



EOSC-hub

D11.5 - Updated training materials about common services, thematic and competence

Lead Partner:	EGI Foundation
Version:	1
Status:	FINAL
Dissemination Level:	Public
Document Link:	https://documents.egi.eu/document/3637

Deliverable Abstract

This document is the final training activity report for WP11. It includes the final status of the online catalogue along with the list of all the training events, materials, datasets, and deployable training environments developed during the project lifetime to target the training needs of data scientists.

COPYRIGHT NOTICE



This work by Parties of the EOSC-hub Consortium is licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). The EOSC-hub project is co-funded by the European Union Horizon 2020 programme under grant number 777536.

DELIVERY SLIP

<i>Date</i>	<i>Name</i>	<i>Partner/Activity</i>	<i>Date</i>
From:	Giuseppe La Rocca	EGI Foundation/WP11	01/04/2021
Moderated by:	Malgorzata Krakowian	EGI Foundation/WP1	
Reviewed by:	Elly Dijk Anabela Oliveira	DANS-KNOW LNEC	19/03/2021
Approved by:	AMB		

DOCUMENT LOG

<i>Issue</i>	<i>Date</i>	<i>Comment</i>	<i>Author</i>
v.0.1	31/08/2020	Initial ToC with content	Giuseppe La Rocca/EGI Foundation
v.0.2	22/02/2021	Full draft of WP11	Giuseppe La Rocca/EGI Foundation
v.0.3	05/03/2021	Integrate inputs and comments from WP11 partners	Alexandre Bonvin/UU, Fabrizio Antonio/CMCC, Dieter Van Uytvanck/CLARIN, Mattia Santoro/CNR, Narges Zarrabi/SURF, Ingemar Häggström/EISCAT_3D, Renè van Horik/DANS-KNAW
v.0.4	19/03/2021	Integrate feedback from reviewers	Anabela Oliveira/LNEC Elly Dijk/DANS-KNAW
v.1	01/04/2021	Final	Giuseppe La Rocca/EGI Foundation

TERMINOLOGY

<https://wiki.eosc-hub.eu/display/EOSC/EOSC-hub+Glossary>

Contents

1	5
2	6
2.1	6
2.2	6
2.3	7
2.4	8
2.5	8
3	9
3.1	9
3.2	12
3.3	13
3.4	17
3.5	21
4	23
4.1	23
4.2	24
5	25
5.1	25
5.2	26
6	28
7	30

Executive summary

In order to leverage the potential of EOSC for open and data-intensive research, a key challenge for Europe is to ensure the availability of skilled people with an excellent knowledge of standards and best practices for delivering, using, sharing and analysing open and FAIR data as well as applications and tools. To realize this vision and reduce the risk of EOSC becoming inaccessible to the majority of researchers within academic institutions, dedicated resources have been allocated to support the development of skills and training to promote the uptake of EOSC and its services.

This document is the final report about the activities carried out by WP11 to sustain the training needs of the different stakeholders involved in the EOSC-hub project. As already reported in past deliverables, WP11 training activities were planned, designed, delivered and monitored with the final goal to raise awareness about the resources and services developed by the project. Thanks to the knowledge sharing promoted with WP11 training activities, and the easy access to advanced digital services, and scientific instruments data, researchers from all disciplines have the possibility to achieve excellence in science, research and innovation. By the time of writing of this report, the promotion of EOSC services, the technical support and the engagement activities carried out during the project's lifetime contributed to integrate new community-specific (from the EOSC-hub Competence Centres) and domain-specific services (from the EOSC-hub Thematic Services) in EOSC and helped SMEs and industries to support the development of innovative applications through the EOSC Digital Innovation Hub. With the EOSC-hub project coming to an end, in collaboration with other trainers of EOSC-related projects working on training areas, WP11 contributed to promote a more focused training strategy to support EOSC in all its phases and ensure its uptake. In this regard:

- EOSC-hub training results were promoted and discussed within the EOSC-wide Training Coordinators' Community of Practice group - a network of trainers composed by representatives of e- and Research Infrastructures and national training initiatives coordinated by OpenAIRE, including CESSDA, DARIAH, ELIXIR, EOSC-hub, FREYA, and many others¹.
- EOSC-hub has also supported training activities organized by other organizations such as: CODATA-RDA, CORBEL, ENVRI-FAIR, INSTRUCT, NGS, NI4OS, PaNOSC and VIRGO.
- A Collaboration Agreement (CA) with the FAIRsFAIR project was agreed on topics of common interest.

EOSC-hub, in collaboration with OpenAIRE and other e- and Research Infrastructures and national training initiatives, contributed to the organization of a training workshop in EOSC. During the workshop all the EOSC-related projects were invited to discuss training strategies for the next years in EOSC. The main outcome from the workshop was a set of recommendations that fed into the EOSC Training and Skills Working Group established by the EC.²

¹ More information on the CoP on training can be found at: <https://www.openaire.eu/cop-training>

² The report "Digital Skills for FAIR and Open Science" can be found at: <https://op.europa.eu/s/oMO0>

Introduction

This document summarizes the final training activities carried out by the project during the period M19-M39. These activities aimed to address the training needs of scientist communities and individual researchers across various countries and scientific disciplines working together in solving today's biggest societal issues, and service providers interested to contribute to the establishment of the EOSC virtual environment. Training activities, and others related to training, reported in this document were organized in collaboration with all the funded partners directly involved in WP7 and the un-funded partners working in WP8.

The document is organized as follows:

- Section 2 reports the activities carried out during the second part of the project to adjust the training strategy and make it more focused.
- Section 3 provides a final report about the training activities carried out during the last part of the project. Section 4 outlines the main outcomes of the webinar programmes organized to continue the promotion of the EOSC-related services and promote the uptake of community-specific/proof-of-concept services.
- Section 5 shows the final architecture of the revamped catalogue of training materials with a quick overview about the training materials, and deployable environments developed by the project. Some statistics (e.g. page views) of the catalogue are reported in Section 6.
- Finally, Section 7 draws some conclusions and provides an outlook to future training activities that will be carried out in the context of the EOSC Future project.

1 Training strategy in EOSC

EOSC is a trusted digital platform which could potentially serve 1.7 million researchers in Europe and progressively expand its user base to include the wider public sector and the private sector. The main objectives of this digital platform are to provide seamless access to a plethora of data and interoperable services that can be used for helping researchers to support the whole research data cycle (e.g. from discovery and mining to storage, management, analysis and reuse). To make EOSC a success, this digital platform and its services must be widely adopted by the European researchers. To make it happen, one of the strategic objectives of EOSC is to focus on Skills and Training activities to reduce, from one hand, the barriers of adoption and, on the other, promote the implementation of a focused training strategy to support EOSC in all its phases and ensure its uptake. In this section we briefly report the initiatives where EOSC-hub contributed to the implementation of this training strategy in EOSC.

1.1 The Community of Practice group

The Community of Practice (CoP) group is an informal network of training coordinators and managers composed by over 60 representatives of e- and Research Infrastructures and national training initiatives, including OpenAIRE, CESSDA, DARIAH, ELIXIR, EOSC-hub, FREYA, GEANT, PRACE, RDA and many others. The main objective of the group, launched in September 2018 and coordinated by the OpenAIRE project, was to stimulate the sharing of training experiences and best practices about training related topics organized in EOSC. The EOSC-hub project contributed to this group: 1.) sharing best practices for implementing the Open Science principles with open-source solutions and open services (See section 3.3.2 for more details), and 2.) providing, in collaboration with the OpenAIRE-Advance project, consultancy and expertise for the implementation of Data Management Plans (DMP). An important milestone for the group was the organisation of the Training Workshop in EOSC organized in Feb. 2020 (see next section).

