D6.3
Training for FAIR services and data – outcomes and experiences

Work Package	WP6
Lead partner	ULUND / ICOS ERIC
Status	Final
Deliverable type	Report
Dissemination level	Public
Due date	30 April 2023
Submission date	30 June 2023

Deliverable abstract

This is the final deliverable of WP6, the training and capacity building work package of ENVRI-FAIR. The objectives of WP6 were to provide training to ENVRIs and key ENVRI stakeholder groups about the FAIR principles, how to implement these in practice in RI services and data management activities at data centre level, how to evaluate the degree of implementation using FAIR metrics, as well as relevant legal and policy requirements. This deliverable report summarises the activities we have undertaken, with special focus on the preparation and running of training events aimed at the ENVRIs' data centre staff. We also describe the infrastructure that was developed to provide a comprehensive Learning Environment for the ENVRI Community, including a catalogue, learning platform and repository and the Training Gateway component of the ENVRI-Hub. Finally, the report also contains an extensive discussion of "lessons learned" that we hope will be of use for future ENVRI Community training activities.

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 Organisation	Date
Main Author	Maggie Hellström	ULUND / ICOS ERIC	26 June 2023
Contributing Authors	Angeliki Adamaki Jacco Konijn Harry Lankreijer Giovanna Maracchia Cosimo Vallo	ULUND / ICOS ERIC LIFEWATCH ERIC UvA INGV	
Reviewer(s)	Giovanna Maracchia Keith Jeffery	INGV UKRI / EPOS	25 June 2023
Approver	Andreas Petzold	FZJ	30 June 2023

DELIVERY LOG

Issue	Date	Comment	Author
V 0.1	19 June 2023	Initial draft	Maggie Hellström
V 0.2	26 June 2023	Included comments and additions from reviewers	Maggie Hellström

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

D6.3 – Training for FAIR services and data – outcomes and experiences

1 Introduction: training "for FAIR", not only "about FAIR"

1.1 The role of WP6 in the project

The overarching mission of the ENVRI-FAIR project is to advance the findability, accessibility, interoperability, and reusability (FAIRness) of digital assets, in particular research data, product and services, provided by the ENVRI Cluster, and to connect them to the emerging service ecosystem of the European Open Science Cloud (EOSC). This requires development and implementation of both technical frameworks and policy solutions that can support environmental researchers working both within and across their scientific disciplines, while at the same aligning Earth system science for the new Open Science paradigm.

As an example, strategic coordination of observation systems is a prerequisite for high-quality and cost-effective interdisciplinary environmental science. This requires wide-ranging harmonisation and standardisation of the collection, curation, and dissemination of data through the implementation of joint data management and access structures at both subdomain and individual research infrastructure level. To complement the adoption of open standards and other interoperability-enhancing solutions, cooperation and co-development of operational services and common sets of policies covering FAIR data stewardship are also needed.

In this context, WP6 "Training and capacity building" objective within the ENVRI-FAIR project is to provide training to ENVRIs and key ENVRI stakeholder groups about the FAIR principles, how to implement them in RI services and data management activities at data centre level, how to evaluate the degree of implementation using FAIR metrics, as well as relevant legal and policy requirements.

It is thus clearly not enough to envisage a training and skills building work package that focuses only on the FAIR principles themselves or basic definitions of Open Science. Instead, the scope must be widened to increase the knowledge in the ENVRIs about the standards, technologies and collaborative services that prepare them for the future FAIR environmental research landscape. At the same time, the learning resources that are developed to assist in this process should also be as FAIR as possible.

1.2 Key pedagogical aspects

Target groups

During the ENVRIplus project, training activities organised for the ENVRI Community were to a large extent focussing on training early-career scientists on topics related to basic research data management. The majority of these participants were still enrolled at, or employed by, institutes of higher education - and thus still very much part of an academic setting where mixing studies and work is normal. During intensive study events such as Summer Schools, instructors were able to apply "true and trusted" pedagogical components, including longer lectures and homework assignments, with confidence in that participants would be able to concentrate their efforts during the entire time at their disposal.

Within the context of ENVRI-FAIR, the focus has shifted to catering to "working professionals," referring to individuals who are fully employed within the ENVRI Community organisations. The aim is to provide them with educational materials and activities that can enhance their skills and knowledge in areas directly related to their work and involvement within the Research Infrastructure environment and the project.

Topics for the training

Events and materials gathered or developed by WP6 aimed to provide detailed and practical instruction on both "general FAIR-related" and "research data management-related" topics. As a first step, we initiated a discussion with representatives of all the ENVRI-FAIR project partners to first learn more

about which FAIR and data management questions and areas they felt should be covered by training during the project. This resulted in 6 “general FAIR” and 16 “RDM for FAIR” topic headings, for which we then in a second step performed a knowledge gap analysis and prioritisation ranking that was reported in deliverable D6.1 [Hellström 2019].

There were some differences between the four subdomains with respect to both estimated existing levels of knowledge and the priorities for training, but also many commonalities. To summarise, the initial prioritisation was as listed in the following; however during the evolution of the project this changed somewhat, with topics related to the development and curation of (data) services coming to the fore. (All topics are listed in Appendix 1 with brief explanations.)

High priority: Metrics for FAIRness evaluation (G2), Performing a FAIRness self-assessment (G3), GDPR (General Data Protection Regulation) issues related to data sharing (G4), Writing (technical) documentation for services (G6), Cloud computing (Virtual Machines & containers) for data processing (R5), Licences & policies for data use (R8), Provenance tracing (R13), Workflow engines for automated data processing (R16)

Medium priority: Certification schemes for repositories (CoreTrustSeal) (R4), Linked Data and ontologies (R9), Metadata standards & schemas (including geospatial, instruments, variables) (R10)
Lower priority: Introduction to FAIR principles (G1), Basic Research Data Management (RDM) (G5), Access control (Authorisation-Authentication-Identification, or AAI) methods (R1), API (Application Program Interface) design for data & metadata access (R2), Cataloguing - design & implementation (R3), Data Management Plans (R6), Landing page design (R7), PID allocation & use (including citation support, bibliometry, provenance) (R11), Portal design & operation (R12), Repository design, operation & sustainability (R14), Virtual Research Environments for data analysis (design & implementation) (R15)

“Shelf-life” and sustainability

To fulfil our mission, WP6 organised and carried out different types of training activities, both face- to-face and, because of the pandemic restrictions, on-line (see the next Chapter for a detailed description). While some of the events have covered topics of interest to wider target audiences, a majority have focused on specific aspects of FAIR and research data management, often presented within very specific project-related contexts such as the need to provide tutorials on how to understand and implement a given cataloguing metadata model.

Following the intentions and plans outlined in the Description of Action (DoA), the training-related resources - recordings, presentation slides, reading suggestions etc. - have been collected, organised and made available in order to support the offered training activities. However, these learning resources are in general not designed to serve as fully-fledged stand-alone training course materials, but can rather be seen to serve as useful post-event reminders of what was presented during a training event, such as a webinar or workshop. The development of courses and comprehensive educational packages, while of potentially great interest also beyond the ENVRI Community, has certainly been beyond the scope of the ENVRI-FAIR project.

Learning outcomes and paths

Thus, both the overall usefulness and "shelf-life" of a training (and the related materials) can vary considerably, especially when considering audiences from the wider EOSC (non-ENVRI) community. These ephemeral aspects influenced the level of efforts originally planned to be invested in composing and disseminating explicit learning paths, learner objectives etc. However, during the project midterm review, the external expert panel expressed the view that much of the training developed and carried out in ENVRI-FAIR would be very valuable also in a wider EOSC user context. The experts stressed that in order to increase both the FAIRness and pedagogical usability of the ENVRI learning resources, it would be crucial to invest time and resources towards annotating them with clearly defined learning outcomes, as well as identifying suitable learning paths.

In order to capitalise on this expert advice while at the same time keeping to a realistic and feasible total workload, we took stock of all the learning resources listed in the catalogue and learning platform and made a selection of what we believe could indeed be of interest beyond the ENVRI-FAIR project lifetime

to prospective learners from both the ENVRI Community and more generally. Based on the selection, we developed a set of demonstrator learning paths and made these available via the Training Gateway component of the ENVRI-Hub. The paths are described in more detail below in Section 1.4.

1.3 Training events organised during the project

Throughout the duration of the project, WP6 organised and/or facilitated 28 training events - a complete list in chronological order is found in Appendix 2. In the following sections we describe some highlights among the webinars, workshops and schools.

1.3.1 Webinars

A majority of the training events were carried out in the form of webinars - typically 60-minute-long video meeting platform-based events with 1-3 presentations by experts followed by a discussion or question & answer session with the audience. Typically, the events would be recorded, and the resulting video (after some light editing) made available afterwards together with slide decks and, if applicable, other documentation such as collaborative notes and meeting chat logs. Because of the project-internal nature of some of these webinars, access to these materials is not always open.

Stand-alone webinars

Examples of stand-alone webinars include: *"How FAIRsharing can help FAIRify your standards, databases and data policies"* (held on 2019-12-13 with 23 participants) with Peter McQuilton who explained the activities of FAIRsharing.org aimed at guiding consumers to discover, select and use resources on data and metadata standards inter-related to databases and data policies; *"FAIR the smart way: Introducing the ENVRI Knowledge Base"* (held on 2020-06-20, with 41 participants) with Markus Stocker et al. explaining the structure and semantic technologies behind the ENVRI Knowledge Base; and *"GDPR in Context"* (held on 2021-02-04, with ca 30 participants) with Keith Jeffery introducing the European Union's General Data Protection Regulation from a research infrastructure point of view.

Webinar series

In response to calls for training and skills building that arose from discussions and work taking place in the common service development and technological support work packages (5 and 7) and the subdomain work packages (8-11), WP6 stepped in to organise or facilitate several series of webinars on important research data management-related topics such as *Provenance Tracing* (4 webinars held in November-December 2020, led by Doron Goldfarb et al.) and the *RDA I-ADOPT Framework* (2 webinars held in January-February 2022, led by Barbara Magagna et al.).

Further webinar series were organised in preparation for the 2021 Winter School and the 2022 Summer School, see below for more information.

1.3.2 Workshops

Workshops in the ENVRI-FAIR sense are typically 90-180-minute-long events that focus on a relatively narrow theme or topic. The meetings combine (introductory) presentations by experts with interactive discussions and/or hands-on activities involving the participants. Workshops benefit from having a set of predefined goals, for example co-creating a shortlist of important points to be worked on towards solving a problem or asking participants to first work individually on a short descriptive text, then discuss it in a breakout group and finally present results to the entire audience.

Workshops on FAIR and RDM topics

Similarly to the case of stand-alone webinars outlined above, WP6 worked with ENVRI-FAIR project participants to organise workshops on a number of different research data management and general FAIR topics, Examples include "Persistent identifiers for instruments" (held on 2021-02-04 as part of the ENVRI Week 2021 training event, with 40 participants) led by Markus Stocker who introduced the Research Data Alliance PIDs for Instruments recommendations, followed by hands-on exercises allowing RI personnel to try out the related metadata model using the B2INST testbed from EUDAT,

and “Introduction to Service Documentation & How to write Service Descriptions” (held on 2023-03-27 with 16 participants) with Maggie Hellström summarising the topic of service documentation before leading the participants in practical exercises towards understanding what to include in the description of a service that is to be onboarded to the EOSC .

