D5.4 - OS and RDM learning paths for Climate Change Communities

Lead Partner:	University of Trento
Version:	1.0
Status:	==Submitted draft not yet approved by the EC==
Dissemination Level:	PU - Public
Document Link:	https://zenodo.org/uploads/12805957
DOI	10.5281/zenodo.12805957

Deliverable Abstract

This report is about the Skills4EOSC D5.4 deliverable on "*Open Science and Research Data Management learning paths for Climate change communities*".

It delivers the main results of Task5.4 on "OS and RDM in Climate Change" at M24. The report provides the context of Task5.4 and it illustrates its three main outcomes: learning path, training material, and pilot event.

COPYRIGHT NOTICE

DELIVERY SLIP

Date	Name	Partner/Activity	Date
From:	Sandro Fiore	UNITN	27/08/2024
Moderated by:	Irakleitos Souyioultzoglou	OPERAS	27/08/2024
Reviewed by:	Alessandro D'Anca Vincenzo Maltese	CMCC UNITN	11/08/24
Approved by:	Sara di Giorgio	GARR	29/08/2024

DOCUMENT LOG

Issue	Date	Comment	Author
v0.1	15/07/24	Initial draft	S. Fiore (UNITN)
v0.2	31/07/24	Draft ready for review	S. Fiore (UNITN), F. Antonio (CMCC), S. Sharma (TU-Delft), P. Martinez Lavanchy (TU-Delft), S. Jyoti Banerjee (UNITN)
v0.3	11/08/24	Document reviewed	
v0.4	26/08/24	Final version	S. Fiore (UNITN)
v1.0	29/08/24	Approved	Sara di Giorgio (GARR)

TERMINOLOGY

Terminology/Acronym	Definition
CMIP	Coupled Model Intercomparison Project
DMP	Data Management Plan
ENES	European Network for Earth System modelling
ESGF	Earth System Grid Federation
ESiWACE	Centre of Excellence in Simulation of Weather and Climate in Europe
FAIR	Findability, Accessibility, Interoperability, Reusability
IPCC	Intergovernmental Panel on Climate Change
IS-ENES	Infrastructure for the European Network for Earth System Modelling
NetCDF	Network Common Data Form
OS	Open Science
PID	Persistent Identifier
RDM	Research Data Management
RI	Research Infrastructure
ToT	Training of Trainers

Contents

Supporting | eosc | Co-funded by the European Union | UK Research and Innovation co-funded by

Executive summary

This report is about the Skills4EOSC D5.4 deliverable on *Open Science and Research Data Management learning paths for Climate change communities.* It presents the main results of Task 5.4 on "*OS and RDM in Climate Change*" at M24.

The report provides the context of Task 5.4 and illustrates its three main outcomes: learning path, training material, and pilot event.

The learning path is described in detail according to the three modules: (i) Introduction to OS and FAIR [R1], (ii) FAIR principles in practice and (iii) Planning for FAIR climate data. Information about the content of each module, its structure (in sub-modules) and discussion/group activities is also provided, jointly with the links to the corresponding training material.

A summary about the pilot held in February 2024 is included in the report, together with a plan for the next period about (i) the organisation of the official training event in Fall 2024, (ii) the review, extension and finalisation of the training material, as well as (iii) the consolidation of the learning path (minor changes only).

1. Introduction

The activities planned for Task 5.4 "OS and RDM in Climate Change" concern the design of learning paths on Open Science (OS) and FAIR (Findability, Accessibility, interoperability, Reproducibility) Research Data Management (RDM) for researchers working in the climate change domain. This task builds on a partnership which was at the core of key community initiatives like the Earth System Grid Federation (which delivers data useful for the IPCC Reports) and European projects such as *IS-ENES* (third phase) [R2] and the *Centre of Excellence in Simulation of Weather and Climate in Europe* (ESiWACE – second phase [R3]) thus ensuring a close link with the European Network for Earth System modelling (ENES) community.

Such initiatives have promoted over time several training events which have been fundamental to foster knowledge and share information within the community. However, they were mainly focusing on scientific topics (thus having a very limited scope with respect to OS and FAIR guiding principles) and they were mostly targeting scientists -without following a Train-of Trainer (ToT) approach, which is key within Skills4EOSC.

Thus, building a learning path for new trainers (ToT) on OS and FAIR RDM represents a more structured approach which fills the gaps with respect to the past training offering. ToT is key to enable the practice of Open Science in the climate change domain with adequate knowledge of standards, applications and tools and relevant best practices for delivering, managing, re-using, sharing and analysing FAIR data, as well as other digital research objects.