1.2 Workshop on Training in EOSC

From Wednesday 26 till Friday 28 February 2020, a workshop³ took place in The Hague, the Netherlands, around the theme of training in the European Open Science Cloud (EOSC). This workshop was funded through a co-creation grant of the EOSC Secretariat⁴, and organised by DANS in cooperation with EUDAT, EGI, representing EOSC-hub, and OpenAIRE. Training coordinators, within the EOSC projects and associated organisations, were invited to collaboratively work on the Rules of Participation for training in EOSC and develop practical recommendations for training service providers. During the three days of workshop the 35 training coordinators, representing more than thirty EOSC-related projects and initiatives, provided, as a result, recommendations in the following two main areas:

³ <https://www.eosc-hub.eu/events/workshop-training-eosc>

⁴ <https://www.eoscsecretariat.eu/>

-
- Rules of Participation for training in EOSC, and
 - Practical Guidance for training service providers.

The main result regarding the Rules of Participation for training in EOSC was the adaptation of the EOSC Rules of Participation (v0.2)⁵ to include training services and training materials. Moreover, additional recommendations and guidance for the on-boarding of training service providers were also gathered and included in the final report. For more details about the workshop, please refer to the public report.⁶

1.3 The EOSC Training and Skills Working Group

The project also participated in the EOSC Skills and Training Working Group (WG) formed in 2020 to identify a framework for building competence and capabilities in EOSC. This WG focused on four priority areas:

- Developing the next generation of FAIR and open science professionals.
- Collaborating to enhance digital skills for FAIR and open science in Europe.
- Building a trusted and long-lasting federated catalogue of training materials and related tools.
- Influencing national open science policy for skills by supporting strategic leaders.

In this WG EOSC-hub contributed to identify the obstacles that are limiting the exploitation of research impact internationally and define proper countermeasures that can be adopted to reduce these barriers. More specifically, the gap analysis has highlighted the lack of experts that can facilitate the promotion of FAIR and Open Science principles. Key profiles such as: Data scientists, data stewards, data curators and research software engineers are needed for the development of data-driven, data intensive science. Another barrier identified by the WG is the fragmentation of training resources. There is a wide plethora of contents developed by different scientific communities, but it is missing the good practise to reuse existing data and keep them up-to-date. The final report from the EOSC Executive Board Skills and Training Working Group along with the key recommendations coming out from the EOSC Training and Skills Working Group activities are published online⁷.

⁵ <https://repository.eoscsecretariat.eu/index.php/s/QWd7tZ7xSWJsesn#pdfviewer>

⁶ <https://zenodo.org/record/3894370>

⁷ <https://doi.org/10.2777/59065>

1.4 Collaboration Partnership with the FAIRsFAIR project

In July 2020, EOSC-hub and FAIRsFAIR⁸ signed a Memorandum of Understanding (MoU)⁹ to strengthen their collaboration and contribution to the European Open Science Cloud. With this agreement the EOSC-hub project offers a discovery and access channel to FAIR-accredited datasets via the EOSC-hub Marketplace. With the MoU the two projects agreed to work together on the following three main areas:

- **Sharing of templates and policies for FAIR data repositories.** The FAIRsFAIR project formulated ten recommendations on practice to support FAIR data principles.¹⁰ A consultation was carried out to assess the relevance of these recommendations for the services developed by the EOSC-hub project. The outcomes of the consultation are published in a working paper¹¹ that on the one hand describes to what extent the EOSC-hub services can be used to implement the recommendations formulated by the FAIRsFAIR project and on the other hand informs the FAIRsFAIR project on the applicability of the recommendations. Moreover, the working paper can help to carry out future activities to support the FAIR data principles.
- **FAIR Certification of trustworthy data repositories.**
- **Training activities.** In this joint activity, the EOSC-hub and FAIRsFAIR projects promoted the uptake of the Open Science and FAIR data principles in their training programmes. Both projects coordinated their planned events to promote project results and reach a wider user base. For this joint activity, FAIRsFAIR leveraged on the FAIRsFAIR Competence Centres to share knowledge and best practices for helping scientific communities requiring support on adopting FAIR data principles, while EOSC-hub relied on the distributed network of expertise to address the training needs of the different research communities.

1.5 Training contents in the EOSC Portal

All the documentations and training materials developed by the EOSC-hub project to target the technical requirements of research communities¹² and service providers¹³ are linked in the EOSC portal. All the documentation is licensed under CC-BY in order to ensure that future initiatives have sufficient IPR rights to curate and further improve the available material.

⁸ <https://www.fairsfair.eu/>

⁹ https://confluence.egi.eu/download/attachments/23233414/EOSC-hub_FAIRsFAIR_MoU.pdf

¹⁰ D3.4 FAIRsFAIR Recommendations on practice to support FAIR data principles.
<https://zenodo.org/record/392413>

¹¹ EOSC-hub and the FAIRsFAIR recommendations on practice to support the FAIR data principles (Working paper) <http://zenodo.org/record/4638551>

¹² <https://eosc-portal.eu/using-the-portal/eosc-hub-scientists-and-scientific-communities-training>

¹³ <https://eosc-portal.eu/using-the-portal/eosc-hub-service-provider-training>

2 Training activities in EOSC-hub

The project has identified a total of 9 Key Exploitable Results (KERs), i.e. outputs which can be taken up, exploited and reused to support the future mature EOSC. One of these KERs is related to the training activities (KER no.9). Overall, structured training programmes represent a prime opportunity to expand the knowledge base of the participants and to promote the uptake of the EOSC services. In the academy domain, a user who receives the necessary training is better able to run his/her research experiment. Also in the EOSC landscape, training activities play a key role in creating awareness about the services and resources offered by EOSC to fit the technical requirements of the diverse stakeholders such as:

- service providers who might benefit from technical assistance on using, integrating and providing services in EOSC, or
- individual researchers possibly encountering the e-Infrastructures for the first time, who are interested in enabling a smooth integration into the EOSC ecosystem and maximising the benefits.

To contribute to creating awareness about the EOSC services and resources, augmenting skills and adapting organisational practices for a full participation in the EOSC ecosystem, the training programme coordinated by WP11 has included the following topics:

- **Training on Data Management Planning (DMP).**
- **IT Federated Service Management training** events to learn the fundamentals of IT Service Management processes.
- Training contents about **Common and Federated services** for supporting scientific activities of Thematic Services (TSs), Competence Centres (CCs) and research communities.
- **IT Security Forensics training.**
- **Domain-specific training** for data providers and data scientists.

An updated status of the training activities and materials produced during the last part of the project is presented in the next sections.

2.1 Training, support and consultancy on Data Management Planning (DMP)

To encourage the creation of a Research Data Management plan to make the research process as efficient as possible and meet expectations, in T11.2 DANS-KNAW (in collaboration with OpenAIRE-Advance project) has continued to offer technical assistance and consultancy, and offered an overview about the main implementation tools that can be used to manage and share research data during the different phases, from its entry to the research cycle through to the dissemination and archiving of valuable results. Broadly speaking, during the last part of the EOSC-hub project, T11.2 focused primarily on recommendations to support FAIR data, support for the Domain Data Protocol (DDP), and providing support for research data management.

