Deliverable D2.2

Training Program

Responsible Partner:	SURF
Status-Version:	Final – v2.1
Date:	29/10/2021
Distribution level (CO, PU):	Public

Project Number:	GA 101017207
Project Title:	DICE: Data infrastructure capacity for EOSC

Title of Deliverable:	Training program
Due Date of Delivery to the EC	31.10.2021
Actual Date of Delivery to the EC	29.10.2021

Work package responsible for the Deliverable:	WP2 – Outreach, stakeholder engagement and service uptake
Editor(s):	Claudia Behnke (SURF)
Contributor(s):	Anna-Lena Fluegel (DKRZ) Rene van Horik (DANS) Ellen Lennards (DANS)
Reviewer(s):	John Favaro (TRUST) Rob Carrillo (TRUST, EUDAT)
Recommended/mandatory readers:	WP4 & WP5

Abstract:	This document includes the inventory of the existing training material related to data storage and management solutions, identifies the gaps and defines the activities and the strategy that will be put in place to fill those training gaps and maximise the uptake of the DICE services. The plan will be developed by liaising with the INFRAOESC-03 (EOSC Future) and INFRAOESC-07 projects.
Keyword List:	Training, Hackathon
Disclaimer	This document reflects only the author's views and neither Agency nor the Commission are responsible for any use that may be made of the information contained therein

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

Document Description

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
v0.1	19.08.2021	Draft Table of Contents and text contributions	Claudia Behnke
v0.2	24.09.2021	Final text	Claudia Behnke
v0.3	30.09.2021	Review by the task members	
v0.4	19.10.2021	Comments and suggestions received by internal reviewers	
v2.0	28.10.2021	New version addressing reviewers' comments	Claudia Behnke
v2.1	29.10.2021	Final cleaned version for submission to the EC	PMT

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

Table of Contents

List of Figures

List of Tables

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

Terms and abbreviations

ASTRON	Astron
BSC	Barcelona Supercomputing Center - Centro Nacional de Supercomputacion
CESNET	CESNET, z. s. p. o.
CINECA	Cineca
CSC	CSC – Tieteen Tietotekniikan Keskus Oy
CyI	The Cyprus Institute
Datacite	DataCite
DKRZ	Deutsches Klimarechenzentrum GmbH
DoA	Description of Action
EC	European Commission
EOSC	European Open Science Cloud
ETHZ	Eidgenössische Technische Hochschule Zürich
EU	European Union
EUDAT ltd	EUDAT ltd
FZJ	Forschungszentrum Juelich Gmbh
GA	Grant Agreement to the project
GRNET	National Infrastructures for research and technology
GWDG	Gesellschaft für Wissenschaftliche Datenverarbeitung mbh Göttingen
INFN	Istituto Nazionale di Fisica Nucleare
IT4I	Vysoka Skola Banska - Technicka Univerzita Ostrava
KIT	Karlsruhe Institut für Technologie
KNAW-DANS	Koninklijke Nederlandse Akademie van Wetenschappen
KPI	Key Performance Indicator
MPG	Max Planck Gesellschaft zur Foerderung der Wissenschaften e.V.
PID	Persistent Identifier
SIGMA	SIGMA2
SNIC	Uppsala Universitet
SURF	SURFsara BV
TRUST	Trust-IT services
UCL	University College London
ULUND	University of Lund
VA	Virtual Access
WP	Work Package

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

Executive Summary

This document includes the inventory of the existing training material related to data storage and management solutions, identifies the gaps and defines the activities and the strategy that will be put in place to fill those training gaps and maximise the uptake of the DICE services.