In terms of key objectives for the training, T5.4 focused on the following topics:
- Fostering the access and use of the IS-ENES data and services
- Fostering Open Science and RDM in Climate Science

Finally, the plan discussed with the partners in Task 5.4 included a pilot in early 2024 and an official ToT event in Fall 2024.

2. Learning path

2.1 Methodology

The learning path has been designed and improved, by iterating several times with the Task 5.4 partners to include inputs from all participants and build upon their comments and feedback.

The final learning path for researchers has been ultimately developed around 3 modules:

- Introduction to OS and FAIR;
- FAIR principles in practice;
- Planning for FAIR climate data.

As part of the 2nd year outcomes, four presentations with transcripts have been delivered as training material for this task and will be used in the official ToT event during Fall 2024 (preliminary drafts have been used in a pilot delivered in February 2024).

The material has been prepared following the FAIR-by-design methodology [R4]. The Skills4EOSC template has been adopted for the slides.

2.2 Description

As reported above, the final learning path for researchers has been ultimately developed around 3 modules (see Section 3.1).

A detailed description about each is reported in the following three subsections.

2.2.1 Module1: Introduction to OS and FAIR

This module is organised into two submodules:

- Introduction to Open Science in climate research
- Open Science and FAIR

2.2.1.1 Introduction to Open Science in climate research

The learning objectives the first submodule focuses on are:

- Identify the different elements of OS

- Recognise examples of how OS enables climate research
- Describe the current OS policy landscape in climate science

In terms of content, the submodule addresses the following aspects: (i) *How open science already works in several climate projects and initiatives at EU and International level (i.e., IPCC)*, (ii) *What kind of openly available projects/data/databases/software/climate models are accessible to researchers in the field* and (iii) *What to take into account or how to make your own research openly available*.

An initial "Buffet on Open Science" acting as an icebreaker and a final discussion on "*How relevant is Open Science and FAIR to climate research (or climate modelling research) specifically? Why so?*" are also activities foreseen in this submodule.

2.2.1.2 Open Science and FAIR

The learning objectives the second submodule focuses on are:

- Identify key elements of FAIR data principles
- Recognise the relationship between FAIR and OS, and how FAIR can enable OS
- Contrast/differentiate/compare Open vs FAIR research outputs

In terms of content, the submodule provides a general introduction to the FAIR principles, and some key elements (like Permanent IDs, licences, and so on) that enable researchers to implement these principles. The distinction between Open Science and FAIR is also emphasised, highlighting that they are not equivalent - showcasing ways in which outputs can be FAIR but not Open, or Open and not FAIR - with the help of some examples (both good and bad examples of how the two concepts can be operationalised).

Activities around these concepts are also part of this submodule and they include the use of climate datasets.

The material about this module is available at the following link: https://zenodo.org/records/12805958/files/OSandRDMforClimateChangeCommunities_Module1.zip

Element	Description
Title	Introduction to OS and FAIR
Abstract / Description	An introductory presentation about Open Science and FAIR principles. The role of OS in Climate Science and policy landscape in this domain.
Author(s)	Paula Martinez Lavanchy and Saba Sharma
Primary Language	English
Keyword(s)	Climate Science, Open Science, FAIR Principles.
License	CC BY 4.0 license https://creativecommons.org/licenses/by/4.0/
Version Date	13/08/2024
URL to Resource	https://zenodo.org/records/12805958/files/OSandRDMforClimateChangeCommunities_Module1.zip
Target Group	Researchers
Learning Resource Type	Presentation
Learning Objectives	- Identify the different elements of open science - Recognise examples of how open science enables climate research - Describe the current OS policy landscape in climate science - Identify key elements of FAIR data principles - Recognise the relationship between FAIR and OS, and how FAIR can enable OS - Contrast/differentiate/compare Open vs FAIR research outputs
Expertise (Skill) level	Beginners

Table 1. Module1: Introduction to OS and FAIR

2.2.2 Module 2: FAIR principles in practice

This module features a submodule about "FAIR principles in Climate Modeling Research".

2.2.2.1 FAIR principles in Climate Modeling Research

The learning objectives foreseen in this submodule are:

- Recognise the challenges in applying FAIR principles within climate modeling research
- Illustrate the key elements of FAIR principles in a specific climate infrastructure
- Recognising community standards available to implement FAIR in climate research (like NetCDF)
- Evaluate FAIRness of datasets in climate modeling research

In terms of content, the submodule presents a use case to showcase FAIR principles in a project or relevant community infrastructure (ESGF/CMIP), showing what data is collected, and the challenges regarding FAIR implementation.