2.1.1 Recommendations to support FAIR data

The formulation of the recommendations for data and infrastructure providers is based on an extensive consultation process in which EOSC-hub cooperated with other H2020 projects including FREYA, FAIRsFAIR, RDA Europe and OpenAIRE. Three workshops were organised in 2019. The initial two workshops examined services in the research data ecosystem and discussed challenges and recommendations for services to support FAIR data through panels and breakout groups. During these two workshops, key needs and areas for improvement were identified by participants. The most notable areas for improvement were a lack of a sustainable ecosystem of independent interoperable services; findability and accessibility, which requires mostly technical expertise and specific domain expertise for increasing interoperability and reuse; and skills and services for data stewardship and preservation. After the community consultations yielded a list of recommendations, the third workshop set out to prioritize those recommendations, define actions to be taken, and suggest stakeholders best suited for taking responsibility to carry these actions forward. While the prioritization exercise showed substantial heterogeneity between different stakeholder groups, essential infrastructure components and socially oriented recommendations around fostering global collaborations and including FAIR in research assessments scored high among all groups. The recommendations and priorities have been published and made available for a wide audience.¹⁴

2.1.2 Domain Data Protocol (DDP)

A Domain Data Protocol (DDP) is a Data Management Plan that is tailored to be used in a specific field or discipline. Users of the DDP should only comply with this DDP or explain why they deviate from it. The DDP principle is introduced by Science Europe¹⁵, the association of major research funding organizations. A main activity of T11.2 in relation to this topic was a contribution to the EOSC-hub week in May 2020 in which the background and benefits of a DDP were shared with the EOSC-hub community. This raised interest on the topic by stakeholders and resulted in activities ranging from providing additional information by email to organising dedicated workshops on the topic (see training register). The DDP principle is introduced in a couple of research infrastructures and stakeholders are motivated to create and apply DDPs for their community. T11.2 developed a workshop template that had the following goals:

- To provide a short introduction on the state of art of Data Management Plans (DMP) and Domain Data Protocols (DDP).
- To recognize the similarities and differences between a DMP and a DDP.
- To gain experience with creating a DMP by filling in a small part of a DMP (hands-on).
- To gain experience with compiling a DDP for a specific scientific field (hands-on).
- To assess the potential value of a DDP for “your” research environment.

¹⁴ “Recommendations for services in a FAIR data ecosystem” <https://doi.org/10.1016/j.patter.2020.100058>

¹⁵ <https://www.scienceeurope.org/>

The hands-on assignment consists of an introduction and break-out groups. The aim of the assignment is to fill in a small part of a DMP template and to evaluate whether the elements of the DMP that are chosen can be used as a basis for a Domain Data Protocol for a specific discipline.¹⁶

2.1.3 Providing support on research data management

During the last part of the project T11.2 continued to provide input on research data management principles and tools (including attention for DDPs) to several stakeholders. Examples were webinars for research infrastructures such as DARIAH, the participation in webinars organised in cooperation with RDA, OpenAIRE national nodes and the FAIRsFAIR project. In these events often Open Science is the main topic that obviously has a close relationship with research data management services and tools. Depending on the background of the audience the support by T11.2 is constructed.

Overall, a total of 20 events (workshops, training, webinars) were organized during the last part of the project (M19-M39) in collaboration with the OpenAIRE-Advance and other projects initiatives. Among them, the most recent events organized were:

- An interactive session during the RDA's 16 (virtual) plenary meeting to introduce the use of ARGOS, the OpenAIRE DMP tool, for creating domain specific Data Management Plans in Archeological research.
- The workshop "Services to support FAIR data"¹⁷ jointly organized by EOSC-hub, OpenAIRE-Advance, FREYA, and RDA during the Open Science FAIR event that took place in Porto, November 2019. The main objective of this workshop was twofold: on one hand, to explore how existing infrastructures can work together and, on the other hand, to understand how to deliver services that support the creation of FAIR research outputs and suggest roadmaps towards the implementation of targeted aspects of Open Science principles. A final report with the main outcomes of these series of three workshops was published in 2020 in the following article¹⁸: Hylke Koers, Daniel Bangert, Emilie Hermans, René van Horik, Maaïke de Jong, Mustapha Mokrane, 'Recommendations for services in a FAIR data ecosystem' in: Patterns, Vol. 1, issue 6 (2020).
- The workshop "Simplifying RDM by using Domain Data Protocols"¹⁹, held in November 2019 in Aveiro, Portugal. During the workshop, the role of Data Management Plans (DMP), and more specifically Domain Data Protocols (DDP) were discussed. This topic is relevant for both researchers that provide the DMP requirement for a specific discipline, and policy makers/funders.

¹⁶ The workshop template can be found at: <https://doi.org/10.5281/zenodo.4588456>

¹⁷ <https://www.opensciencefair.eu/workshops-2019/services-to-support-fair-data-formulating-recommendations-for-eosc>

¹⁸ <https://doi.org/10.1016/j.patter.2020.100104>

¹⁹ <https://forumgdi.rcaap.pt/5forum/workshop-tail/>

- Workshop on Domain Data Protocols in January 2021 for NI4OS²⁰ based on the workshop template.
- A domain-specific Data Management Plan event in March 2021 for the ExPaNDS facilities. As part of this event an info-sheet was created to foster the awareness of the importance of data management planning for facility managers and end users.
- A webinar, in cooperation with OpenAIRE, on March 16, 2021 introducing FAIR research data management for 70 data supporters, librarians and researchers in Israel²¹.

The number of (cumulative) DMP training events planned and delivered by WP11 during the period (M19-M39) is monitored by the project specific metric M11.2.

Table 1 - WP11 metric for T11.2

ID	Definition	Baseline	Value M06	Value M12	Value M24	Value M39
M11.2	Number of DMP events organized	0	5	9	18	32

2.2 IT Federated Service Management Training

During the second part of the project T11.3 continued to support the implementation of a new Service Management System (SMS) for EOSC based on the lightweight FitSM standard. Starting from June 2019, T11.3 has extended its initial training offer including also Advanced level courses. In total, during the period (M19-M39), T11.3 organized additional 11 formal FitSM training courses of which 9 were Foundation and 2 were Advanced level. According to the DoA, the plan was also to run one Expert level course before the end of the project, but this depends on having enough critical mass with participants holding the Foundation and both Advanced level certificates. Overall, the total courses organized during the period (M19-M39) were fewer than planned due to COVID-19. However, T11.3 task members worked on adapting the courses to be run remotely. The first ever remote course was held in July 2020. Lessons learned were applied to both courses that were co-located with 2 conferences in Nov. 2020. Total participants training to date are 212 with 206 successfully receiving certification in FitSM (backed by Certification Authority ICO-Cert).

The number of (cumulative) FitSM training events planned and delivered by WP11 is monitored by the project specific metric M11.2.

²⁰ <https://ni4os.eu/>

²¹ <https://dans.knaw.nl/en/current/dans-trains-israeli-data-supporters-in-research-data-management>

Table 2 - WP11 metric for T11.3

ID	Definition	Baseline	Value M06	Value M12	Value M24	Value M39
M11.3	Number of FitSM training events delivered during the project	0	7	8	16	22

2.3 Training about Common and Federated services

During the second part of the project T11.4 continued with the organization of specialized training events to promote the uptake of the common and federated services offered by European e-infrastructures of EOSC-hub for computing, data management, AAI and other topics. These services are relevant for service providers who need to enhance their service before enabling access for new communities via the EOSC Portal. The list of events organized grouped per categories are reported below:

2.3.1 AAI services

A total of 3 events were organized by T11.4 during the period (M19-M39) to update the scientific communities about the status of the EOSC AAI service. The most important updates were provided during the remote EOSC-hub week²². During the event, a focused training session was organized to provide an overview of enabling the use and management of community identities for accessing resources made available through the EOSC-hub AAI. More specifically, the session put the focus on the Community AAI which streamlines researchers' access to services, both those provided by their own infrastructure (if they have one) as well as the services provided by infrastructures that are shared with other communities. The target audience of this training session were community managers interested in enabling the members of their communities to access services and resources in a secure and user-friendly way through the EOSC-hub AAI.