The policy workshops

ENVRI FAIR Work Package 6 had the objective of providing training to ENVRI personnel and key ENVRI stakeholder groups on the FAIR principles. The training aimed to implement these principles in RI services and data management activities at the data centre level, evaluate the level of implementation using FAIR metrics, and consider relevant legal and policy requirements.

Task 6.3 focused specifically on providing legal and policy training to RIs on the FAIR requirements. The target audience for this training activity primarily included the system managers of RIs within administrative staff of ENVRI, and those responsible for data policy. Within the context of Task 6.3, a series of four policy workshops has been organised in collaboration with Work Package 4, which aims to establish consistent policies for the RIs in ENVRI to permit interoperation and policies for the ENVRI-Hub catalogue and its usage. The goal also was to extend this to the interaction between the ENVRI-Hub and EOSC. Users accessing the catalogue are able to discover and contextualise the digital assets offered by the ENVRI RIs. From the chosen metadata records, they will be able to access services or other digital assets at the servers of individual ENVRI RIs. It is therefore crucial to have appropriate policies at the RIs to permit and control such access. The main aim is for ENVRI RIs to have control over which of their digital assets are visible and under what conditions (e.g., licensing) when providing metadata to ENVRI-Hub.

The *“First legal and policy training event for ENVRI administrators”* was held online on September 24, 2021 (Milestone 25) https://training.envri.eu/course/view.php?id=55. The workshop, attended by about 40 participants, focused on presenting and discussing general concepts on writing and developing policies in the ENVRI RIs, their aspects and their relationship to the RI operations. In particular, the workshop had two main goals:

1. Discuss “backbone” issues of policies in RIs: Differences between policies, best practices, and technological solutions, what kind of processes are involved in each in a distributed RI. Also, the workshop aimed to make these terms commonly agreed by the RI participants.
2. Get engagement with the policy-involved people in all ENVRI RIs to discuss further the key aspects of the ENVRI Policy framework (e.g., access policies, service policies, PID policies), etc.

Results of the workshop have been used to develop further the ENVRI policy recommendations, used as a basis for training materials on the subject, and as the backbone of more detailed workshops on specific RI policies.

The second ENVRI policy workshop, *“Policies Required by EOSC and FAIR”*, held on February 8, 2022, had a particular emphasis on the recently developed ENVRI Policy Framework and was attended by about 50 participants. This comprehensive framework comprises 31 policy statements that address a broad range of topics relating to data, metadata, and service-related decisions. It highlights the interconnections between these policies and provides guidance on their implementation. The Policy Framework identifies key policy drivers for ENVRI data service providers, drawing on sources such as the EOSC Rules of Participation, the FAIR Principles, and the ENVRI-Hub. It also outlined the requirements that ENVRI RIs policies must meet to ensure compliance with these regulations and indicated a way forward to create policies and guidelines from the policy statements framework.

The third ENVRI policy workshop *“ENVRI-FAIR policy Workshop #3”* was held online on November 30, 2022 https://envri.eu/event/envri-fair-policy-workshop-3/

Workshop 3, which had approximately 20 participants, was designed to build upon the policy concepts and framework established in workshops 1 and 2. The primary goal was to provide examples of policy documents, compare the various approaches adopted by ENVRI RIs, and reach a consensus on the policies necessary for the ENVRI-Hub catalogue. This would ensure that the policies of individual ENVRI RIs are compatible with those of the ENVRI-Hub catalogue.

The *"Fourth ENVRI policy workshop"* was held at ENVRI week 2023 on 1st February. There was a large attendance (>50) and the topic covered a review of the progress made in ENVRI on policies, a repetition of the process to create policies and guidelines, and agreement on the need for the three key policies (personal data privacy, terms and conditions, cookies) to be implemented with a pop-up check box page on the ENVRI-Hub.

1.3.3 Schools

The events that make up the ENVRI Community International Schools series have been quite varied in format, to a large extent depending on the organisational constraints imposed by the Covid-19 pandemic. However, they have in common a focus on how to increase the FAIRness of data and services as provided by and for the ENVRIs. In general, the ENVRI School events have comprised 5-6 days of intense activity (either concentrated to a calendar week or spread out over ca 14 days), combining instructor-led presentations and practical demonstrations with exercises for the participants.

2019 School

International Summer School "Data FAIRness in Environmental and Earth Science Infrastructures: theory and practice" took place July 1-5, 2019, in Lecce, Italy, hosted by University of Salento and LifeWatch Italy. While this event was financed by the ENVRIplus project (precursor to ENVRI-FAIR), from the topic coverage it essentially served as the kick-off for the ENVRI-FAIR training activities.

The school sessions "RDM in a Nutshell", "Findability", "Accessibility", Interoperability" and "Reusability" gave the ca 20+ participants a thorough introduction into FAIR data concepts, terminology and basic underlying research data management technologies. A majority of the participants were early-career scientists, but data centre staff were also represented.

2021 School

Originally this was planned as a summertime 2020 in-person event in Lecce, Italy, but the pandemic restrictions forced us to both go completely online and to entirely rethink the school format. The *ENVRI Community International Winter School on data FAIRness* was thus composed of two components: a series of preparatory webinars that took place in the summer and autumn of 2020, and the main event which took place over a period of two weeks in January 2021.

The preparatory webinars were open to all interested persons (from the ENVRI Community and beyond), and on average gathered 60+ participants who were treated to expert presentations on "Cloud computing and application development", "Workflows Orchestration and Execution" and "An introduction to Jupyter".

The school itself had 30+ participants, a majority of which was staff working at the data centres of the ENVRI Community research infrastructures. Themes covered included "Semantics", "Virtual Research Environments, Data Analysis & Visualisation", "Resource access tools" and "Cloud computing for developing and operating data management services". The pacing of the instruction was adapted to fit the work schedule of these predominantly ICT specialists who had to balance learning activities with their day-to-day tasks.

2022 School

In the summer of 2022, the school returned to Lecce, Italy as ENVRI Community International Summer School 2022: Road to a FAIR ENVRI-Hub - Designing and Developing Data Services for End Users. The main event was again complemented with preparatory webinars, allowing on-site participants to focus on hands-on exercises and discussions related to the webinar topics.

The open preparatory webinars on "Service validation & evaluation" and "Service documentation & tutorials" gathered 30+ participants on average. The in-person School following a few weeks later saw 20 participants representing ENVRI-FAIR project partners engaging with topics including "Service user requirements elicitation", "Creating high-quality service documentation and usage examples to support service end users", "Developing services and fostering reusability/interoperability among them", and "Validating and evaluating your services".

1.4 Preparation & curation of learning materials

Recordings and slide decks

Slide decks can be of varying quality from an educational and informational perspective. Some may be text-heavy or include detailed notes, while others may utilise concise bullet points and visuals such as images. To maximise their effectiveness, especially for individuals who did not attend the presentation, it is advantageous to provide a recording alongside the slide deck. This additional resource enhances the overall usefulness and comprehension of the material.

WP6 has made best efforts to adapt presentation materials according to specific needs, including investing efforts in post-event processing of videos to remove irrelevant footage, enhance findability of specific passages and optimise the overall structure of the recordings. This process aims to strike a balance between maintaining the completeness of the record and minimising potential distractions. Furthermore, it is important to consider the sensitivity of certain project-internal content, ensuring that such parts are not shared with external audiences unless deemed appropriate. Finally, frame rate and screen resolution often needed to be adjusted to bring down file sizes. Many of the resulting videos are available via the LifeWatch ERIC YouTube channel (https://www.youtube.com/@lifewatcheric9455).

Exercises, tutorials, and other texts

Having accompanying textual materials takes the learning experience further, as even a brief summary text from the instructor(s) goes a long way to help convey the messages and reduce confusion or information loss. By recommending preparatory materials (book chapters, articles etc. or recordings) an instructor can both help trainees to structure their learning and save time that can be better used during an event. Easy-to-follow instructions for exercises or "homework" also adds value as such activities help build new knowledge and strengthen retention of skills.

It is important to make sure that all relevant materials are gathered, organised, annotated and properly curated for post-event dissemination, and WP6 has worked closely with instructors and trainers to ensure that this has been done expediently after each event.

Learning paths and learning outcomes

As mentioned already in Chapter 1.2, a set of learning paths were developed with the aim to demonstrate how selected learning resources and associated materials currently indexed in the ENVRI Community's Training Catalogue and Learning Platform can be utilised by learners with specific learning goals in mind. Each path is aimed at a different target audience and comprises 3-5 learning resources that together should provide learners with a good introduction to the given theme. As part of each path description, we indicate a recommended order of progression, as well as associated learning outcomes, both at path and component levels. The paths are "FAIRify yourself" (for researchers and end users of ENVRI), "FAIRify your data" (for the producers of data and related services) and "FAIRify your infrastructure" (for RI administrators and managers). For details, including path components and associated learning goals, please see Appendix 3.

1.5 Implementing the ENVRI Community learning environment

A priority of WP6's activities has been to not only create training materials on FAIR-related topics, but to also make our own learning resources as FAIR as possible. A central point has therefore been to develop a complete Learning Environment for the ENVRI Community. The original plans for this, as outlined in the WP6 Description of Action, was to have three components: a catalogue for learning resources, a community learning platform, and an underlying repository for curating materials. The successful completion of the corresponding implementation work was reported on in D6.2 ([Vaira 2019])

However, following the discussions with the external experts during the midterm review, we came to realise that the design and instantiation of the first version of the Environment did not allow for hosting easily accessible introductory and descriptive information about the WP6 activities and outputs. We therefore reached out to WP5's ENVRI-Hub development team to create the Training Gateway, which can be considered to be a fourth component of the Learning Environment.

Training catalogue

The catalogue (https://trainingcatalogue.envri.eu/), based on the international IEEE standard for learning resources, is operated by LifeWatch ERIC on behalf of the ENVRI Community. It is a FAIR training service that facilitates findability (for both humans and machine processes) and the reuse of the materials. At the time of writing, it holds metadata on close to 50 learning resources, of which half were created by us during the ENVRI-FAIR project and the other half have been harvested from external sources and catalogues.

The catalogue implementation and the underlying metadata model have received much attention from outside of the project, including from the FAIRsFAIR and EOSC Future projects as well as the Research Data Alliance interest group on education and training on handling research data (See also the Impact chapter below).

In the time after the launch of the first instance of the Catalogue in 2020, testing by WP6 members and others resulted in a comprehensive review and update of the underlying metadata model, which has been customised in order to be compliant with the EOSC Training Resource Profile - Data Model. In particular, a deep analysis and comparison between the two metadata models has been performed in order to identify the mandatory metadata fields with the appropriate label and value space. Appendix 4 contains a detailed description of all the metadata fields of the so-called ENVRI LOM (Learning Object Metadata model). For more details about the technical specifications of the Catalogue and its underlying database, as well as the motivations for the selection of metadata model fields, please see deliverable D6.2 ([Vaira 2021]).