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

1 Introduction and goal of the deliverable

As described in the Description of Action, "The Data Infrastructure Capacities for EOSC" (DICE) is a network of computing and data centres and research infrastructures. The goal is to enable a European wide storage and data management infrastructure for EOSC, providing generic services and building blocks to store, find, access and process data consistently and persistently. Specifically, DICE partners will offer 14 state-of-the-art data management services together with more than 50 PB of storage capacity. The service and resource provisioning will be accompanied by enhancing the current service offering to fill the gaps still present to support the entire research data lifecycle. DICE will provide solutions for increasing data quality and re-usability, supporting long-term preservation, managing sensitive data, and bridging between data and computing resources."

DICE published "the Outreach, Communications and Dissemination Plan"[1] to reach this goal and ensure uptake of the services. This deliverable completes the plan and describes the objectives of training actives in DICE.

This document includes the inventory of the existing training material related to data storage and management solutions, identifies the gaps, and defines the activities and the strategy that will be put in place to fill those training gaps and maximise the uptake of the DICE services.

The main activities of the task will be the organisation and facilitation of four hackathons and one datathon, which will try to fill the gaps identified in this document. Currently, the hackathons are planned in the months M15, M20, M24 and M28 of the project (see Figure 1), but since they strongly depend on the needs of the communities and development, they might slightly move in time and will be adapted in the content. The hackathons are planned and budgeted as physical events, but hybrid settings are also considered with the knowledge of the CoVID19 pandemic.

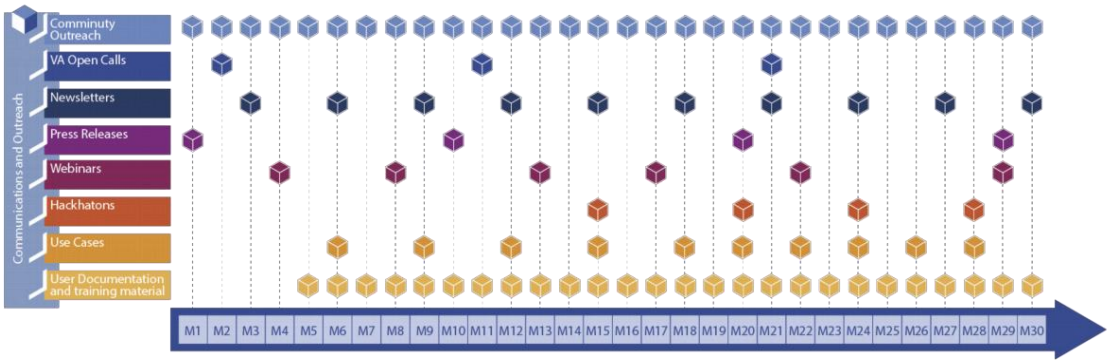

Figure 1: DICE overall communications timeline

1.1 Collaboration with EOSC Future

From the start, a close collaboration with EOSC Future is anticipated since this project has much more resources for developing training material and organising training. EOSC Future will focus on the content of the EOSC Portal[2].

[1] https://www.dice-eosc.eu/deliverables/outreach-communications-and-dissemination-plan
[2] https://marketplace.eosc-portal.eu/services/c/training-support

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

An EOSC Future & INFRAEOSC-07-2020 Collaboration Agreement is set up, and the training activities are explicitly mentioned in this. The goals are:

- Joint EOSC training activities,
- Promotion of training and training materials developed by the EOSC Future and the INFRAEOSC-07 projects,
- Co-development of EOSC training materials and the EOSC Future Knowledge Hub,
- Joint hackathon/s or summer school/s.

Training of EOSC future will align with the EOSC technical roadmap and fill gaps in the disciplines. This is complementary to the planned activities in DICE and will also help users of DICE virtual access resources to use those more efficiently.

1.2 Document

The deliverable is organised in the following section:

- An overview at the consortium level and within EUDAT on the current training plans and material.
- An overview of the trainings planned by DICE tasks.
- Requirements in terms of training from the DICE use cases.
- Conclusions and training events planned from DICE.

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

2 Existing Training

In this part, we will give an overview of existing training and material. At this moment, the focus is on training for data management and storage solutions to distinguish them from more general courses on computing facilities and software or programming languages.