2.3.2 Computing services

In the past years, the vision of Open Science has emerged as a new paradigm for transparent, data-driven science capable of accelerating competitiveness and innovation. To promote the implementation of Open Science principles with open-source solutions and open services, EOSC

²² <https://www.eosc-hub.eu/eosc-hub-week-2020/>

promoted the integration of the following services: Jupyter²³, Binder²⁴, GitHub²⁵ and Zenodo²⁶ (see Fig. 1).

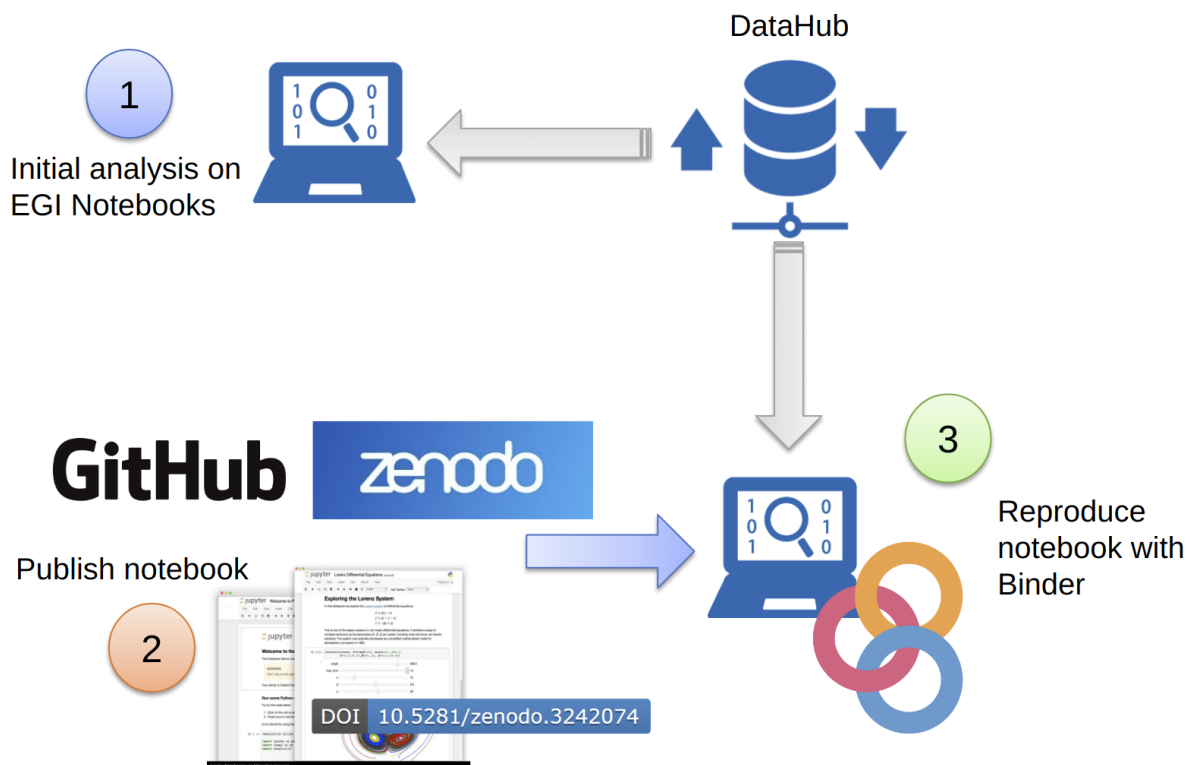


Fig. 1 - A practical example of service integration in support of Open Science

EGI Notebooks²⁷ is a service of the EGI e-infrastructure, providing a user-friendly and highly flexible Jupyter-based web environment for the development and execution of data analysis and visualisation ‘notebooks’. More specifically, notebooks can contain programming codes in various languages, HTML scripts, dynamic visualization, equations as well as images and explanatory text to provide guidance and context for the analysis. Through notebooks users can easily *share* concepts, ideas and working applications, capturing the full analytical methodology, connections to data and descriptive text to interpret those data. Combined with Binder technology, the Jupyter notebooks are allowed to be *reproducible* and reusable by anyone, anywhere. Lastly, in combination with

²³ <https://jupyter.org/>

²⁴ <https://jupyter.org/binder>

²⁵ <https://github.com/>

²⁶ <https://zenodo.org/>

²⁷ <https://www.egi.eu/services/notebooks/>

GitHub and Zenodo²⁸, the open access repository for research publications, scientific data and other research objects, users are supported to easily engage with the concept of Open Science.

During the second part of the project T11.4 continued to foster the adoption of open-source solutions to implement Open Science principles in science in several training events jointly organized in collaboration with CODATA-RDA and EaPConnect²⁹. Among them, the most recent events organized were:

- The CODATA-RDA Research Data Science Summer School (2019)³⁰.
- Introduction to Jupyter and Open Science Training (2019)³¹.
- Training sessions during the EGI Conference 2019³² on how system administrators can deploy a JupyterHub instance for their users on top of Kubernetes and explore some of the possible customisations that can improve the service towards your users like integration with authentication services or with external storage systems.
- Open Science with Jupyter, Zenodo and Binder (2019)³³.
- Reproducible big data analytics with Jupyter notebooks and Binder technology (2019)³⁴.
- The 2020 CODATA-RDA School of Research Data Science (virtual)³⁵.

To respond to the requests of using advanced computing solutions for supporting data-intensive research in EOSC, additional training events were organized to introduce:

- The PaaS solutions (e.g. the Infrastructure Manager and the INDIGO-DataCloud PaaS Orchestrator) available in EOSC to facilitate the development and the deployment of complex applications in a cloud-based environment.
- The Workload Management service to distribute computing tasks in an efficient way while maximising the usage of computational resources.
- The status of the HTCondor and ARC CE open-source high-throughput computing software.
- CVMFS the solution to provide a scalable, reliable, and low- maintenance software distribution service.

²⁸ <https://blog.jupyter.org/binder-with-zenodo-af68ed6648a6>

²⁹ <https://www.eapconference.org/eapec-2019>

³⁰ <http://indico.ictp.it/event/8706/>

³¹ <https://indico4.twgrid.org/indico/event/8/session/9/?slotId=0#20190402>

³² <https://indico.egi.eu/indico/event/4431/timetable/#20190507>

³³ <https://www.eapconference.org/co-located-events>

³⁴ <https://lnu.se/mot-linneuniversitetet/aktuellt/kalender/2019/open-science-with-jupyter-and-zenodo/>

³⁵ <https://codata.org/initiatives/strategic-programme/research-data-science-summer-schools/2020-codata-rda-school-of-research-data-science-virtual/>

- The VMOps Dashboard³⁶ that provides transparent access to federated clouds.
- How to use Accelerated Computing in EOSC.

For more details, please refer to Section 4.1.

2.3.3 Data Management services

A valuable opportunity for researchers and communities to improve their skills and understanding about the services provided by EOSC to manage scientific data from analysis to long term archiving was the EUDAT - PRACE summer school³⁷. Through dedicated lectures and hands-on sessions the 26 Masters and PhD students and researchers from diverse disciplines discovered how EUDAT and PRACE services can contribute to support a typical research data life cycle (from data discovery, data processing, and data analysis to data preservation and publishing).