Learning platform & repository

The ENVRI Community Learning Platform (https://training.envri.eu/) is a fully-fledged Learning Management System (LMS) based on Moodle. It is operated by LifeWatch ERIC on behalf of the ENVRI Community. The Platform allows easy creation of individual items that can either just contain descriptive metadata and links to relevant learning resources (slides, recordings, documents) related to an event, or be configured with plugin modules for, e.g., quizzes, discussion forums and student material uploads and grading. To support the Learning Platform, a dedicated secure storage space has been allocated to hold copies of training materials.

Training Gateway component of the ENVRI Hub

Because of technological limitations of the software platforms selected for the catalogue front end and the community learning platform, it proved challenging to use these to disseminate general information about learning resources, the training topics classification scheme and pedagogical aspects such as learning paths and associated learning objectives. We initially relied on the ENVRI web site (https://www.envri.eu and specifically https://envri.eu/training/) to disseminate news about and general descriptions of training events (webinars, workshops, and schools). However, feedback from learners suggested that they struggled to locate relevant information.

As the development of the ENVRI-Hub progressed, WP6 therefore took the opportunity to develop the Training Gateway. This Hub component is intended to offer a "first stop" entry point to all ENVRI training and capacity building activities and materials in a sustainable manner also after the end of ENVRI-FAIR.

The Training Gateway consists of a simplified search interface that allows a user to enter words in the resource title or free-text keywords. Further filtering of search results can be made by selecting from a list of training topics, or by choosing one of the three predefined learning paths. To facilitate the search experience as well as provide relevant background information, a number of “About…” web pages are provided, easily accessible via “info button” symbols.

Finally, for users wishing to make use of more complex queries, links are provided to the search interface and informational pages of the underlying ENVRI Training Catalogue. There, instructions for using the catalogue’s REST API can also be found, supporting owners of external catalogues and others who wish to harvest the catalogue entries programmatically.

2 Lessons learned

Through their participation in skills building and training activities during previous projects including ENVRIplus, the members of WP6 came into ENVRI-FAIR with considerable experience of creating, curating and providing training. However, while these lessons from previous activities have been learned and taken into account, the rapid evolution in community expectations, online training platform technologies and governance that have taken place during the project lifetime have resulted in us identifying many new challenges. In the following we outline and discuss the most important of these.

2.1 Ensuring relevance & timeliness

It was important for WP6 to not only follow our own original plan - based on the initial gap analysis - but to also continuously engage with other Work Packages to learn of their current status ideas and needs for training and be able to respond quickly to develop or facilitate training. One example of a "spontaneous" initiative concerned data provenance. Here a discussion in one of the ENVRI subdomain working groups resulted in a wish to organise some webinars on provenance and how to trace it effectively. This was passed on to WP6 which was able to quickly facilitate a multi-event series of presentations and a workshop, open to all of the subdomains, thus broadening the target audience of the knowledge exchange to the whole project. These training events were led by a topic expert together with representatives of ENVRIs wanting to share their experiences with collecting and curating provenance information.

Capturing the training & skills building needs of the ENVRI-FAIR subdomains

As briefly mentioned above (in section 1.2), one of the first actions of WP6 was to organise a gap analysis in order to map out which FAIR and data management training topics that the project partner RIs found to be relevant. We then asked the RIs to indicate their current level of knowledge on the 6 "general FAIR" and 16 "RDM for FAIR" topics identified, and to which of these they assigned the highest priority, hoping to obtain a solid knowledge of needs and requirements on which to base the WP6 activities.

However, we noted that it was challenging to obtain informed and comprehensive responses, especially from those RIs that were very distributed in their organisation or in a relatively early stage of maturity. Partly this was due to the RI representatives responding to the gap survey often being quite unfamiliar with the skills and experience levels of e.g. their own organisations' data centre staff. In addition, some of the subdomains provided more responses than others, somewhat skewing the analysis. However, it was nevertheless possible to identify a common set of prioritised learning topics.

We draw the conclusion from this that it is very important for ENVRI Community training organisers to ensure that their contact points in the individual RIs are able to give accurate and relevant information on the status in their respective organisations; in addition, the contact points may change over time. Also, low or subdomain-skewed response frequencies will hamper the ability to extract statistically significant conclusions or trends and may limit evaluators to make mostly qualitative conclusions.

Collecting & leveraging individual participant feedback

Finding out useful information about what participants in a training activity actually felt about the contents, delivery and relevance may sound easy at first glance, but is in practice quite challenging. A clear prerequisite is having a clear picture of what aspects to collect feedback on, and then to apply the same framework of questions in a consistent manner. WP6 chose to apply two online questionnaires: a simple one for webinars and workshops, based on a template developed by RDA Europe, and a more detailed one for Schools, following a set of questions set up for similar LifeWatch training events. Appendix 5 contains a sample webinar feedback form, showing both the questions and the responses received.

But no matter how useful and to the point the questions are, feedback surveys fail if not enough participants engage. The WP6 experience has unfortunately been that the average response frequencies have unfortunately been very low, typically below 50%. This has been seen regardless of e.g., creating easy-to-use "short links" (such as bit.ly URLs or QR codes), and/or sending out several reminders, for example in connection with informing participants of recordings and other materials having become available post-event. Trying other approaches, such as Including a few questions on usefulness,

applicability etc. in the form of Mentimeter polls shown at the end of events would probably elicit a much higher response rate, but this method cannot be expected to give the same depth of feedback as a questionnaire can.

Engage with ENVRIs as training providers

Some of the ENVRIs are not only consumers of training and skills building, but operate their own programmes and initiatives. Although they are many times receiving funding to host and curate related materials and information, they can increase the discoverability and simultaneously also the FAIRness of their learning resources by exposing these through community or regional training catalogues or knowledge hubs. Here the ENVRI Community's Training Catalogue can play an important role, especially for those RIs that do not have similar resources available through their typical subdomain contexts and collaborations.

2.2 Organising training events efficiently

Virtual events vs physical meetings

The Covid-19 pandemic has had a significant impact on travel, resulting in the widespread adoption of 100% virtual events in 2020-2021 and beyond. This shift has had ripple effects on the structure and duration of training events. We have placed a strong emphasis on interactivity, including Q&A sessions, group work through breakouts, polls, and quizzes. In addition, to combat the potential loss of attention and prevent "zoombieism", we have incorporated frequent breaks to keep participants engaged and focused.

However, it is very difficult to create a fully welcoming and social environment online that encourages the spontaneous exchanges and discussions among learners, or between learners and instructors, that are so characteristic of physical training events. During the virtual School of 2021 we set up a virtual social area using the Wonder platform, in which people could interact throughout the entire event, for example for coffee break, group work and discussions with instructors and group mentors. While this functionality was used to some extent, the interactions were nowhere near as spontaneous as what was typical for the 2019 and 2022 in-person Schools.

Communication & publicity

Timely and effective advertisements and dissemination is often crucial in the realm of training events in the realm of training events. Training organisers need to maintain a direct line to communicators and be aware of the existing channels of communication, including social media, newsletters and means of announcing events and news on the project's website. Ideally draft publicity announcements for events should be prepared well in advance (at least 2-3 weeks for webinars and workshops, 2-3 months for schools) and sent to the communicators.

However, the information exchange must also be bi-directional to avoid confusion over event titles, target groups, etc. It is especially important to agree beforehand on who has the responsibility for participant registration, (virtual) meeting hosting, materials dissemination and feedback collection. As we learned first-hand, it is especially important to ensure that preparations for and management of event recordings are managed with great care.

Event organisational workflow & checklists

Having well-established workflows and checklists in place for different types of events and learning materials management is crucial to an effective and smooth delivery of training, allowing trainers and organisers alike to focus on the content and best way of delivery.

To standardise the training event planning process, the WP6 core team suggested a process consisting of four phases. In the first phase, calendar planning would be conducted, including venue identification and defining registration deadlines. Target groups, themes, access rules, and goals should also be established. Instructors can then be invited, and a program proposal with support tools and materials should be created. A registration form should also be implemented. In the second phase the organisers should focus on training event promotion through newsletters, social media, project websites, and email promotions.

Registration collection and monitoring can be then carried out, with additional promotion cycles if needed. In the next phase the organisers will have to share connection details and upload relevant materials to the ENVRI Training Platform giving access to the registrants. After the training event, the organisers should distribute the feedback form and collect feedback from the participants for further analysis which can be used to improve future events. In the last phase the metadata of the training event is registered in the ENVRI Training Catalogue creating a new learning object.

2.3 Pedagogical challenges

Adapting to pandemic conditions

The most difficult challenge to organising the training and capacity building activities during ENVRI-FAIR was adapting and adjusting to the global pandemic caused by the coronavirus. In early 2020, everything on very short notice had to be adapted to the fact that travel and person-to-person contacts became almost impossible, and all plans for schools and workshops had to be reworked to fit an online-only format. While the restrictions for international travel started to be lifted in spring of 2022, WP6 continued to organise its training events mostly in hybrid mode, at least giving instructors the option to connect remotely. This situation continued until the end of the project.

Changing to an online-only format had different effects depending on the type of training. Obviously asynchronous self-study of written materials or recordings is not affected per se, but the fact that at least for some ENVRI staff the daily routine suddenly became cluttered with numerous video meetings brought along the unforeseen consequences of mental overload, screen fatigue and "zoombieism" which are all detrimental to the concentration required for effective, deep learning.

For workshops and schools, when planning these it soon became evident that activity blocks had to be drastically shortened, with plentiful breaks added at least hourly. In addition, exercises and other assignments often had to be split up into smaller parts (as well as over longer periods, in the case of schools) to facilitate for students to organise their work. Finally, the fact that some participants actually fell ill with covid-19 during School activities meant that they were unable to complete the courses as planned, which both impacted their own learning and that of other students as group work was disrupted.

Adapting to new audiences

An aspect that is often forgotten, or at least partially ignored, when designing training and related materials in the context of RIs and similar research performing organisations is that the target audiences are only rarely (very) early-career scientists from academic environments. Instead, just like ENVRI-FAIR WP6, the majority of potential learners are adult professionals who are participating in training as part of their journey of life-long learning. Not only do their studying motivations, contexts and environments tend to differ from those of (master or graduate) university students, but they also bring along experiences and non-academic knowledge that can be very useful - both to themselves, fellow learners and instructors. As a consequence, the contents, formats and teaching methods must be adjusted, always keeping in mind that topics and presentations must be felt to be of direct relevance and hence worth the time and effort to engage with.

During ENVRI-FAIR, the main focus groups for our training activities have been the ENVRIs' data centre staff and administrators/managers, followed by researchers belonging to the ENVRI Community's traditional end user groups. This differed from e.g. ENVRIplus, where training was mainly targeting early-career scientists in and beyond the ENVRI Community.