For this, all partners of DICE were asked to submit a form about the training they are currently planning to provide. Furthermore, we approached training coordinators of the EUDAT B2services.

2.1 Training from DICE Partners

Seven partners confirmed that they provide training courses on data management and storage. Four partners do not give any training related to data management or storage. Naturally, most of the time, partners are focusing on the infrastructure they provide. This offers possible users the chance to get acquainted with the services but can also be used to increase the uptake since potential users might not have been aware of an offer until they have attended the training. It is very positive that many partners also provide training about the EUDAT Collaborative Data Infrastructure (or EUDAT CDI)[3] services.

Additionally, many of them provide trainings also in the framework of Partnership for Advanced Computing in Europe (PRACE)[4]. However, those training often focus on pure "HPC- skills" and not necessarily on data management and solution. Since one of the goals of DICE is to bridge between data and computing resources, a closer collaboration also on the partners' level seems positive.

The table on the following pages gives an overview of the existing and planned training from the DICE Partners, and the table will be periodically updated as new information becomes available.

[3] https://eudat.eu/
[4] https://prace-ri.eu/training-support/training/

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

Table 1. Training offered by DICE partners

Name Partner	Country Partner	Name of Training	Type (General training/service-related training)	Name of service	Target Audience (End-Users, Data Stewards, Researchers, Other [please specify])	Link to training	Status (in Preparation, Ready, Needs Updates)
SURF	NED	Introduction to Data Management and Research Drive	both	Research Drive	End-Users, Data Stewards	https://www.surf.nl/en/agenda/introduction-to-data-management-and-research-drive	Ready
SURF	NED	Using Persistent Identifiers for Research Data Management	Service	EPIC / B2HANDLE	data stewards	https://www.surf.nl/en/training-courses-for-research	Ready
SURF	NED	3x IRODs usage	Service	IRODS	data stewards	https://www.surf.nl/en/training-courses-for-research	Ready
RDNL (SURF/ DANS)	NED	Essentials 4 Data Support	general training		data stewards	https://datasupport.researchdata.nl/	Ready
SURF	NED	Introduction to Data Management and EUDAT services	service	EUDAT b2 services	End-Users, Data Stewards	https://www.surf.nl/en/training-courses-for-research	Ready
EUDAT (DKRZ)	GER	B2FIND User Training and B2FIND Integration	both	B2FIND	End-Users, Data Stewards	https://www.eudat.eu/b2find-training-suite-0	needs update
DataCite	GER	Onboarding	both	DOI registration	Librarians, data stewards, repository managers.	https://drive.google.com/drive/folders/1OP_kDRchkRHG9AyJvraw0ca9UIa4PSct?usp=sharing	ready
DataCite	GER	Getting started	Service	DOI registration	New organisations	https://datacite.org/assets/Getting%20started%20Guide.pdf	ready
CSC	FI	CSC's services for different stages of the research data life cycle	general training		End-Users, Data Stewards	https://ssl.eventilla.com/event/mkZ4k?_ga=2.217721054.910657602.1631794878-1808089618.1631080222	Ready
CSC	FI	CSC Training activities	general training		Various	https://csc.fi/en/web/guest/training-offering	Ready
UCL	UK	Writing data management plans	both	Research Data Management	All staff and research students	https://moodle.ucl.ac.uk/course/view.php?id=20015&section=1	Ready
UCL	UK	Open Science and Scholarship	general training	Research Support and Open Access	All staff and research students (UCL credentials required)	https://moodle.ucl.ac.uk/mod/lesson/view.php?id=2534343	Ready
GWDG	GER	Working with GRO.data	both	Research data repository/Dataverse	End-Users, Data Stewards, Researchers	https://www.gwdg.de/academy/course?course=10000	Ready
GWDG	GER	Working with GRO.plan	both	Research data planning/RDMO	End-Users, Data Stewards, Researchers	https://www.gwdg.de/academy/course?course=10004	Ready