A dedicated training about sensitive data management and EUDAT services was organized by SURF³⁸. The aim of the course was to raise awareness for the users about the security and privacy aspects of sensitive data and for them to obtain a practical understanding of security measures of our services. The training session included a hands-on part in order to show how to work with tools that facilitate secure data sharing and storage. The target audience of the course were Data Stewards, Data managers, and interested Researchers. More than 30 participants from research institutes around Europe have registered for the course.

Recent updates related to the status of the EGI DataHub were also introduced during the EGI Webinar Programme. EGI DataHub is the service based on Onedata³⁹ software stack that provides distributed data management platform providing eventually consistent, POSIX-compliant virtual filesystem.

2.3.4 IT Security Forensics

Dedicated IT Security Forensics training sessions were organized to present the potential security threats related to the use of containers in EOSC. The number of (cumulative) training events on common, security forensics and federated services planned and delivered by WP11 is monitored by the project specific metric M11.4.

³⁶ <https://dashboard.appdb.egi.eu/>

³⁷ <https://eudat.eu/summer-schools/eudat-prace-summer-school-2019>

³⁸ <https://www.surf.nl/en/agenda/webinar-sensitive-data-management-and-eudat-services>

³⁹ <https://onedata.org/>

Table 3 - WP11 metric for T11.4

ID	Definition	Baseline	Value M06	Value M12	Value M24	Value M39
M11.4	Number of training events on common, security forensics and federated services	0	4	5	20	36

2.4 Domain-specific training for data providers and data scientists

In this section we report the domain-specific training events organized by T11.5 to target the data providers and data scientists within structured scientific communities linked to the Thematic Services and Competence Centres. The number of domain-specific modules produced by WP11, to target the needs of data providers and data scientists is monitored by the metric M11.5:

Table 4 - WP11 metric for T11.5

ID	Definition	Baseline	Value M06	Value M12	Value M24	Value M39
M11.5	Number of training materials delivered to target data providers and data scientists linked to Thematic Services and Competence Centres	0	13	103	160	170

2.4.1 Thematic Services (TSs) training activities

2.4.1.1 CLARIN (T7.1)

CLARIN organized an online training session⁴⁰ in March 2020 on the topic of connecting language processing tools to the Language Resource Switchboard. In addition, also the integration of catalogues (resource providers) was presented. All of the training material remains available online in the form of screencasts and textual instruction material.

⁴⁰ <https://www.clarin.eu/event/2020/centre-meeting-2020#switchboard>

Additionally, a screencast⁴¹ featuring the CLARIN thematic services, was published in November 2020. The idea behind this was to provide a low-threshold introduction for potential new users, which is also accessible and citable as Virtual Collection⁴².

2.4.1.2 DODAS (T7.2)

This Thematic Service organized a total of 8 events. More specifically, the 3rd. edition of the School on Open Science Cloud⁴³ was organized over two tracks: “Machine Learning Methods and Applications” and “Computing Infrastructure”. The first track provided a comprehensive introduction to basic ML and DL concepts as a way to set a common ground between participants, while the second focused more on providing students both with basic and advanced knowledge on how to build and use cloud-based infrastructures to implement workflows based on technologies and tools such as Hadoop, HDFS and Spark.

The main objective of the second event “Istanziamento e utilizzo di batch system on demand su infrastrutture Cloud”⁴⁴ (in Italian) was to provide researchers support on how to perform data analysis on cloud-based infrastructures.

During the “Big Data Management infrastructures and Analytics”⁴⁵ event (in Italian) participants were instructed on how to use cloud-based infrastructure and high-level solutions for big data analytics (e.g.: HDFS, Spark and Kafka), and adopt ML technology to extract valuable insights from datasets.

Additionally:

- The status of the DODAS TS was presented during the Jennifer workshop organized at CERN;
- A 2-day training event was organized for Lebanon researchers;
- Dedicated technical lectures and hands-on sessions on clouds, and related scientific services, were also presented during the CODATA-RDA Schools in 2019⁴⁶ and 2020⁴⁷ and during the “Multiscale, Machine learning and QSAR (MM-QSAR) Methods applied to biomolecules” schools in 2020 as part of the Elective course Master in Theoretical Chemistry and Computational Modelling (EMTCCM)⁴⁸.

⁴¹ <https://www.clarin.eu/blog/clarin-services-european-open-science-cloud>

⁴² <http://hdl.handle.net/11372/VC-1034>

⁴³ <https://agenda.infn.it/event/19049/>

⁴⁴ <https://agenda.infn.it/event/20268/>

⁴⁵ <https://agenda.infn.it/event/20847>

⁴⁶ <http://indico.ictp.it/event/8706/>

⁴⁷ <https://codata.org/initiatives/strategic-programme/research-data-science-summer-schools/2020-codata-rda-school-of-research-data-science-virtual/>

⁴⁸ <https://www.master-up.it/ml.php>

2.4.1.3 ECAS (T7.3)

In the last part of the project the ECAS Thematic Service organized a training event during the “EUDAT CDI - PRACE Summer School⁴⁹ on managing scientific data from analysis to long term archiving” (23-27 September 2019). During the session, the ENES Climate Analytics Service was introduced to help researchers in the development and evaluation of state-of-the-art in climate and Earth system models, encourage exchanges of software and results, and help in the development of high-performance computing facilities dedicated to long high-resolution and multi-model ensemble integrations.

A training about the features offered by the ECAS Thematic Service was given in the context of the “Online Training on Data Analytics”⁵⁰ (8-9 March 2021) jointly organized with the IS-ENES3 H2020 project. The online training event was open to anyone interested in processing multidimensional data, mainly targeting Earth system science and data science. The aim of the training was to show how to exploit ECAS features to load, filter, concatenate, perform spatial and time averages on data extracted from a huge catalog, plot CMIP data for multi-model comparisons and use CMIP models to calculate climate indices for any location and period.

2.4.1.4 GEOSS (T7.4)

CNR-IIA co-organized the ERA-PLANET GEOEssential workshop held on September the 1st. 2021 where the team presented the VLab service and its support of EOSC as a computing infrastructure. The service was also presented in other virtual conferences organized in the context of ERA-PLANET project. In addition, several dedicated online activities were organized to support VLab users requesting to use EOSC via VLab in the context of different H2020 projects, including ERA-PLANET and E-shape projects. In the first half of 2021, CNR-IIA also organized a dedicated training webinar for the use of VLab in the context of the H2020 E-shape project. Finally, the VLab service was presented as part of an Early Adopter Use Case in the context of the recently kicked-off project EGI-ACE.

2.4.1.5 OPENCoastS (T7.5)

OPENCoastS organized a total of 6 events. In all these events, the OPENCoastS Platform for On-demand Real-Time Forecasting of Coastal Zone was presented. Training events included hands-on demos and lectures. Since the new version of OPENCoastS was released in July 2020, a new e-training event was organized in Jan. 2021. This last training event included both sessions for users, devoted to the use of the platform, and a service deployment, for IT personnel.

Additionally, OPENCoastS organized a hands-on course on OPENCoastS integrated in the Master’s Degree in Coasts and Ports of the University of Cantabria (Santander, Spain), course 2019-2020.

This course covered the relevant topics to help everyone using the service, in 4 modules. Goals: (a) Empower potential users by providing an introduction to the relevant physical processes,

⁴⁹ <https://eudat.eu/summer-schools/eudat-prace-summer-school-2019>

⁵⁰ <https://is.enes.org/events/trainings-and-education/joint-is-enes3-eosc-hub-online-training-event-on-data-analytics-with-enes-climate-analytics-service-ecas>

the numerical model SCHISM and unstructured grid generation; (b) Introduce OPENCoastS, an innovative and free platform to generate on-demand forecasts, and (c) Minimize the learning effort through a step-by-step tutorial on the use of OPENCoastS.