For data centre and management staff, the contiguous amount of time that for example a developer of data services or a head office strategist can dedicate to training is often short and must be fitted into what can be very complex work schedules. In addition, study techniques and related practical learning tasks may have been forgotten and need brushing up. Researchers on the other hand, regardless of whether they fork in an RI or outside in academia, may find it easier to fit training into their schedules, but choose courses based on how quick they can find answers to specific questions or learn a new skill. People from outside of the ENVRI Community could benefit from those parts of the training that cover general aspects of FAIRness and FAIRification but would likely struggle with content that is focussed on ENVRI-specific contexts and needs.

Experts as instructors

Experts are not necessarily used to teaching, especially not in "non-academic" settings (outside of higher education institutions, or HEIs). It is beneficial in all situations to organise a briefing session ahead of time, especially in preparation for school events, where the schedule - with time limits! - can be clarified, hints for how to best present e.g. coding or demonstration are shared, and concrete suggestions for e.g., engaging the event participants with interactive contents can be made.

A special challenge concerns the inclusion of "live" demonstration blocks during expert lectures, especially when instructors want to show examples of programming or coding. To do this successfully requires detailed planning and a good knowledge of the skill level of the learners. Sharing "live" and interactive desktop or browser windows in a video meeting is extremely difficult. Font type, size and colour must be specially chosen for maximum legibility. The speed of typing and accompanying commentary must be slowed down, especially if students are expected to follow along on their own computers. Optimal to prepare a video recording, either in advance (for people to prepare ahead of the training event) or for later (re-)viewing. Also recommended to create a step-by-step tutorial document covering the most important steps.

2.4 Dissemination of materials: FAIR or not?

Ensuring optimal Findability of learning resources

While it is convenient to create both Training Catalogue and Learning Platform entries for complex multi-part events such as Schools or Webinar Series, we have come to realise that in a majority of cases this is not enough to ensure that people can Find, Access, and Reuse the individual parts - for example in the contexts of different Learning Paths. Instead, it is necessary to set up and maintain "parallel" catalogue & platform entries for those course or event segments. It is much easier and less time consuming to prepare for and do this "splitting" already from the beginning. In addition, one must remember to indicate qualified references such as "is part of" or "has parts" in the metadata of the concerned learning resources.

Controlling Access to materials

Access to learning resources should be as easy and convenient as possible for all interested parties. The default for all learning resources created by WP6 with ENVRI-FAIR partners is open access, but in some cases all or parts of the materials related to given training events and activities needed to be controlled - for example by requiring users to register accounts on the Learning Platform and have these configured with appropriate access rights. Examples include recordings and slides from project-internal workshops where non-public information was shared, as well as participant-created documents from schools which are produced as part of individuals' personal learning processes.

Obstacles preventing Interoperability of resources

An important aspect of learning resources is their context dependence, i.e., whether they can be used as stand-alone entities or rather need to be considered as components of a more complex matrix. An interesting intermediate position is taken by materials developed as parts of Schools; these typically form lessons that cover certain aspects of the overarching theme for the whole event, but at the same time they comprise an introductory lecture, interactive exercise materials and often participant assignments. When these are only curated and disseminated as parts of the whole event package, the individual lesson parts can be difficult to identify and retrieve. That not only lowers their findability and accessibility, but also makes it complicated to repackage and incorporate them into other courses and activities, i.e., their Interoperability suffers.

Promoting fair and FAIR Reuse

Just as with all other types of digital objects, learning resources should be assigned a suitable licence and relevant Terms of Use. These should make clear what limitations, if any, are imposed on who can reuse the materials, in what context and for what purpose. Examples of common licences include Creative Commons Zero (no limitations) and Creative Commons By Attribution (no limitations for reuse, but the creator must be attributed).

3 Impacts

3.1 On the project

Through the training activities that WP6 organised, staff in the project partners have been offered means to extend their knowledge of both the FAIR principles as well other basic research data management practices, and existing and emerging technologies that can be implemented in the ENVRIs to enhance the interoperability and sustainability of their data repositories and the services built on these. Based on the event-by-event feedback received, the majority of participants expressed high levels of satisfaction with regard to the content, level of difficulty, and delivery style of the training sessions.

As a follow-up to the gap analysis performed at the beginning of the project, during the 2023 ENVRI WeekWP6 organised a feedback collection round with representatives from all the four subdomains. Poll participants were asked to rate how well their RI thought that the training offered during the project had addressed high-priority topics, whether the training event formats (webinars, workshops, schools) had been suitable, and what training topics would they like to see offered in the future in a common ENVRI Community training context. While it came to light that only some of the individuals representing the project partners had themselves been participating in the training events, the responses indicated that RI personnel who had received training were positive to the balance of topics presented and the overall training program offered by WP6.

3.2 On stakeholders

While the training activities and learning resources developed during the project were not specifically targeted at the stakeholders of the ENVRIs, we believe that the increased awareness achieved within the ENVRIs regarding the importance of making environmental and Earth science outputs as FAIR and Open as possible will have positive and long-lasting impacts on them as well. Thus, our funders who play a crucial role in make the work of ENVRI possible, our research partners and collaborators in governmental agencies and across academia, as well as citizens and society at large will take notice of these advancements

3.3 Beyond ENVRI

FAIRsFAIR

Members of WP6 were invited to actively participate in several training & learning-related events organised by the FAIRsFAIR project, including a workshop on learning resource metadata models and co-authorship of a book, *How to be FAIR with your data* [Engelhardt 2022].

EOSC Future project

Partners of ENVRI-FAIR (LifeWatch ERIC, ICOS) have had a strong involvement in EOSC Future WP9 ("Skills & Training"), including leading and contributing to Task 9.3 which produced the specifications for the EOSC Knowledge Hub. This work benefited very much from the experiences gained from the building up of the LifeWatch/ENVRI training catalogue that was undertaken in the ENVRI-FAIR context. Conversely, the ENVRI Training Catalogue model was adjusted and updated in 2023 to improve the alignment with the EOSC Knowledge Hub, thus facilitating future harvesting of learning resource metadata from our catalogue. In addition, materials from the ENVRI Community International 2022 Summer School are being successfully adopted by EOSC Future WP9 into a course for EOSC service providers on writing effective service documentation, illustrating that cluster project training activities have high relevance across research disciplines.

Research Data Alliance Education and Training on Handling Research Data (ETHRD) Interest Group

As individuals several WP6 members played very active roles in the work to create minimal and extended metadata profiles for learning resources, as well as standards for documenting training material repositories. This activity - which brought much positive attention to the ENVRI version of the Learning Object Model (the "ENVRI LOM") that was developed in the ENVRI-FAIR context - resulted in several RDA outputs including *Recommendations for a minimal metadata set to aid harmonised discovery of*

learning resources [Hoebelheinrich 2022a] and *Core Characteristics of Learning Resource Collectors* [Hoebelheinrich 2022b].

4 Conclusions & outlook

During ENVRI-FAIR it became evident that the effectiveness of FAIR and research data management education relied on several key factors. These included the provision of relevant and timely training that addressed the specific needs of different subdomains, while actively gathering feedback from participants and involving ENVRIs. WP6 achieved significant milestones and delivered various training events and valuable resources to enhance training and learning experiences for the project participants, implementing communication strategies and well-defined workflows.

The implementation of the ENVRI Community Learning Environment, including the training catalogue, learning platform, repository, and ENVRI-Hub Training Gateway, aimed to make the training materials as FAIR as possible and facilitate findability, accessibility, and reuse of the resources. The catalogue and learning platform provided a variety of training resources that were collected or developed to support the ENVRI training activities and capacity building within the project. These resources were not limited to the FAIR principles but helped increase knowledge in ENVRIs about standards, technologies, and collaborative services for the future FAIR environmental research landscape.

The ENVRI learning resources are expected to have long-term benefits, not only for the ENVRI Community but also for the wider EOSC user context. To ensure their FAIRness and usability, the ENVRI-FAIR project dedicated efforts to defining clear learning outcomes, identifying suitable learning paths, and ensuring the pedagogical quality of the materials. With the learning paths designed by WP6, structured and cohesive educational journeys for learners are provided, focusing on specific skills or knowledge areas while being designed to align with desired learning outcomes. By following these paths, learners can progress gradually and gain a deep understanding of the topics while developing the necessary competencies to apply their knowledge effectively.

All training digital assets are offered to the ENVRI-Hub users via the ENVRI-Hub Training Gateway, the entry point to training activities and materials for both the ENVRI Community and wider audiences. The Training Gateway, being a dynamic platform, can be continuously updated and enhanced to meet the evolving needs of users, ensuring a sustainable and future-proof training ecosystem on the Hub. Moving forward, ongoing collaboration with ENVRIs, continuous adaptation to community needs, online platform advancements, and robust governance will be vital for future training endeavours.

5 Acknowledgements

The authors wish to acknowledge the support and contributions from all instructors and participants in the training events arranged under the framework of ENVRI-FAIR - together we made the ENVRIs a FAIRer community! We also thank our extended WP6 members representing the subdomains, all training contact persons in the ENVRI Community, and the core team members who moved on to other duties - without you we couldn't have completed our goals!

6 Bibliography

[Engelhardt 2022] Engelhardt, C., R. Barthauer, K. Biernacka, A. Coffey, R. Cornet, A. Danciu, Y. Demchenko, S. Downes, C. Erdmann, F. Garbuglia, K. Germer, K. Helbig, M. Hellström, K. Hettne, D. Hibbert, M. Jetten, Y. Karimova, K. Kryger Hansen, M. E. Kuusniemi, V. Letizia, V. McCutcheon, B. McGillivray, J. Ostrop, B. Petersen, A. Petrus, S. Reichmann, N. Rettberg, C. Reverté, N. Rochlin, B. Saenen, B. Schmidt, J. Scholten, H. Shanahan, A. Straube, V. Van den Eynden, J. Vandendorpe, S. Venkataram, A. Vieira, C. Wiljes, U. Wuttke, J. Yeomans & B. Zhou. (2022). How to be FAIR with your data. Universitätsverlag Göttingen. https://doi.org/10.17875/gup2022-1915

[Hellström 2019] Hellström, M., Johnsson, M., Konijn, J., Fiore, N., Quimbert, E., Boulanger, D. & Baker, G.R. (2019). ENVRI-FAIR deliverable report D6.1 Inventory & gap analysis of FAIR training materials (Version 1). Zenodo. https://doi.org/10.5281/zenodo.3885122

[Hoebelheinrich 2022a] Hoebelheinrich, N. J., Biernacka, K., Brazas, M., Castro, L. J., Fiore, N., Hellström, M., Lazzeri, E., Leenarts, E., Martinez Lavanchy, P. M., Newbold, E., Nurnberger, A., Plomp, E., Vaira, L., van Gelder, C. W. G., & Whyte, A. (2022). Recommendations for a minimal metadata set to aid harmonised discovery of learning resources (Version 1.0). Research Data Alliance. https://doi.org/10.15497/RDA00073.