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

Name Partner	Country Partner	Name of Training	Type (General training/service-related training)	Name of service	Target Audience (End-Users, Data Stewards, Researchers, Other (please specify))	Link to training	Status (in Preparation, Ready, Needs Updates)
DANS	NED	Introduction to the European Open Science Cloud (EOSC)	general training	Examples: CoreTrustSeal, FAIR data services, RDM	researchers, research institutions, research funders, data professionals	https://dans.knaw.nl/en/about/services/training-consultancy/training	Ready
DANS	NED	Integrating Resources in the European Open Science Cloud (EOSC)	both				
DANS	NED	Prepare your data for publishing and archiving	both				
DANS	NED	Tools for FAIR data	both				
DANS	NED	Using Knowledge Organisation Systems for FAIR data	both				
DANS	NED	Towards a FAIR enabling organisation	both				
DANS	NED	Train-the-trainer on RDM and Open Science	both				
DANS	NED	Managing a Dataverse repository and community	Service				
DANS	NED	Hands-on Data Management Planning	general training				
DANS	NED	Reviewing DMPs	general training				
ICOS @ ULUND	SWE	Introduction to ICOS data management	service	Data portal	internal and end-users	https://www.icos-cp.eu/cp-webinar#introduction-to-ICOS-data	Ready
ICOS @ ULUND	SWE	Notebooks on ICOS Exploredata	service	Exploredata	end-users	https://www.icos-cp.eu/cp-webinar#notebooks-on-exploredata	Ready
ICOS @ ULUND	SWE	Notebooks for atmospheric stations	service	Exploredata+Jupyter	end-users	https://www.icos-cp.eu/cp-webinar#notebooks-for-atmospheric-stations	Ready
ICOS @ ULUND	SWE	In depth: ICOSCP python library	service	Exploredata+Jupyter	end-users	https://www.icos-cp.eu/cp-webinar#icoscp-python-library	Ready
ICOS @ ULUND	SWE	Drop-in Q&A	general training	Exploredata+Jupyter	end-users	https://www.icos-cp.eu/cp-webinar#drop-in	Ready
ICOS @ ULUND	SWE	ENVRI Community International School	general training	Services for FAIRness	data stewards, researchers	https://envri.eu/event/save-the-date-envri-community-international-school-services-for-fairness/	29 Sept 2021

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

2.2 Training from B2 services

The EUDAT B2Services are vital components in the DICE service offer[5]. Therefore, we focus here on the current training material for these services and identify needs for updates, since the services might develop during the project and the existing training material might become outdated.

The page https://eudat.eu/training gives an overview of the existing material for the EUDAT B2services. The training contact of B2FIND noted that they plan to update the existing resources since they had significant changes in the service. The other services have not planned significant changes in the near future, so the training material should be up to date.

Table 2. Training material offered by EUDAT

Service	Trainings	Contact
B2ACCESS	No training material available	Sander Apweiler (JSC)
B2DROP	https://www.eudat.eu/b2drop-training-suite	Sander Apweiler (JSC)
B2FIND	existing, but outdated → needs updating	Anna-Lena Flügel (DKRZ), Heinrich Widmann (DKRZ),
B2HANDLE	https://github.com/EUDAT-Training/B2SAFE-B2STAGE-Training/blob/master/07c-Working-with-PIDs_B2HANDLE.md	Themis Zamani (GRNET)
B2SAFE	https://www.eudat.eu/b2share-training-suite	
B2SHARE	https://www.eudat.eu/b2share-training-suite	Chris Ariyo (SCS)

The training of the EUDAT suite is (like the services themselves) domain-agnostic. On the one hand, this caters for a wide range of users. On the other hand, this might also reduce the uptake since potential users might not know if the training/service suits them.