2.4.1.6 *WeNMR (T7.6)*

WeNMR organized a total of 27 training events, including workshops, summer schools and webinars, to update the structural biology communities about the latest capabilities of the platform developed by the EOSC-hub Thematic Service. This platform used for the modelling of biomolecular complexes is freely accessible to non-profit users as part of the WeNMR services. In 2020, thanks to the cooperation already established between EGI, the Open Science Grid (OSG), and various high energy physics sites that committed to support COVID-19 related research, the platform has also been used to support COVID-19 research. During this research activity the HADDOCK platform was able to more than double its processing capacity, serving on average ~550 active users per months, >11,000 simulations related to COVID-19 (the equivalent of ~1.5 million HTC jobs, ~2.7 million CPU hours) on the EGI and OSC grid resources over the months of April to September 2020. In the context of EOSC-hub, WP11 also contributed to coordinate the development of REST APIs to allow extracting users' statistics from the WeNMR portal. In terms of training activities:

- The new release of HADDOCK platform was introduced with lectures and hand-on sessions in several training events, including the last International Symposium on Grids & Clouds 2021 (ISGC2021)⁵¹.
- A new tutorial about the modelling of antibody-antigen complexes was developed.
- Tutorials and lecture recordings are available online⁵².
- Recordings of lectures given at the BioExcel summer school are available from the BioExcel YouTube channel⁵³.
- A recorded webinar introducing the new features of the WeNMR HADDOCK2.4 webportal is also available online⁵⁴.
- A HADDOCK workshop was organized at the International Symposium on Grid and Cloud computing in Taipei⁵⁵ (online) in March 2021 for which a new training material was developed in the form of a tutorial on modelling membrane complexes⁵⁶.

2.4.1.7 *DARIAH (T7.8)*

DARIAH organized a one-day training event in Feb 2020 (in Croatian) with 30 participants. The aim of the event was to promote the solutions developed in the context of the EOSC-hub Thematic

⁵¹ <https://indico4.twgrid.org/indico/event/14/session/0/?slotId=0#20210322>

⁵² <https://www.bonvinlab.org/education/>

⁵³ <https://www.youtube.com/playlist?list=PLzLqYW5ci-2clo9v1AmhxuxlMQTj3sVVVo>

⁵⁴ <https://www.youtube.com/watch?v=9dWdaJ5jBqo>

⁵⁵ <https://indico4.twgrid.org/indico/event/14/page/4>

⁵⁶ <https://www.bonvinlab.org/education/HADDOCK24/LightDock-membrane-proteins>

Services to target the needs of the Digital Humanities community and to demonstrate the data analysis and knowledge extraction tools with application in Digital Humanities.

2.4.1.8 *LifeWatch (T7.9)*

LifeWatch organized a total of 6 meetings. These meetings focused on explaining how to use the GBIF.ES Biodiversity Data Portal, relying on the performance of guided exercises on its main functionalities.

2.5 Competence Centres (CCs) training activities

In this section we report the training activities delivered by the EOSC Competence Centres during the final reporting period.

2.5.1 **Fusion (T8.2)**

The Competence Centre held internal training events with individual users on helping them to migrate their workloads from traditional batch systems to the cloud.

2.5.2 **EISCAT_3D (T8.4)**

The Competence Centre organized the International EISCAT Radar School. The school was organized with support from EISCAT and NSF in the United States. For this reason, an equal number of EU and US students were accepted. The school was intended for new users of incoherent scatter radars at any stage of their research career. Participants were expected to have a background in space physics, ionospheric physics, plasma physics, or radar (radio science). During the 5 full day school covered all essential aspects of incoherent scatter radar theory and practical use. In total 15 teachers, representing both EU and US organizations including: The University of Oulu (Finland), EISCAT Scientific Association (Sweden), MIT (US), SRI International (US), and STFC (UK), were present at the school. The school was attended by 37 participants (25 male, 12 female) from 8 different countries. A user guide to access the DIRAC portal for EISCAT data was prepared for this event and made available in the EOSC-hub training catalogue. Three additional events on how to use the EGI Check-In service and management were organized in August-October 2020.

2.5.3 **EPOS-ORFEUS (T8.5)**

To promote the main results achieved in the context of the EOSC-hub project, as part of the intensive outreach and training activities, the EPOS-ORFEUS Competence Centre organized a total of 4 events during the period (M19-M39):

- In September 2019, the Competence Centre organized the International training course on ocean bottom and amphibian experiment seismological data.
- In October 2019, during the EPOS Seismology Workshop a keynote talk was given on EIDA services and how to benefit from them.
- Two webinars have been organized for the community at a global scale in the last months of 2020. In November, the “EIDA Data Access” with an attendance of 140 scientists from all around the world; and in December, the “EIDA Authentication and Authorization” with

around 50 participants registered. The recordings and presentations of these webinars can be found in the ORFEUS website.

The EPOS-ORFEUS Competence Centre also contributed with an article in the EOSC-hub Magazine (Issue no.7, Jan. 2021).

2.5.4 Radio Astronomy (T8.6)

The Competence Centre organized 3 internal training events to facilitate the knowledge sharing on the usage of EOSC infrastructure in the radio astronomy communities. Additional 2 internal training events were organized to promote the dCache authentication tokens and introduce the Advanced dCache API. A training session on the LOFAR processing workflow was organized in March 2021. The event was attended by 40 participants.

2.5.5 ICOS (T8.7)

Delays and COVID-19 situation complicated the work at the ICOS Ecosystem Thematic Center (ETC) in Italy considerably. For this reason about $\frac{1}{3}$ of the available budget, reserved for the implementation of the web-based interface, and training activities has not been used.

2.5.6 Disaster Mitigation (T8.8)

The Competence Centre continued to promote the collaboration framework between EGI and APAN (Asia Pacific Advanced Network) and promote the uptake of services, tools and best practices to support the Disaster Mitigation Working Group (DMWG). During the last part of the project, the EOSC-hub Competence Centre organized 4 training events in Malaysia (2), Myanmar and Bangladesh. During these events several training modules, based on case studies, and case study discussions, were produced to share the experiences of disaster risk analysis. Overall, a total of 9 training events were organized during the whole duration of the project. Dedicated sessions were also organized during the Annual International Symposium on Grids & Clouds 2021.

3 Webinar Programmes

In 2020 the severe measures introduced by different European countries to contain the outbreak of the coronavirus contributed to drastically reduce the number of training events. Most of the planned training events originally scheduled early in 2020 were cancelled or postponed to the second part of the year. To mitigate the impact of COVID-19 on the WP11 training programme and respond to the *“Recommendation 1.2: run a good number of workshops and training activities”*, online webinars were organized, when possible, to continue the engagement and the promotion of the EOSC-related services and promote the uptake of community-specific/proof-of-concept services developed by the Thematic Services and Competence Centres. In the following section, we report a brief summary of the webinar programmes organized by EGI and EOSC-hub to remain in contact with the targeted scientific communities and continue with the training and outreach activities during the pandemic.