[Hoebelheinrich 2022b] Hoebelheinrich, N. J., Newbold, E., Whyte, A., Johnsson, M., Helbig, K., Hellström, M., Kuchma, I., & Jetten, M. (2022). Core Characteristics of Learning Resource Collectors (Version 1.0). Research Data Alliance. https://doi.org/10.15497/RDA00072.

[Vaira 2021] Vaira, L., Manca, M.T., Vallo, C., Fiore, N. & Konijn, J. (2021). ENVRI-FAIR deliverable report D6.2: FAIR training materials catalogue & integration with Common Training Platform (Version 1). Zenodo. https://doi.org/10.5281/zenodo.5052426

7 Appendix 1: ENVRI-FAIR Training Topic Codes

In this appendix we list and briefly explain the training topic codes, as defined in the D6.1 deliverable report ([Hellström 2019]). These codes are used in the Training Catalogue and have been implemented as a filtering parameter in the ENVRI-Hub Training Gateway search form.

General FAIR-related training topics

Introduction to FAIR principles (G1)

Covers basic explanations of the FAIR (Findable, Accessible, Interoperable, Reusable) guiding principles, and provides best practices that facilitate data producers, data repositories and data end users to make sure their data & metadata achieve a high degree of FAIRness.

Metrics for FAIRness evaluation (G2)

There exist many approaches to estimate the degree of compliance of a given (data) object with the FAIR principles, some based on subjective evaluations by human experts and others relying on machine-actionable "automatic" tests.

Performing a FAIRness self-assessment (G3)

Once one or several metrics for FAIR assessment have been chosen, a RI needs to decide how and when to collect information, with whom to share the results, etc.

GDPR (General Data Protection Regulation) issues related to data sharing (G4)

ENVRIs produce and manage many types of data and metadata, some of which contain personal information. The European GDPR covers what personal data are, what may be collected and for what purposes, where and for how long to store the collected information.

Basic Research Data Management (RDM) (G5)

Basic research data management covers introductions to the different "pillars" and "cross-cutting beams" defined by the ENVRI-Plus project:

Writing (technical) documentation for services (G6)

Clear and comprehensive documentation is absolutely essential for a service to be discoverable and successfully used. Writing effective documentation starts with understanding the target audience(s) and addressing the aspects that are important to them in a relevant format - for example description fields in catalogues, promotional materials, usage instructions and tutorials, technical specifications and richly commented code.

Other general FAIR-related topics (G7)

A catch-all classification, covering any other FAIR topic that isn't directly connected to concrete implementations of, or technologies for, data management components.

Research Data Management (RDM) training topics

Access control (Authorisation-Authentication-Identification, or AAI) methods (R1)

Providers of services must be able to control who has access to them in general, and what functionalities or resources they can use. To this end, mechanisms need to be in place that allow users to state who they are (identification), if needed backed by some proof of their identity (authentication), and then grant access to relevant resources based on this authenticated identity (authorisation).

API (Application Program Interface) design for data & metadata access (R2)

Application programming interfaces (APIs) are software interfaces designed to facilitate computer-to-computer communications, for example between a repository server hosting data and related metadata, and a script running on the computer of an end user. In the ENVRI context, APIs offer an important means to enhance interoperability between services provided by different subdomains.

Cataloguing - design & implementation (R3)

Many of the ENVRIs operate repositories of their data holdings - combining storage with a catalogue that contains all relevant metadata describing the digital objects. It is important to build the catalogues

on relevant standardised metadata schemas and models, in order to achieve maximum findability and interoperability.

Certification schemes for repositories (CoreTrustSeal) (R4)

Repositories rely on their end users having trust in the management, storage and overall quality of their data and metadata holdings. There exist several expert bodies and initiatives that can evaluate research data repositories, including CoreTrustSeal.

Cloud computing (Virtual Machines & containers) for data processing (R5)

Cloud computing utilises networked resources (storage, computation power etc.) typically provided on-demand by e-infrastructures or other large service providers. Users do not have to install software or work on their local machines, but instead connect to cloud services that often are operated in virtual machine environments that can be scaled to meet demands.

Data Management Plans (R6)

Data management plans (DMPs) are formal documents describing how data and other digital assets are to be handled during and after research projects. The contents and level of detail of a DMP may vary between funders, research disciplines and regions, but there are templates and instructions available to guide researchers through the preparation and maintenance phases.

Landing page design (R7)

In Research Data Management contexts, a "landing page" is a web page containing relevant metadata describing a digital object such as a dataset. In the case that the digital object has been assigned a persistent identifier (PID), resolving the PID will redirect the user to the landing page., rather than to the bitstream representing the object itself. Designing landing pages requires knowledge on both the metadata itself and the ways to encode this in a way that is machine-interpretable and -actionable.

Licences & policies for data use (R8)

Any (digital) asset that is intended for reuse or sharing should be assigned an appropriate usage licence, for example a Creative Commons one. In addition, data producing organisations should set up and define clear policies that describe what it is doing to data throughout all parts of the research data lifecycle, as well as by end users.

Linked Data and ontologies (R9)

Linked data (LD) is a way of expressing information in a way that is directly interpretable and actionable by machine-based processes, for example via subject-predicate-object triples that can be serialised using Resource Description Framework (RDF) statements. Linked Data is one of the core pillars of the Semantic Web, which makes use of ontologies, vocabularies and thesauri.

Metadata standards & schemas (including geospatial, instruments, variables) (R10)

Being able to identify and implement relevant metadata standards and schemas both on subdomain and higher levels is a crucial first step towards achieving interoperability of data products and services from the ENVRIs. Once these are in place, qualified references and crosswalks can be defined to link together terminologies from the different domains.

PID allocation & use (including citation support, bibliometry, provenance) (R11)

Globally unique and resolvable persistent identifiers (PIDs) play central roles throughout the entire research data lifecycle, and are essential for FAIR research data management. They provide identity, indirection & machine actionability of digital objects, supporting search, referencing and accurate citation. PIDs can be assigned to any digital object, not just data.

Portal design & operation (R12)

Many of the ENVRIs operate their own portals where their data and related services are made available to human and machine-based users. To be optimally useful to the intended end user communities, the design of a portal's functionalities and user interfaces should be based on elicitation of user requirements.

Provenance tracing (R13)

To enable trustability, reusability and reproducibility of research objects, it is necessary to collect and store detailed information on how they were produced, by whom and when. This lineage is referred to as provenance.

Repository design, operation & sustainability (R14)

A repository for digital objects is typically based on a storage system, a catalogue, and some user interface. When designing and implementing a repository at data centre level, the most sustainable and easy-to-operate options are building on existing commercial or open-source solutions.

Virtual Research Environments for data analysis (design & implementation) (R15)

A Virtual Research Environment is a digital platform that offers end users tools and storage for the analysis of data. Importantly, VREs often provide functionality that supports collaboration between researchers, as well as procedures for proper management of results and outcomes.

Workflow engines for automated data processing (R16)

A workflow is defined as an orchestrated and repeatable pattern of activity. In the research data context, workflows are typically built up from software components that are controlled by scripts that can run analysis code, capture provenance information and produce new outputs.

Other RDM training topics (R17)

A catch-all classification, covering any other RDM topic that relates to concepts or technologies used to process or manage digital assets.

8 Appendix 2: Training Events Organised by ENVRI-FAIR WP6

Table 1: Training events organised by WP6, in chronological order

Date	Name/title	Link(s)	Event type	Format
2019-07-01	International Summer School for Environmental and Earth Science Infrastructures "Data FAIRness"	Training Catalogue	School	F2F
2019-12-13	How FAIRsharing can help FAIRify your standards, databases and data policies	Not available	Webinar	online
2020-02-05	Terminologies for ENVRIs: why, what and how (ENVRI Week 2020 training event)	Training Catalogue	Workshop	F2F
2020-06-03	FAIR the smart way: Introducing the ENVRI Knowledge Base	Training Catalogue	Webinar	Online
2020-07-13	Cloud computing and application development for research infrastructures (Towards ENVRI Community International Winter School DATA FAIRness)	Training Catalogue	Webinar	Online
2020-07-14	Workflows Orchestration and Execution (Towards ENVRI Community International Winter School DATA FAIRness)	Training Catalogue	Webinar	Online
2020-09-22	An introduction to Jupyter (Towards ENVRI Community International Winter School DATA FAIRness)	Training Catalogue	Webinar	Online
2020-11-19	Training Provenance Tracing in ENVRI Research Infrastructures: Introductory course	Training Catalogue	Webinar	Online
2020-12-02	Training Provenance Tracing in ENVRI Research Infrastructures: Technical demonstrators	Training Catalogue	Webinar	Online
2020-12-09	Training Provenance Tracing in ENVRI Research Infrastructures: Technical demonstrators	Training Catalogue	Webinar	Online

2020-12-16	Training Provenance Tracing in ENVRI Research Infrastructures: Technical demonstrators	Training Catalogue	Webinar	Online
2021-01-11	ENVRI Community International Winter School on DATA FAIRness	Training Catalogue	School	Online
2021-02-04	PIDs for instruments (ENVRI Week 2021 Training Event)	Training Catalogue	Workshop	Online
2021-02-04	GDPR in Context - Introducing for Research Infrastructures (ENVRI Week 2021 Training Event)	Training Catalogue	Webinar	Online
2021-02-18	Training Provenance Tracing in ENVRI Research Infrastructures: Bring your own provenance cases	Training Catalogue	Workshop	Online
2021-09-24	1st ENVRI-FAIR policy workshop for the ENVRI community Research Infrastructures	Training Catalogue	Workshop	Online
2021-09-27	ENVRI Community International School: Services for FAIRness	Training Catalogue	School	Online
2022-01-25	Making the ENVRIs “FIP for purpose” through FAIR Convergence (ENVRI Week 2022 training event)	Training Catalogue	Workshop/ Webinar	Online
2022-02-08	2nd ENVRI-FAIR policy workshop for the ENVRI community Research Infrastructures (ENVRI Week 2022)	Training Catalogue	Workshop	Online
2022-03-17	ENVRI-FAIR I-ADOPT Framework Training - event 1	Training Catalogue	Workshop	Online
2022-03-31	ENVRI-FAIR I-ADOPT Framework Training - event 2	Training Catalogue	Workshop	Online
2022-06-17	Service validation & evaluation: making sure your services are up to the task (webinar associated with the ENVRI Community International School 2022)	Training Catalogue	Webinar	Online
2022-06-23	Service documentation & tutorials: rolling out the red carpet for end users (webinar associated with the ENVRI Community International School 2022)	Training Catalogue	Webinar	Online
2022-07-10	ENVRI Community International School 2022	Training Catalogue	School	F2F
2022-11-30	3rd ENVRI-FAIR policy workshop for the ENVRI community Research Infrastructures	Training Catalogue	Workshop	Online
2023-02-01	4th ENVRI-FAIR policy workshop for the ENVRI community Research Infrastructures (ENVRI Week 2023)	Training Catalogue	Workshop	Hybrid
2023-02-01	Eliciting service requirements workshop (ENVRI Week 2023 training event)	Not available	Workshop	Hybrid
2023-03-27	Introduction to Service Documentation & How to write Service Descriptions	Not available	Workshop	Online

9 Appendix 3: Sample Learning Paths

Here we present details around the three sample Learning Paths that we have implemented into the ENVRI-Hub Training Gateway (In the search form at https://envri-hub.envri.eu/training select the appropriate Learning Path in the filter box, and click Search).