[5] https://www.dice-eosc.eu/call-service-requests

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

3 Trainings planned by DICE tasks

3.1 DICE Digital Preservation Service (DDPS)

There is one dedicated work package (WP4) in the DICE that focuses on the integration of data services offered via DICE with European platforms and infrastructures (or vice versa) and will make these services available in the other European platforms and infrastructures increasing their uptake and use in research workflows. One of the goal is to provide relevant EUDAT services with adequate long-term preservation (LTP) and implement these policies for at least the B2SHARE and B2SAFE services.

The DICE policies for long-term preservation will provide core EOSC data services with a basic and fundamental solution and associated business model to sustain research data in the long run. Without such policies and cost-recovery models, the future accessibility to and assessment of FAIR research data cannot be guaranteed. The LTP policy and service is named DDPS (DICE Digital Preservation Service).

For this implementation, the involved people have set up four training events which the Training task will support.

Training 1

Title: DICE Long Term Preservation Policy

Learning goal: To get acquainted with the DICE LTP policy

Target group: Policymakers, infrastructure managers, researchers

Trainers: Compilers of the DICE LTP policy

Format: Webinar (reusable)

Training 2

Title: Notification service for LTP

Learning goal: To understand DICE notification service principles (Signposting and LDN (linked data notification))

Target group: Developers, IT managers

Trainers: Creators of the notification service

Format: Webinar (reusable)

Training 3

Title: DDPS for B2SHARE instances

Learning goal: To be able to apply the DICE LTP policy (e.g., as “demonstrator”)

Target group: Developers (related to B2SHARE services)

Trainers: Creators of the DDPS service

Format: Workshop / Hackathon

The DICE project has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

Training 4

Title: DDPS as a robust service

Learning goal: To be able to assess the relevance of the DDPS service and to know how to implement it.

Target group: Policymakers, Service managers, Users

Trainers: Developers of the DDPS service

Format: Workshop (reusable)

While the first three events will focus on developers and IT managers, the last training is well suited as a hackathon from the training task since it will focus on the end-users of this newly developed service.

Possible users of the service are the LOFAR community which is also a use case; see Chapter Needs from Use Cases.

3.2 Sensitive Data

EUDAT services are not by default well suited for privacy-sensitive data. However, communities, especially from the health and social science domain, have an increasing demand for storing and managing this type of data. Within DICE, a dedicated task within WP4 will try to enhance the existing services such that they might be usable for this sensitive data, which might also create options for a datathon. Two partners in this task (SURF and DANS) contribute to the RDNL course, which recently piloted training on managing personal data. It will be targeted especially at data supporters[6]. Integrating (parts of) this course in the datathon, as mentioned earlier, is anticipated.

[6] https://researchdata.nl/en/curious-about-new-rdnl-courses-stay-tuned-for-new-material-on-gdpr/

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

4 Needs from Use Cases

One goal of the deliverable is also to understand the needs of the selected communities. Therefore, the use cases participating in DICE were contacted, and we asked mainly two questions:

- Is there already sufficient user training material on the single components/topics/challenges of this workflow?
  - If yes and openly accessible, we would like to know where to include it in our training inventory.
  - If not, which training topics (or training on the whole workflow?) could be interesting for DICE/EOSC users within the scope of your task?

In the following, the use-cases are introduced using the description of the work package from the Description of Action.

4.1 CompBioMed data platform integration (biomedicine)

CompBioMed[7] is a Centre of Excellence focused on the use and development of computational methods for biomedical applications. They have a training portal[8] that offers regular webinars about various topics to help the community users. Furthermore, CompBioMed offers training on-demand, which are collected in a training repository[9]. The repository holds 71 entries, where nine are using the tag "data management". However, when reviewing the remaining training, it becomes clear that only three cover data management. At the same time, the others focus on analysis methods for certain biomedical data types, which give the impression that data management is not in the focus of the community. This finding was supported during interviews with Marco Verdicchio and Carlos Teijeiro Barjas (both SURF) that agreed that there is a need for more data management training for the community.