3.1 The EGI Webinar Programme

EGI in EOSC-hub is contributing not only by offering the core services, but also with some of the federated and the common services constituting the EOSC-hub service portfolio. During the pandemic, a Webinar Programme⁵⁷ was organized in order to remain in contact with project members and other scientific communities interested in learning the latest updates of the services and understand how they can be used for supporting research. The first part of the EGI Webinar Programme started in April 2020 included 11 webinars and was attended by more than 270 participants from 30 different countries worldwide. During the first part of the webinar programme, the following topics were presented:

- Introduction of the EGI Cloud Compute Service.
- The EGI Notebooks service: Support for analytics and big data visualisation in the cloud.
- The EGI AAI Check-In service for scientific communities.
- The EGI Datahub to federate distributed data sets for data-intensive applications in the cloud.
- Managing Docker containers and Kubernetes clusters in the EGI Cloud.
- Introduction to HTCondor-CE service.
- The Infrastructure Manager (IM).
- Create elastic virtual clusters in the EGI Cloud with the EC3/IM.
- Rolling out ARC6 CE.

⁵⁷ <https://www.egi.eu/webinars/>

The second part of the EGI Webinar Programme restarted in October with new additional topics:

- CernVM-FS for Containers.
- DIRAC services for EGI users.
- EGI Operations and responsibilities of an NGI.
- Chipster – a user-friendly bioinformatics software for analyzing high-throughput data.
- Accessing federated clouds in EGI with VMOps Dashboard.
- How to use Accelerated Computing in EGI Federated Cloud.

The full Webinar Programme was attended by more than 420 participants. All the training materials and video recordings are also available on the EGI YouTube channel⁵⁸.

3.2 The EOSC Webinar Programme

Similarly to EGI, EOSC-hub organized a series of webinars to showcase the work and key results of the project. The EOSC-hub webinar programme officially started on September 15th with 6 webinars:

- EOSC business model recommendations.
- Integrating services into EOSC through EOSC-hub.
- EOSC-hub Technical Architecture.
- EOSC Portal orders - The user and provider perspectives.
- Using compute services in EOSC: Experiences and advice from EOSC-hub.
- Managing services and resources for EOSC.

The EOSC-hub webinar programme was attended by a total of 157 participants from 20 different countries worldwide.

⁵⁸ <https://www.youtube.com/user/EuropeanGrid/videos>

4 The EOSC-hub catalogue of training materials


As reported in D11.2, the architecture of the EOSC-hub catalogue of training materials was completely revamped in 2020 to address the *Recommendation 1.4 - "The Portal should include a set of training material focused on EOSC use, distinguishing between different audiences, as a minimum, users vs. service providers"*. In the new catalogue, dedicated paths were created for helping scientific communities and service providers to get easy access to the relevant documentation. The front-page of the catalogue of training materials is shown in Fig. 2.

Resources for training

[View](#)
[Edit](#)


This section provides a list of training materials being used during the EOSC-hub training events.

Researchers & Communities



- What is EOSC
- Fundamentals of Open Science
- Overview of the Research Data Management
- Access to domain-specific platforms
- Common & Federated services

Service Providers



- Overview of the EOSC-hub service architecture
- IT FitSM
- IT Security Forensics
- Operations
- Documentation






[Training Calendar](#)

[Guidelines](#)

[Ask for support](#)

[Search](#)

Fig. 2 The frontpage of the EOSC-hub catalogue of training materials

Over the last year the EOSC-hub catalogue of training materials has been further improved and populated with additional contents developed by the partners. Moreover, with the recent development of the EOSC Portal and the automatic procedure for onboarding new service providers in EOSC, additional information has been made available in the catalogue.

4.1 Training events

During the EOSC-hub project WP11 facilitated the organization and the delivery of 204 training events:

Training events during the period 2018 - 2021

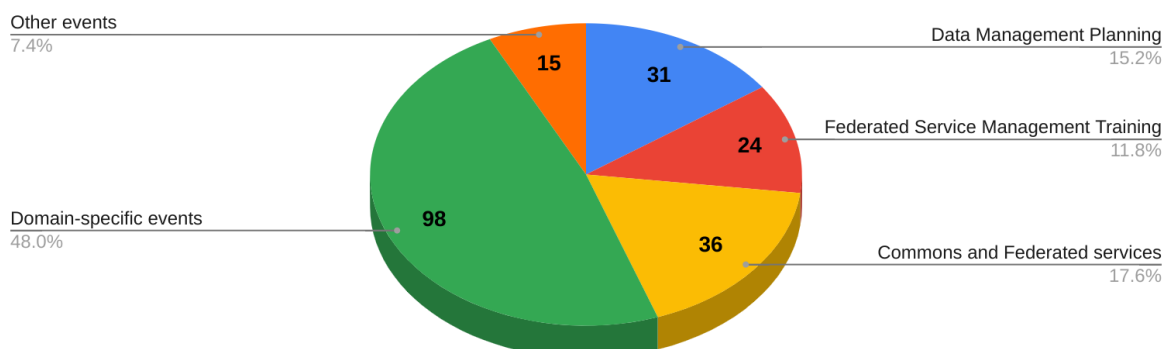


Fig. 3 - Total's training events organized (and/or supported) by WP11

4.2 Training materials

By the time of writing of this report a total of 246 training modules⁵⁹ have been produced. The breakdown of these training modules developed by the EOSC-hub Competence Centres and Thematic Services grouped per categories is reported in Fig. 4 and 5:

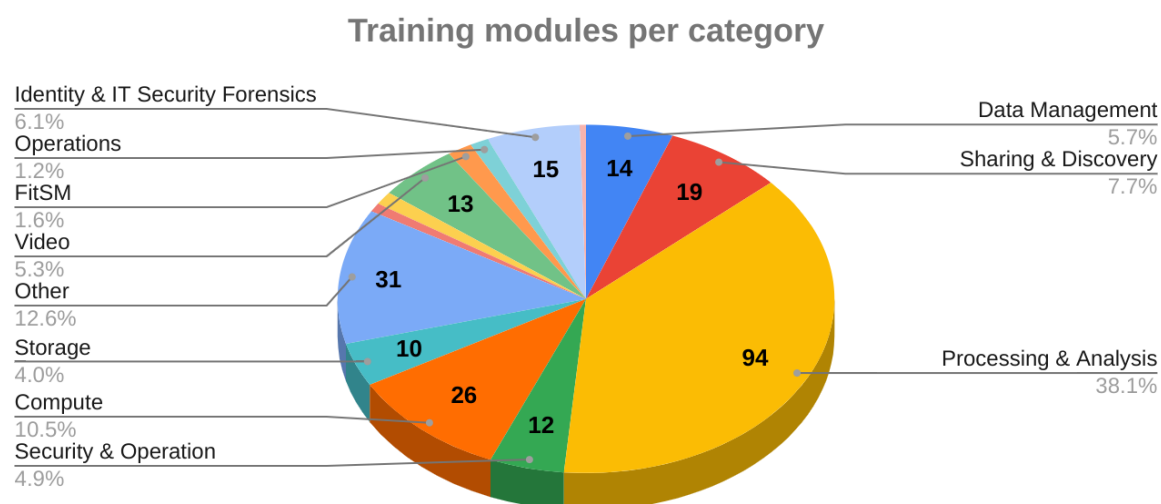


Fig. 4 - Training modules produced grouped per category

⁵⁹ Some modules were used several times during training events.