Learning Path #1: FAIRify yourself

- ***Introduction to the FAIR principles for RI staff***

Target group: all staff, but with focus on the researchers who 'generate' data.
Learning outcome: Along this path a number of training resources are presented that gives you an introduction to the FAIR principles and presents an overview why FAIR data are important for research and researchers. In steps along the path, you will obtain detailed information about the so-called Data Life Cycle which is only a circle when data is made accessible for re-use. The path ends with material from the ENVRI-FAIR summer school on FAIRness: how you can test if your data are FAIR enough.
Components in suggested learning order:

1. "FAIR for beginners" from the Danish e-Infrastructure Cooperation (DeIC): you will obtain an overview of the FAIR principles
   Learning outcome: you will be able to present an overview of the FAIR principles.
2. "FAIR Data Training - ANDS" from the Australian National Data Service: repeats the principles of FAIR data, but presents also how the principles can be implemented with many examples from different resources.
   Learning outcome: you examine further the principles of FAIR data, discussing their implementation with many examples from different resources.
3. "DataOne Best Practices of Data Management" from the US DataONE collaboration
   Learning outcome: You will be able to discuss the Data Life Cycle and its most relevant aspects, connecting FAIR principles to data management.
4. "International Summer School Data FAIRness in Environmental and Earth Science Infrastructures: theory and practice". From the material only the following lecture presentations are relevant:
   a. Data Fairness by Erik Schultes
   b. Assessing FAIRness by Margareta Hellstrom and Barbara Magagna
   c. ENVRI-FAIR Procedure to extract FAIR Answers from FAIR Questionnaires by Barbara Magagna
   Learning outcome: you will be able to illustrate concepts and solutions to evaluate data FAIRness

Learning Path #2: FAIRify your data

- ***Principles, technologies and recommendations on how to achieve FAIRness in practice***

Target group: IT staff, developers and researchers who manage and process data
Learning outcome: You will be able to present principles, approaches and frameworks to support learners and professionals in discussing and adopting solutions, technologies and recommendations to enrich data resources, advance their FAIR implementation and ultimately achieve FAIRness in practice.
Components in suggested learning order:

1. "ENVRI Community International Winter School on data FAIRness" (ENVRI FAIR Project): Subsections of interest
   1. Semantics;
   2. VREs, data analysis and visualisation;
   3. Data access tools;
   4. Cloud computing for developing and operating data management services

   Learning outcome: After examining semantic navigation, Jupyter environments for visualisation and data discovery, resource access tools and cloud computing, you will be able to discuss solutions and technologies to enrich data resources and advance their FAIRness.
2. FAIR Implementation Profile (FIP) workshop (ENVRI FAIR and GO-FAIR)
   Learning outcome: You can introduce the FIP approach, discussing its ontology and the FAIR enabling resources (technical solutions, implementation strategies, standards and models) which address each of the FAIR Guiding principles.
3. Provenance training

Learning outcome: You will be able to present the theoretical background of provenance, its rationale and values, how it may be recorded, and how it interrelates with FAIRness.

4. I-Adopt workshop (InteroperAble Descriptions of Observable Property Terminologies) framework

Learning outcome: You will be able to describe the RDA I-ADOPT frameworkand discuss how to use it as a way to compile clear and unambiguous definitions of variables in a standardised way, advancing FAIRness in many ways

Learning Path #3. FAIRify your infrastructure

- ***Policies for FAIR web services***

Target group: managers and staff in supporting positions in RIs aiming to implement FAIR principles in the work and presentation of the infrastructure's data and services. The material is based on the work of WP4 in ENVRI-FAIR on common FAIR policies.
Learning outcome: The main theme of this path is the formulation of the policies needed for making data and services FAIR. You will receive an overview of drivers and contexts for policies and tips for how to write and curate them. Once the policies are defined, step 4 will give you examples of different web services that can be developed for your infrastructure.
Components in suggested reading order:

1. "1st ENVRI-FAIR policy workshop for the ENVRI community Research Infrastructures": motivates why a policy framework is required for a research infrastructure

   Learning outcome: you will be able to motivate the need for the different policies, and connect formal policy definitions with their practical implementation using examples from EPOS.

2. "ENVRI-FAIR 2nd policy workshop for the ENVRI community Research Infrastructures": discusses around access, admin, metadata and data related policies

   Learning outcome: You will examine the ENVRI Policy Framework, presenting some of the key policy drivers relevant to ENVRI data service providers and discussing the necessary requirements and best practises for ENVRI RIs policies.

3. "3rd ENVRI-FAIR Policy Workshop": from policy, to guidelines to implementation

   Learning outcome: You can discuss the main aspects of creating policies (also examining the three key documents related to policy) and guidelines from policy statements and their consequent IT implementation.

4. "4th ENVRI-FAIR Policy Workshop": from policy, to guidelines to implementation

   Learning outcome: You will examine the progress made on policies in ENVRI and define a roadmap, an integrated plan for policy, and the policy target for ENVRI RIs and ENVRI-Hub.

5. "ENVRI Community International Summer School 2022 (developing FAIR web services)": here different examples are given of web services for FAIR data

   Learning outcome: you will examine approaches and solutions to designing and developing user-centred interfaces for services, validating and evaluating them and fostering reusability/interoperability among them.

10 Appendix 4: The ENVRI Learning Object Metadata Model

Based on content from https://trainingcatalogue.envri.eu/about, as accessed on 2023-06-23.

The ENVRI FAIR training catalogue's accurate and descriptive metadata allow all users to find the most appropriate and well-suited educational resources for their needs. This catalogue is based on the IEEE Standard for Learning Object Metadata that has been customised in order to be compliant with the EOSC Training Resource Profile - Data Model. It currently consists of 30 elements, most of which are mandatory. In the

The metadata elements used in the ENVRI Training Catalogue at the conclusion of the ENVRI-FAIR project (June 2023) are listed below, together with their definitions, acceptable values (if there are controlled vocabularies) and cardinality limitations (mandatory or optional & single/multi-value). In addition, major changes with respect to the original model (see deliverable D6.2 [Vaira 2021]) are indicated in **bold red** text.

1. **General**: this category groups the general information that describes the learning object as a whole.
   1.1 **Identifier:** *(mandatory, single value)* a globally unique label that identifies the learning object.
   1.2 **URL type:** *(mandatory, single value)* the designation of identifier scheme used for the URL. It represents the type of the URL of the learning object, that is the used scheme. It can be:
   - ARK (Archival Resource Key)
   - DOI (Digital Object Identifier)
   - HANDLE
   - URN (Uniform Resource Name)
   - URI (Uniform Resource Identifier)
   - URL (Uniform Resource Locator)
   - NOT APPLICABLE

   1.3 **URL:** *(mandatory, single value)* the URL that resolves to the learning object or to a "landing page" that contains important contextual information including the direct resolvable link to the learning object, if applicable.
   1.4 **Title:** *(mandatory, single value)* the name given to the learning object.
   1.5 **Language:** *(mandatory, multi-value)* the primary human language or languages used within the learning object.
   1.6 **Description:** *(mandatory, single value)* the textual description of the content of the learning object.
   1.7 **Keywords:** *(mandatory, multi-value)* the keyword(s) describing the topic of the learning object.
   1.8 **Geographical availability:** ***(mandatory, single value)*** **the locations where the learning object is offered. It can be:**
   - **WW: Worldwide, all countries**
   - **EO: Europe, all European Countries**
   - **EU: European Union, all countries of the European Union**
   - **EZ: Euro Zone, all countries of the Eurozone**
   - **AH: Schengen Area, all Schengen Area countries**

2. **Lifecycle:** this category describes the history and current state of this learning object and the entities that have affected it during its evolution.
   2.1 **Version:** *(mandatory, single value)* the edition of the learning object. Example: 1.2. Specify "Not available" if needed.
   2.2 **Status:** *(mandatory, single value)* the completion status or condition of the learning object. It can be
   - Draft
   - Final,
   - Revised
   - Unavailable

   2.3 **Contribute:** those Entities (i.e., people, organisations) that have contributed to the state of the learning object during its life cycle (e.g., creation, edits, publication).
   2.3.1 **Role:** *(mandatory, single value)* the kind of contribution. It can be

- Author
- Content provider
- Editor
- Educational validator;
- Graphical designer
- Initiator
- Instructional designer
- Publisher
- Script writer
- Subject matter expert
- Technical implementer
- Technical validator
- Terminator
- Other

2.3.2 **Entity:** *(mandatory, single value)* the identification of and information about entities (i.e., people, organisations) contributing to this learning object.

2.4 **Date:** *(mandatory, single value)* the date of the contribution. E.g., 2001-08-23. Specify "Not available" if needed.

3. **Educational:** this category describes the key educational or pedagogic characteristics of the learning object.

3.1 **Interactivity type:** *(mandatory, single value)* the predominant mode of learning supported by the learning object. It can be

- Active
- Expositive
- Mixed

3.2 **Learning resource type:** *(mandatory, single value)* the specific kind of the learning object. It can be

- Diagram
- Exam
- Exercise
- Experiment
- FAQ
- Index
- Lecture
- Narrative text
- Problem statement
- Questionnaire
- **School**
- Self-assessment
- Simulation
- Slide
- Table
- Video
- Webinar
- Other

3.3 **Interactivity level:** *(mandatory, single value)* the degree of interactivity characterising the learning object. It refers to the degree to which the learner can influence the aspect or behaviour of the learning object. It can be:

- Very low
- Low
- Medium
- High
- Very high

3.4 **Semantic density:** *(mandatory, single value)* the degree of conciseness of the learning object. The semantic density of a learning object may be estimated in terms of its size, span, or - in the case of self-timed resources such as audio or video - duration. It can be:

- Very low
- Low
- Medium

- High
- Very high

3.5 **Target group:*(mandatory, multi-value)* the principal user type for which the learning object was designed. It can be:**

- **Researchers**
- **Research groups**
- **Research communities**
- **Research projects**
- **Research networks**
- **Research managers**
- **Research organisations**
- **Students**
- **Innovators**
- **Businesses**
- **Providers**
- **Funders**
- **Policy Makers**
- **Research Infrastructure**
- **Managers**
- **Provider Managers**
- **Resource Managers**
- **Publishers**
- **Other**

3.6 **Context:***(mandatory, single value)* the principal environment within which the learning and use of the learning object is intended to take place. It can be:

- Higher education
- School
- Training
- Other.