> *In the context of DICE, CompBioMed seeks to exploit the third pillar of science to render predictive models of health and disease more relevant to clinical practice by providing a personalised aspect to treatment. One of the clear trends in the biomedical community is its ever-increasing demand for storing more data as well as the transfer, management and longer-term preservation of this data. Frequently, large data sets need to be moved closer to High-Performance Computing (HPC) services before performing computational work. Once the computational work is done, the resulting data is then transferred to somewhere else or kept closer to the HPC services for post-processing work.*
>
> *Examples of such types of computational workflows are cardiovascular simulations, in which researchers create high fidelity simulations of the human cardiac conduction system along with the full biventricular description of the cardiac muscle. Alya (Alya - High-Performance Computational Mechanics | BSC-CNS) is one of the main codes used within CompBioMed for cardiovascular modelling. Developed at the Barcelona Supercomputing Centre, Alya performs cardiac electro-mechanics simulations from tissue to organ level, focusing on a FEM-based electro-mechanical coupling solver optimised for HPC infrastructures.*
>
> *This work focuses on the development of a data management workflow to facilitate Alya simulations using large datasets and to explore the capabilities and challenges towards*

[7] https://www.compbiomed.eu/
[8] https://www.compbiomed.eu/training-3/
[9] https://www.compbiomed.eu/training/training-repository/

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

the use of Alya for Exascale simulations. This use case addresses the need for safe data replication and large data transfer within a system that can support a FAIR data cycle, an important data requirement within this international community. Parallel activities will also support this work within the CompBioMed consortium, where we will extend the results of this work and promote access to the workflow to different HPC (e.g., LRZ, EPCC), research and medical centres in the future.

The focus will be on research data (mainly synthetic and simulation data) services, and sensitive data will be not treated at this stage.

No training exists for this newly developed platform, so that the task will dedicate one of the hackathons for an (onboarding) user training. A possible combination with the yearly organised CompBioMed all-hands meeting, held yearly in June, is anticipated.

Even though the focus within DICE will be on research data (mainly synthetic and simulation data) services, there is a need for training on which services are suited for processing sensitive data by the CompBioMed community. In collaboration with the task on sensitive data, another hackathon or the datathon shall be held. This hackathon will not only focus on CompBioMed but also try to engage partners from other communities as well as industry.

4.2 LOFAR radio astronomy platform for advanced data discovery and preservation

The LOFAR Observatory operates LOFAR[10], a unique radio astronomical instrument with stations distributed over Europe, interconnected through a 10 Gbps wide area network and a central processing facility hosted by the Rijks Universiteit Groningen. Following initial processing on the central processing facility, observation data is stored in the Long-Term Archive (LTA) to assess the quality and reduce data volume. No training resources are mentioned on the dedicated website for this archive, but it contains exhaustive documentation.

ASTRON has started preparing for offering further data services associated with the instrument data archives. Initially in the EOSCpilot project, followed up in the EOSC-Hub project where LOFAR is at the core of the Radio Astronomy Competence Center, and now also as a partner in the ESCAPE project, various aspects of integration with the EOSC infrastructure are being evaluated and prototyped. At the end of 2020, this was concluded with a demonstration of a user requested data processing service and an advanced data product repository, built on integrating the EUDAT B2SHARE service and dCache. Service access will be offered via AARC compliant federated AAI solution, currently considering EGI Check-In. LOFAR community management is implemented to handle community and science project membership for use by the integrated services.

The proposed case will particularly follow up on the deployment and further development of the advanced data product repository, supporting initial operations for safe storage, discoverability, and distribution of science-level data generated by a user- and ASTRON managed processing services connected to the LOFAR LTA. As such, the case will develop a data service that supports manual and automated ingestion of processed data, in particular integrating with automated data processing services running on compute clusters co-located with the LTA data storage infrastructure. Since the processed data will typically be orders of magnitude smaller than the source (instrument) data, it is considered that the advanced data-product repository can, in its initial phase, be centrally hosted (i.e. hosted by a single partner). It offers data analysis services with low demand on computing

[10] https://www.astron.nl/telescopes/lofar/

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

resources in conjunction with the repository using an EOSC cloud compute service, e.g. a cutout service to generate specific cut out slices from data stored in the repository.