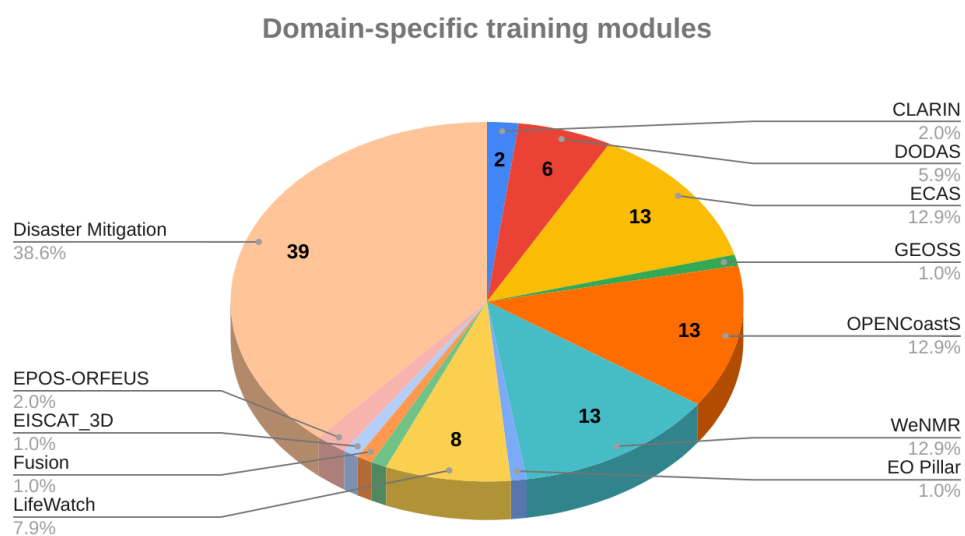


Fig. 5 - Training modules produced by Thematic Services and Competence Centres

5 Statistics from the training catalogue

The updated monthly view of the new EOSC-hub Training catalogue and the top 10 page views of the EOSC-hub training catalogue are reported in Fig. 6 and Fig. 7.

#Views of the EOSC-hub Catalogue of Training materials

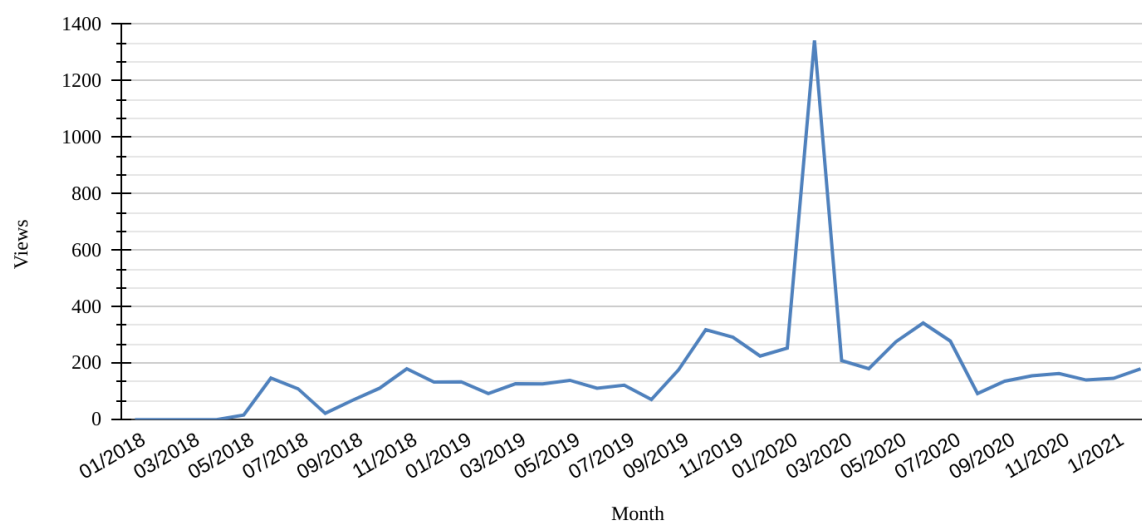


Fig. 6 - Monthly views of the training catalogue over the 30 months of the EOSC-hub project

From the statistics shown in Fig. 6 a significant increase of the number of views of the training contents in Feb. 2020 is reported. This happened when the Workshop on Training in EOSC was organized in The Hague in collaboration with other training coordinators involved in EOSC projects and associated organizations. The second increase was recorded in June 2020 when the revamped catalogue was officially announced (see EOSC-hub Magazine, Issue no. 6⁶⁰).

⁶⁰ <https://eosc-hub.eu/news/eosc-hub-magazine-issue-6>

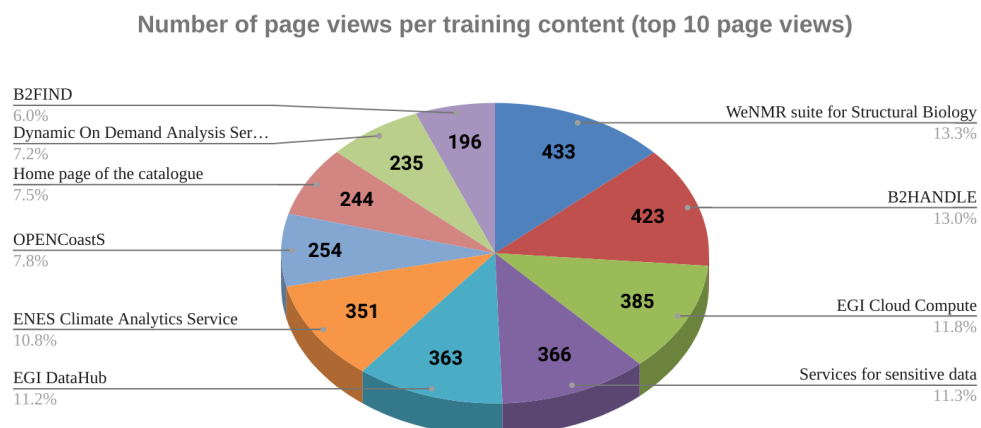


Fig. 7 - Top 10 EOSC-hub Training page view

From Fig. 7, we report a significant increase of the number of visits for some of EOSC-hub Thematic Services such as: WeNMR, ENES, DODAS and OPENCoastS. This positive trend is also justified by the registration of these thematic services in the EOSC Portal and the organization of targeted training activities to promote the new solutions for reaching a wider user base. From the figure we also spot a progressively interest in using EOSC-hub common services to operate with data such as: B2FIND and B2HANDLE and services to analyse sensitive data.

6 Conclusions and future outlooks

To contribute to the implementation of an effective training strategy, all the scientific communities involved in the project (mainly from WP7, 8 and 9) were invited to provide training needs. Over the years, these inputs were analysed by WP11 to identify possible skills gaps, update the training offer prioritizing the training topics to be delivered in the short and long terms, and deliver specialized training programmes to address the needs of service providers who might benefit from technical assistance on using, integrating and providing services in EOSC, or individual researchers possibly encountering the e-Infrastructures for the first time, enabling a smooth integration into EOSC ecosystem and maximising the benefits. Overall, during the three years the project organized a total of 204 training events which were attended by more than 5200 participants.

During the last part of the project EOSC-hub has also played an active role to promote the implementation of a focused training strategy to support EOSC in all its phases and ensure its uptake. In 2021, with the kick-start of the EOSC-Future project, the coordination of the training activities in EOSC will be led by OpenAIRE. In EOSC-Future, the main objective of WP9 - Training and Skills - is to set up a comprehensive and pervasive training infrastructure to enable EOSC users to share, publish, discover, and use resources (data, publications, software, services, etc.) and apply the OS/FAIR principles. All training activities will address skills development on two fronts:

- Data: Providing researchers and data practitioners with consolidated cross-infrastructure training packages for data skills, data science and data stewardship as key enablers for EOSC; and
- Services: Providing IT and library personnel of RPOs, RIs with training packages for aligning with EOSC-Core services (e.g., PIDs, AAI) and for providing services in EOSC-Exchange.