Buck, Justin, Queric, Antoine, & Harscoat, Valérie. (2021). ENVRI-FAIR D9.6: Marine subdomain EOV product specification (Version 1). Zenodo. https://doi.org/10.5281/zenodo.4766796

Carval, Thierry, & Queric, Antoine. (2022). ENVRI-FAIR D9.7: Marine subdomain EOV product version 1 (Version 1). Zenodo. https://doi.org/10.5281/zenodo.6410538

Y. Chen, B. Grenier, M. Hellström, A. Vermeulen, M. Stocker, R. Huber, B. Magagna, I. Häggström, M. Fiebig, P. Martin, D. Vitale, G. Judeau, T. Carval, T. Loubrieu, A. Nieva, K. Jeffery, L. Candela and J. Heikkinen (2017). Service deployment in computing and internal e-Infrastructures. ENVRIplus Deliverable 9.1. Available at http://www.envriplus.eu/wp-content/uploads/2015/08/D9.1-Service-deployment-in-computing-and-internal-e-Infrastructures.pdf

M. Hellström, M. Johnsson, C. D'Onofrio, M. Selander and A. Vermeulen (2018). Serving key data service stakeholders and policy initiatives version 2. ENVRIplus deliverable report D9.4. Available at https://www.envriplus.eu/wp-content/uploads/2019/07/D9.4.pdf

Magagna, Barbara, Adamaki, Angeliki, Liao, Xiaofeng, Rabissoni, Riccardo, & Zhao, Zhiming. (2020). ENVRI-FAIR D5.1 Requirement analysis, technology review and gap analysis of environmental RIs (Version 1). Zenodo. https://doi.org/10.5281/zenodo.3884998

Petzold, Andreas, Gomes, Ana Rita, Bundke, Ulrich, Schleiermacher, Chris, Adamaki, Angeliki, Vermeulen, Alex, Zhao, Zhiming, Stocker, Markus, Lund Myhre, Cathrine, Boulanger, Damien, Hienola, Anca, & Bailo, Daniele. (2023). ENVRI-Hub Design and Architecture White Paper (Version 1). Zenodo. https://doi.org/10.5281/zenodo.8046894

Runting, R.K., Phinn, S., Xie, Z. et al. Opportunities for big data in conservation and sustainability. Nat Commun 11, 2003 (2020). https://doi.org/10.1038/s41467-020-15870-0

Schultes, E., Magagna, B., Hettne, K.M., Pergl, R., Suchánek, M., Kuhn, T. (2020). Reusable FAIR Implementation Profiles as Accelerators of FAIR Convergence. In: Grossmann, G., Ram, S. (eds) Advances in Conceptual Modeling. ER 2020. Lecture Notes in Computer Science(), vol 12584. Springer, Cham. https://doi.org/10.1007/978-3-030-65847-2_13

Thijsse, Peter, Jeffery, Keith, Fiebig, Markus, Fiore, Nicola, & Zhao, Zhiming. (2022). ENVRI-FAIR D7.7: Test and validation at subdomain and cluster report (Version 1). Zenodo. https://doi.org/10.5281/zenodo.5873528

Vaira, Lucia, Manca, Maria Teresa, Vallo, Cosimo, Fiore, Nicola, & Konijn, Jacco. (2021). ENVRI-FAIR D6.2: FAIR training materials catalogue & integration with Common Training Platform (Version 1). Zenodo. https://doi.org/10.5281/zenodo.5052426

Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJ, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PA, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone SA, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar 15;3:160018. doi: 10.1038/sdata.2016.18

Zhao, Zhiming, Liao, Xiaofeng, Goldfarb, Doron, Stocker, Markus, Farshidi, Siamak, & Magagna, Barbara. (2020). ENVRI-FAIR D7.3: ENVRI-FAIR Knowledge Base for RI Service Interoperation and Competence (Version 1). Zenodo. https://doi.org/10.5281/zenodo.4311047

7 Glossary

AAI	Authentication and Authorization Infrastructure
API	Application Programming Interface
DCAT	Data Catalogue Vocabulary
DCAT-EPOS-AP	Data Catalogue Vocabulary - European Plate Observing System - Application Profile
EGU	European Geosciences Union
EOSC	European Open Science Cloud
ENVRI	Environmental Research Infrastructures
ENVRI-FAIR	Environmental Research Infrastructures - Findable, Accessible, Interoperable, and Reusable
ENVRI-Hub	The concept of a central gateway that provides access to services and resources for the ENVRI Community
ENVRIplus	A Horizon 2020 ENVRI cluster project
ERIC	European Research Infrastructure Consortium
EU	European Union
FAIR	Findable Accessible Interoperable Reusable
FAQ	Frequently Asked Questions
FIP	FAIR Implementation Profile
GitLab	A web-based platform for software development and collaboration
ICOS	Integrated Carbon Observation System
KB	Knowledge Base
PostgreSQL	An open-source relational database management system
RDF	Resource Description Framework
RDFlib	A Python library for working with RDF data
RI	Research Infrastructure
SPARQL	SPARQL Protocol and RDF Query Language
TRL	Technology Readiness Level
TF	Task Force
Ttl	Turtle, a syntax for writing RDF data
VRE	Virtual Research Environment
WP	Work Package