3.7 **Expertise level:*(mandatory, single value)* the target skill level in the learning object for the typical intended target audience. It can be:**

- **Advanced**
- **Intermediate**
- **Beginner**
- **All.**

3.8 **Typical learning time:***(mandatory, single value)* the approximate or typical time it takes to work with or through the learning object for the typical intended target audience. Example: PT1H30M, which means 1 hour and 30 minutes. Specify "Knowledge-dependent" if the learning time depends on the familiarity with the context.

3.9 **Learning outcome(s): *(mandatory, multi-value)*the descriptions of what knowledge, skills or abilities the target group should acquire on completion of the learning object. The Bloom's Taxonomy can be used to write effective learning outcomes.**

3.10 **Access rights:***(mandatory, single value)* the access status for the learning object. It can be:

- Open access: it refers to a learning object that is immediately and permanently online, and free for all on the Web, without financial and technical barriers.
- Restricted access: it refers to a learning object that is available in a system but with some type of restriction for full open access.
- Metadata only access: it refers to a learning object in which access is limited to metadata only. The resource itself is described by the metadata, but neither is directly available through the system or platform nor can be referenced to an open access copy in an external source.
- Paid access: it refers to the need to pay a fee to access the learning object.

3.11 **Cost:** *(mandatory, single value)* whether use of the learning object requires payment. It can be:

- Yes
- No

3.12 **Copyright and other restrictions:** *(mandatory, single value)* whether copyright or other restrictions apply to the use of the learning object. It can be:

- Yes
- No

3.13 **Condition of use:** *(mandatory, single value)* comments on the conditions of use of the learning object.

4. **Technical:** this category describes the technical requirements and characteristics of the learning object.

4.1 **Size:** *(mandatory, single value)* the actual size of the digital learning object in bytes not Mbytes, GB, etc. If the learning object is compressed, then this data element shall refer to the uncompressed size. Specify "Not available" if needed.

4.2 Scientific domain and subdomain: *(mandatory, single value)* the branch and sub-branch of science, scientific discipline that is related to the learning object. It can be:

i) Natural Sciences:
- **Mathematics**
- **Computer and information sciences**
- **Physical sciences**
- **Chemical sciences**
- **Earth and related environmental sciences**
- **Biological sciences**
- **Other natural sciences**

ii) Engineering & Technology:
- **Civil engineering**
- **Electrical, electronic and information engineering**
- **Mechanical engineering**
- **Chemical engineering**
- **Materials engineering**
- **Medical engineering**
- **Environmental engineering**
- **Environmental biotechnology**
- **Industrial biotechnology**
- **Nanotechnology**
- **Other engineering and technology sciences**

iii) Medical & Health Sciences:
- **Basic medicine**
- **Clinical medicine**
- **Health sciences**
- **Medical biotechnology**
- **Other medical sciences**

iv) Agricultural Sciences:
- **Agriculture, forestry, and fisheries**
- **Animal and dairy sciences**
- **Veterinary sciences**
- **Agricultural biotechnology**
- **Other agricultural sciences**

v) Social Sciences:
- **Psychology**
- **Economics and business**
- **Educational sciences**
- **Sociology**
- **Law**
- **Political sciences**
- **Social and economic geography**
- **Media and communications**
- **Other social sciences**

vi) Humanities:
- **History and archaeology**
- **Languages and literature**
- **Philosophy, ethics and religion**
- **Arts**
- **Other humanities**

vii) Generic: Generic

viii) Other: Other

4.3 **Location <removed>**

4.4 **Topic codes:** *(mandatory,* ***multi-value****)* the code and title of the covered topics according to the list of the training topics identified in Table 1 of the ENVRI-FAIR project Deliverable 6.1. Topic codes are divided into two main categories: "general FAIR-related" (from G1 to G7) and "research data management-related" (from R1 to R17).

5. Relation: this category defines the relationship between this learning object and other learning objects, if any.

5.1 Kind: ***(optional, single value)*** **the nature of the relationship between this learning object and the target learning object. It can be isPartOf, hasPart, requires, isRequiredBy.**

5.2 Entry: ***(optional, multi-value)*** **the value of the identifier within the identification or cataloguing scheme that designates or identifies the target learning object. A namespace specific string, e.g., http://www.ieee.org/documents/1234.**

11 Appendix 5: Training event participant feedback form

Here we reproduce an example of the questionnaire used to collect feedback from participants of training webinars and workshops. We have interleaved the questions with the responses we received.

—----------

Service validation & evaluation webinar - feedback form

On Friday, 17 June 2022, ENVRI-FAIR organised a webinar "Service validation & evaluation: making sure your services are up to the task". This event was associated with the 2022 ENVRI Community International Summer School, and featured presentations by Mark van de Sanden from EUDAT (Validating services & assessing their Technical Readiness Level), Yin Chen from EGI (Service evaluation: why & how) and Maggie Hellström from ICOS (Evaluating ENVRI services: experiences from the ENVRIplus).

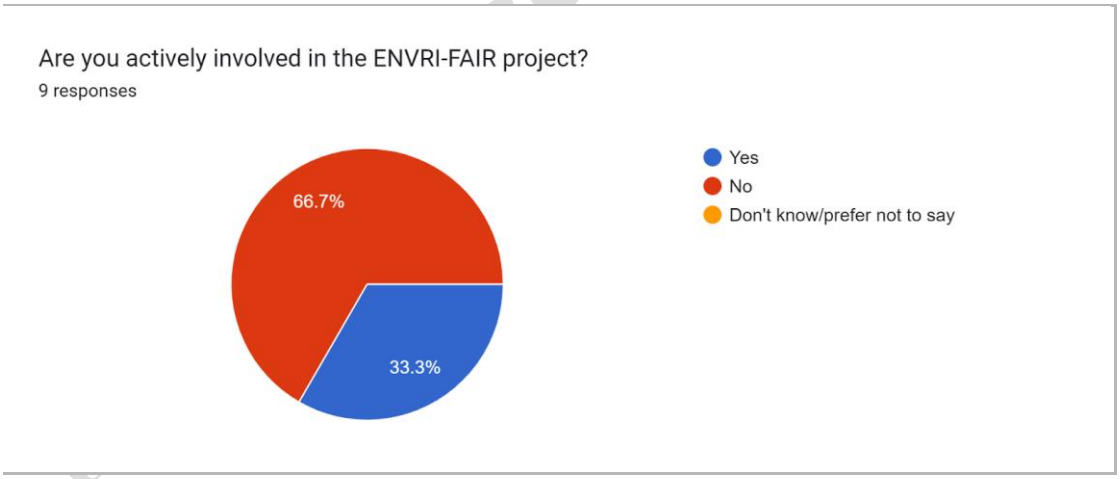
If you took part in this webinar, we invite you to provide us with some quick feedback - it should only take a few minutes to complete this form, and we really appreciate you sharing your views!

Thanks in advance!
Maggie Hellström, Nicola Fiore and Jacco Konijn - co-chairs of the ENVRI-FAIR training work package

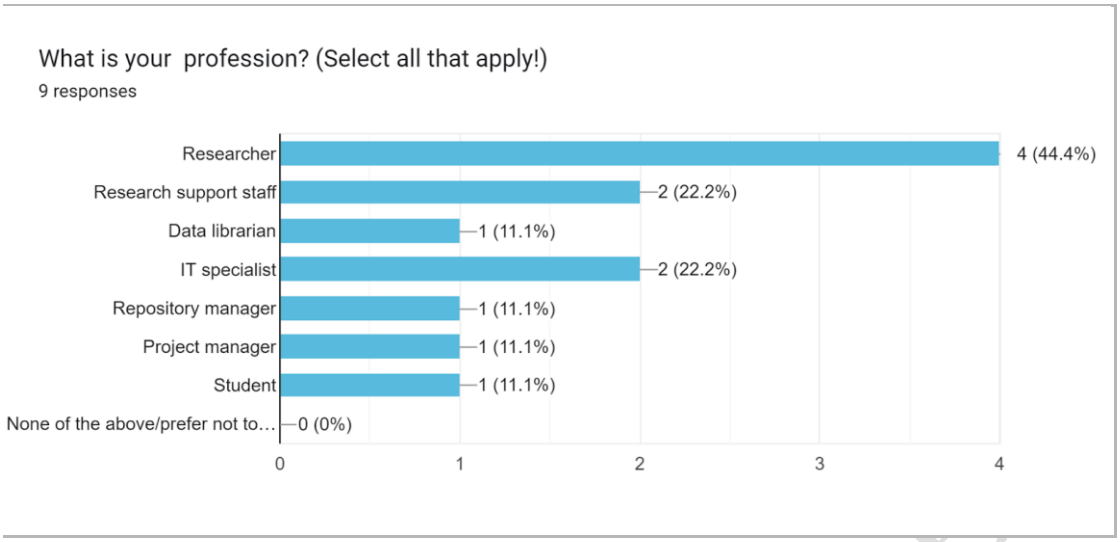
* Indicates required question

==========

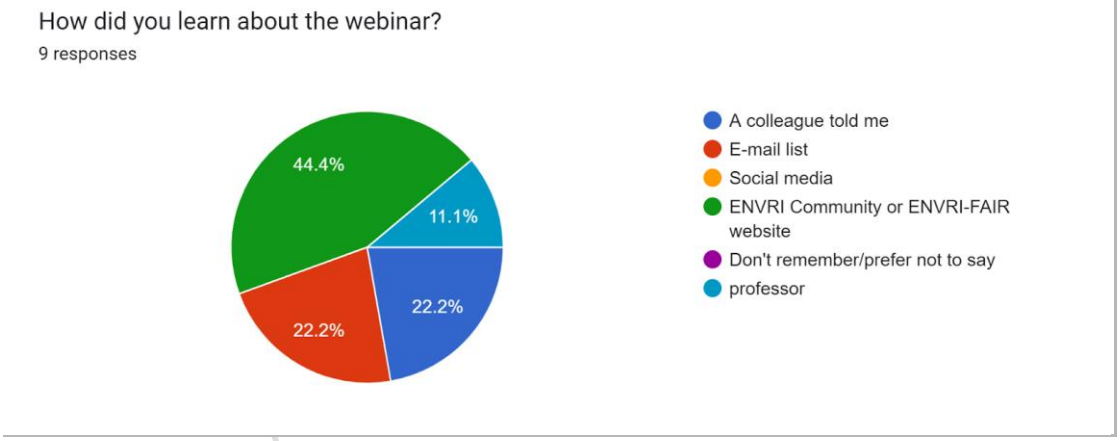
Are you actively involved in the ENVRI-FAIR project? *
O Yes
O No
O Don't know/prefer not to say



What is your profession? (Select all that apply!) *
O Researcher
O Research support staff
O Data librarian
O IT specialist
O Repository manager
O Project manager
O Student
O None of the above/prefer not to say
O Other: _____