The proposed case addresses all aspects from Findability, Accessibility, Interoperability, and Reproducibility principles, focusing on the first two (FA).

The use case wants to have a production-ready service by March 2022, which will combine a DICE provided science data repository (B2SHARE or SDR), B2HANDLE PID and the usage of B2FIND for data discovery. This gives an excellent opportunity to host a hackathon for the users of this service.

4.3 ICOS community inversion benchmarking tool

The Integrated Carbon Observation System[11], ICOS provides standardised and open data from more than 140 measurement stations across 13 European countries. The stations observe greenhouse gas concentrations in the atmosphere and carbon fluxes between the atmosphere, the land surface and the oceans. Thus, ICOS is rooted in three domains: Atmosphere, Ecosystem and Ocean. We have been in touch with Alexander Vermeulen, the task leader of the related DICE task. On the website of the project, regular webinar training[12] is provided. This series of webinars will guide users to find, preview and use ICOS data. However, they seem not to focus on data management or storage or FAIR data related aspects.

Following the methane emission analysis in this task, more complex cases are planned to demonstrate the system's modularity by applying the approach to the assessment of natural CO2 fluxes over Europe using two models' setups from an earlier comparison in the framework of the EUROCOM project. Three inversion model setups (Carbon-tracker, LUMIA and STILT) are already in operational use at the institutes that work together in Carbon Portal (ULUND and Wageningen University).

This task will utilise the EUDAT data service B2SAFE to store the data from ICOS and to stage it for analysis at computing platforms. The suitability of the B2SHARE service for publishing results in addition to the ICOS Carbon Portal is investigated. ICOS data in B2SAFE will be given persistent identifiers via the B2HANDLE service in addition to the main Handle PIDs directly minted by ICOS that are based on the data object checksum. The case addresses all aspects from Findability, Accessibility, Interoperability, and Reproducibility principles, focusing on reproducibility.

The B2SAFE-ICOS-B2SHARE workflow for the inversion demonstrator is developed in an interactive Jupyter notebook. Therefore, the task is working on the design and cannot tell how the training material needs to be.

The development goal is to make the whole process seamless, so users only need training on the system's specific scientific ins and outs and the user interface.

Next, a general introduction to FAIR data was requested since the community lacks some awareness here.

[11] https://www.icos-cp.eu/
[12] https://www.icos-cp.eu/cp-webinar

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.

5 Conclusions

The training team started to look first at which training events and materials are available already among the partners and EUDAT initiative, so to avoid replicating material while focusing the effort on filling gaps in the current training plans.

From the analysis of what is missing, it was evident that two specific development of new data management features might benefit from specific trainings. At the same time, after identifying the communities' needs, it is clear that each use case should benefit from a hackathon dedicated to their work in DICE. The summary of the training events planned in DICE are reported in Table 3. Please note that the list might be still updated if other opportunities will become evident.

Since all use-cases should be in a production phase by next year (2022), the hackathons for the user uptake can be planned during this time.

In some cases, it might be beneficial to move the dates to create synergies with events from EOSC Future or the communities themselves. Furthermore, delays of development need to be taken into account, which means that the month and topics can be understood as suggestions.

Table 3. Suggested topics for DICE training events

	Month	Topics	Collaboration with
Hackathon 1	M15, Mar22	LOFAR platform for advanced data discovery and preservation	LOFAR
Hackathon 2	M20 Jun22	CompBioMed data platform integration	CompBioMed
Hackathon 3	M24 Aug22	Community inversion benchmarking tool	ICOS
Hackathon 4	M28 Feb23	DDPS as a robust service	DDPS
Datathon	TBC	Sensitive data with EUDAT services	Sensitive Data

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207.