This submodule includes two activities, which are detailed as follows:

- **Discussion activity** divided in two parts:
  - *before the lecture*: climate modelling research is briefly introduced, and then participants are engaged in a discussion concerning the main challenges in applying FAIR principles in such a domain;
  - *after the lecture*: considering the concepts about large climate experiments (i.e., CMIP), climate data services and climate data infrastructures (i.e., ESGF) presented during the lecture, the discussion concerns how FAIR principles can be applied and address challenges in climate research communities. For instance, a NetCDF file from the ESGF infrastructure could be taken as an example and related FAIR practices and their implementation could be discussed.
- **Practical activity:** by using datasets from the ESGF infrastructure, learners analyse different datasets in small groups of 3 or 4, and evaluate their FAIRness using the different elements and tools of FAIR.

The material about this module is available at the following link: https://zenodo.org/records/12805958/files/OSandRDMforClimateChangeCommunities_Module2.zip

Element	Description
Title	FAIR Principles in Practice
Abstract / Description	A presentation about the ENES Community, relevant experiments like CMIP, the underlying infrastructure ESGF to manage the outputs of climate models simulations and the role of FAIR principles in such a federated data management ecosystem.
Author(s)	Fabrizio Antonio
Primary Language	English
Keyword(s)	ENES, Climate data, NetCDF, ESGF, CMIP, FAIR principles.
License	CC BY 4.0 license https://creativecommons.org/licenses/by/4.0/
Version Date	13/08/2024
URL to Resource	https://zenodo.org/records/12805958/files/OSandRDMforClimateChangeCommunities_Module2.zip
Target Group (Audience)	Researchers
Learning Resource Type	Presentation
Learning Objectives	- Recognise the challenges in applying FAIR principles within climate modeling research - Illustrate the key elements of FAIR principles in a specific climate infrastructure - Recognising community standards available to implement FAIR in climate research (like NetCDF) - Evaluate FAIRness of datasets in climate modeling research
Expertise (Skill) Level	Beginner

Table 2. Module 2: FAIR principles in practice

2.2.3 Module 3: Planning for FAIR climate data

This module is organised into two submodules:

- Introduction to Research Data Management
- Creating a Research Data Management strategy and Data Management Plans

2.2.3.1 Introduction to Research Data Management

The learning objective the first submodule focuses on is:

- Recognise the relevance of RDM to enable FAIR and OS

In terms of content, the submodule includes (i) a *Recap on how OS, FAIR and RDM connect, RDM as a basis for FAIR and making your data open (but going beyond as well, like data organisation), data lifecycle*, (ii) *What is research data (in climate modeling)*, and (iii) *Knowledge about the research and the data life cycle.*

In terms of activities, learners will be engaged in gradually implementing the different elements of RDM, and identifying/using tools required at each stage.

2.2.3.2 Creating a Research Data Management strategy and Data Management Plans

The learning objectives the second submodule focuses on are:

- Identify different types of research data in climate modeling research
- Build an RDM strategy to manage research data in a climate modeling research project
- Test a relevant tool to create a DMP

In terms of content, the submodule provides (i) Knowledge about DMP and FAIR principles to understand/develop data management, (ii) Guidance on research planning, concerning all aspects of a research project preparation; that includes: identification of the data that will be collected or used to answer the research question; plan for data management throughout the lifecycle; limitations (what you can do with the data, commercial restrictions, etc.);applying FAIR principles for data management; role of metadata standards & vocabularies; how to make OS and FAIR compliant choices to preserve data; OS/FAIR best practices and tools

for data publication, citation, dissemination & sharing; infrastructure for the management of community-based experiments; components and services, data formats, metadata standards, conventions; search & discovery; processing, analysis, preservation; sharing and exploitation.

A wide set of activities around these concepts are also part of this submodule. In particular, going through 4 activities, learners will be asked to:

Activity 1 (in small groups): describe the kind of data they work with in their research (data types, descriptive name and adding their descriptions, and also the formats, how it will be collected, whether it will be new or reused (from where);
Activity 2: (linked to Activity 1) provide information about data documentation, metadata; data organisation (folder structure or naming conventions, version control, any measures for data organisation); data quality measures;
Activity 3: share considerations about ethical and legal issues around data, reuse of data, data policies and terms of use of the RIs they are reusing data and software from;
Activity 4: (from Activity 1 and Activity 2) discuss where/in what repository they will publish this data; check whether the repository will give their dataset a PID; kind of license and conditions of access and reuse; possible restrictions to data publication, or an embargo and the reasons about it.