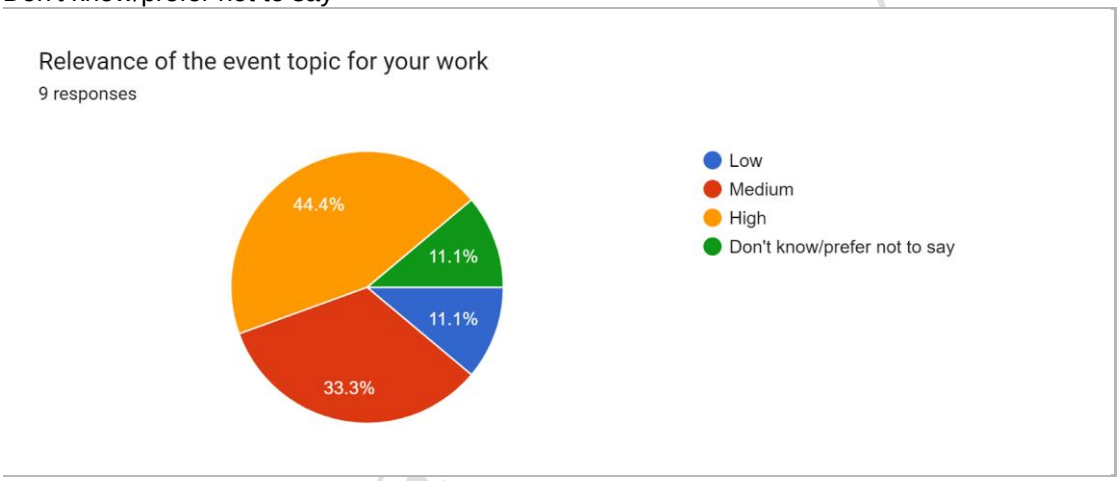


How did you learn about the webinar? *
O A colleague told me
O E-mail list
O Social media
O ENVRI Community or ENVRI-FAIR website
O Don't remember/prefer not to say
O Other: _____

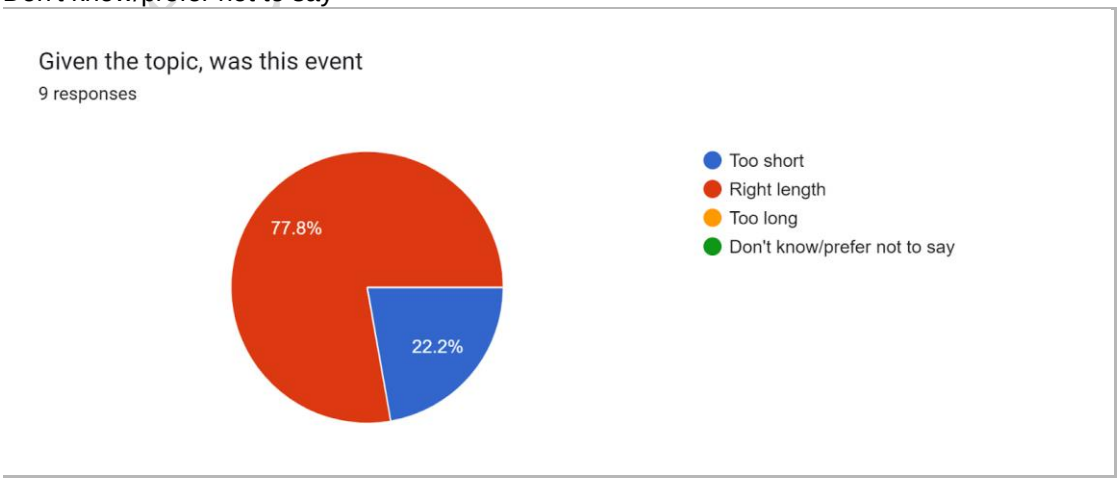


Please rate the following aspects *
Use the scale 0: don't know / prefer not to say, 1: poor, 2: reasonable, 3: good, 4: very good, 5: excellent
Overall satisfaction with the event: ___
The presentations were: ___
Quality of discussions and dialogue: ___
Usefulness of shared materials (video, slides): ___
Virtual meeting facility (Zoom): ___

Relevance of the event topic for your work *
Low
Medium
High
Don't know/prefer not to say

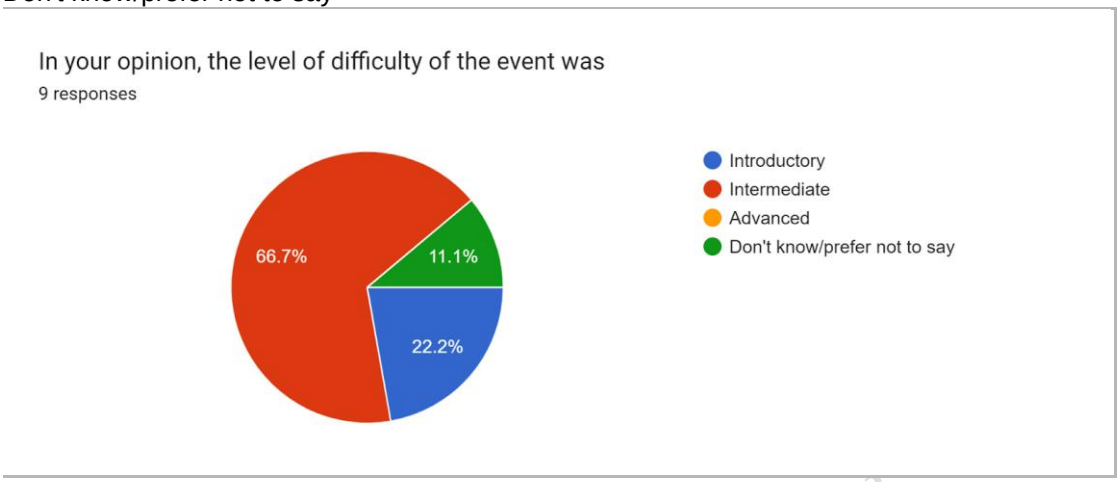


Given the topic, was this event *
Too short
Right length
Too long
Don't know/prefer not to say

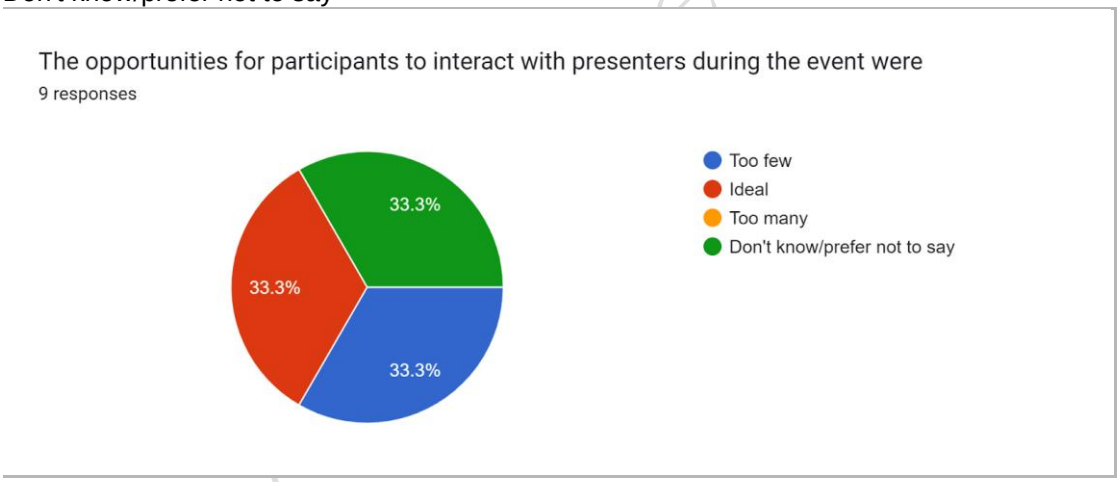


In your opinion, the level of difficulty of the event was *

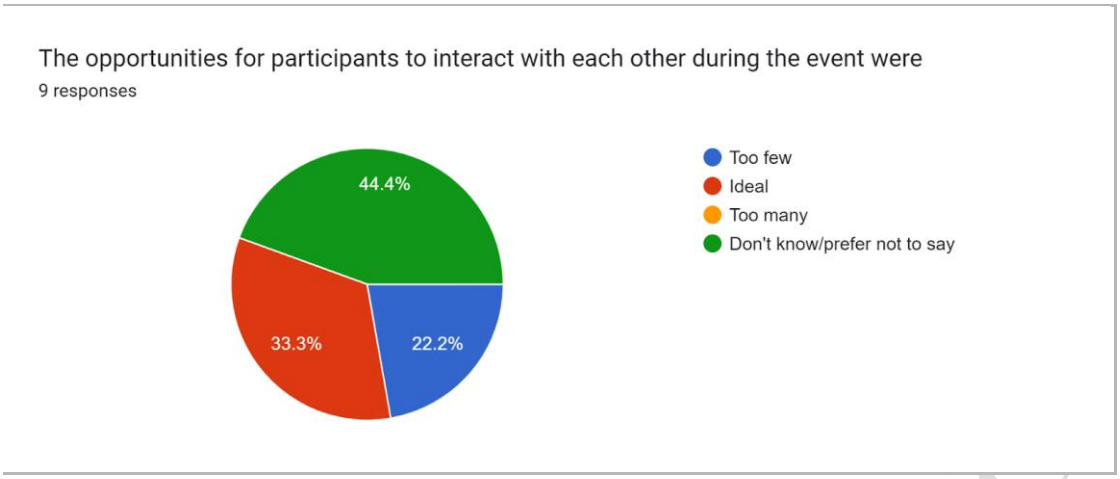
Introductory
Intermediate
Advanced
Don't know/prefer not to say



The opportunities for participants to interact with presenters during the event were *
Too few
Ideal
Too many
Don't know/prefer not to say



The opportunities for participants to interact with each other during the event were *
Too few
Ideal
Too many
Don't know/prefer not to say



What is your top take-away from the event?

- Discuss with project leads the relevance and effort needed for providing our services to marketplace.eosc-portal.eu for us
- getting to know EOSC initiative
- importance of the process of validation and evaluation of a service and introductory information on how to do it.
- to plan both validation life cycle and evaluation well in advance when designing a new service
- Learning the difference between service validation and evaluation.
- Validation and evaluation are vital parts of an infrastructure to provide better services

What did you most appreciate or enjoy about the event?

- 1) It was unclear how good a match for me this was, so I am happy to have this chance to join and find out 2) It let me map out part of the science service ecosystems
- the topic per se... that for me was absolutely new.
- appreciate sharing not only knowledge about the topic, but also experiences from former projects
- Collaborative notes by presenters and participants. Excellent coordination overall.
- Complete information in relative short time

Please let us know how we could improve!

- I think this was more relevant for project leads, managers, ... than for the IT people in the trenches.
- The webinar was fine. One thing that maybe can be improved is the need of some examples about services evaluation-validation issues etc at a more practical level in order to fix better and understand deeply all the concepts and topics explained during the webinar.
- a little more time for Q&A session
- We could have more time for Q&A next time.

Thank you! We really appreciate your feedback, which will help ENVRI-FAIR WP6 to continue providing high quality training on topics relevant to the ENVRI Community and beyond!