The material about this module is available at the following link: https://zenodo.org/records/12805958/files/OSandRDMforClimateChangeCommunities_Module3.zip

Element	Description
Title	Planning for FAIR climate data
Abstract / Description	A presentation about research data (i.e., in climate modelling) and data life cycle. An in-depth view about research data management and Data Management Plan.
Author(s)	Sandro Fiore
Primary Language	English
Keyword(s)	ENES, Climate data, Research Data Management, Data Lifecycle, Data Management Plan, FAIR principles.
License	CC BY 4.0 license https://creativecommons.org/licenses/by/4.0/
Version Date	13/08/2024
URL to Resource	https://zenodo.org/records/12805958/files/OSandRDMforClimateChangeCommunities_Module3.zip
Target Group (Audience)	Researchers
Learning Resource Type	Presentation
Learning Objectives	- Recognise the relevance of RDM to enable FAIR and OS - Identify different types of research data in climate modeling research - Build an RDM strategy to manage research data in a climate modeling research project - Test a relevant tool to create a DMP
Expertise (Skill) Level	Beginner

Table 3. Module 3: Planning for FAIR climate data

3. Pilots

The pilot event was led by the University of Trento and held in February 2024; it aimed to equip participants with essential skills and knowledge in OS and RDM tailored to the challenges of Climate Change research. The pilot comprised a module spanning two half-day online sessions. It was mainly targeting early career researchers and focused on essential aspects of OS and RDM. The pilot was attended by researchers affiliated to CMCC, the University of Trento and TU-Delft. On the first day, the trainers delivered an introductory session on OS and FAIR principles, which was followed by a theoretical webinar and a hands-on session. The second day featured a three-hour webinar, incorporating a Q&A session and practical activities. Key topics covered include the data lifecycle, RDM practices, OS principles, FAIR principles, preparation of data management plans, and significant infrastructural initiatives in Climate Science.

Key objectives of the pilot were (i) *Fostering the access and use of the IS-ENES data and services* and (ii) *Fostering OS and RDM in Climate Science*.

A total number of 29 people registered for the event; 16 participants attended on day 1, and 13 participants attended on day 2. Dissemination of the event was performed on the [Skills4EOSC website][1] and [LinkedIn of instructors][2].

Feedback was gathered from participants at the end of the event to understand what could be improved for the official training workshop. The pilot was overall a successful event.

As a follow-up activity, a preliminary discussion about the agenda of the training event in Fall 2024 took place; a draft agenda of the event is expected to be developed in September.

---

[1] https://www.skills4eosc.eu/news/successful-delivery-of-pilot-training-on-open-science-and-research-data-management-in-climate-change-2

[2] https://www.linkedin.com/posts/sandro-fiore-ph-d-080ba38_openscience-fair-climatechange-activity-7167169933968097280-s6qs/

4. Discussion and next steps

The task activities during the second year related to (i) the preparation and the successful delivery of the pilot event held in Feb 2024, and (ii) the updating of the training materials and the supplementary documentation that describes in a comprehensive way the main rationale, content and structure of the learning path for climate change researchers.

The delivered learning path is meant to serve as an initial framework that can be further extended and re-used across different areas of Climate Science. Along the same line, slides and exercises that complement this report serve as a concrete example for trainers, with respect to the preparation of new material.

These two aspects jointly with the dissemination within the climate change research community, the organisation of training events and the participation in the activities of Competence Centres , are critical to promote the adoption of the proposed learning framework and the preparation of new material as well as to support new trainers.

Regarding the foreseen activities, these will mainly focus on (i) the organisation of the official training event in Fall 2024, (ii) the review, extension and finalisation of the training material, (iii) the consolidation of the learning path, (iv) the dissemination within the climate community and (v) and the collaboration with the Competence Centre network.

References

No	Description/Link
R1	Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data 3*, 160018 (2016). https://doi.org/10.1038/sdata.2016.18
R2	The European Network for Earth System Modelling https://portal.enes.org/
R3	ESiWACE Center of Excellence on Climate and Weather https://www.esiwace.eu/
R4	FAIR by Design Methodology: https://fair-by-design-methodology.github.io/microlearning/latest/