Supporting eosc

Skills 4 eosc

D5.6 Recommendations for the set-up of Open Science and Research Data Management thematic trainings

Lead Partner:	OPERAS
Version:	1.0
Status:	Final
Dissemination Level:	Public
Document Link:	https://doi.org/10.5281/zenodo.16992361
DOI:	10.5281/zenodo.16992361

Deliverable Abstract

D5.6 “Recommendations for the Set-up of Open Science and Research Data Management Thematic Trainings" outlines a strategic framework for integrating Open Science (OS) and FAIR Research Data Management (RDM) practices into research workflows. Developed under the Skills4EOSC project, the report provides recommendations for designing effective, tailored training programmes. It is intended for Competence Centres, six thematic communities and external training providers. The recommendations stem from an iterative co-creation process, including workshops, pilot training activities, and train-the-trainers sessions and aim to foster collaboration between researchers, infrastructures, and experts. The report also identifies challenges and proposes solutions for the long-term sustainability of OS/RDM training within the European research ecosystem.

Co-funded by the European Union

co-funded by UK Research and Innovation

Skills4EOSC has received funding from the European Union’s Horizon Europe research and innovation Programme under Grant Agreement No. 101058527 and from UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant number 10040140]

COPYRIGHT NOTICE

This work by Parties of the Skills4EOSC Project is licensed under a Creative Commons Attribution 4.0 International License. The Skills4EOSC project is co-funded by the European Union Horizon Europe programme under grant n° 101058527 and by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee under grant n° 10040140.

DELIVERY SLIP

	Name	Partner/Activity	Date
Moderated by:	Fotis Mystakopoulos Elena Sokolova	OPERAS	
Contributor	Kasper Drazewski	ULeuven	7/2025
Contributor	Sandro Luigi Fiore	UTrento	7/2025
Contributors	Isolde Gotwald Megumi Kiesel Heimo Reiner Elena Ritschard	NHMW	7/2025
Contributors	Mario Locati Federica Tanlongo	EPOS ERIC	7/2025
Contributors	Karla Avanco Arnaud Gingold	AMU	7/2025
Contributor	Irakleitos Souyioultzoglou	OPERAS	7/2025
Reviewed by:	Bianca Gualandi	University of Bologna	8/2025
Reviewed by:	Clara Boavida	Iscte-Instituto Universitário de Lisboa	8/2025

Approved by:	Sara di Giorgio, Claudio Prandoni	GARR	29/08/ 2025
Acknowledgment	This is a note of appreciation for all participants in the Milano Workshop that was held in February 2025. Each participant provided input to the overall discussion of what recommendations are important and they significantly shaped the discussions and the overall direction of this work.		

DOCUMENT LOG

Issue	Date	Comment	Author
v0.1	20/05/2025	Early Draft, Open for Comments and Input	Fotis Mystakopoulos, Elena Sokolova
v0.2	27/06/2025	First deadline for partners to provide input	WP5 Partners
v0.3	31/07/2025	Document enriched with input from all partners.	WP5 Partners
v0.4	05/08/2025	Document sent to reviewers with a deadline for 15/8/2025	Peer Reviewers provided comments (Clara Boavida, Bianca Gualandi)
v0.5	25/08/2025	Comments implemented into deliverable	Irakleitos Souyioultzoglou, Fotis Mystakopoulos, Elena Sokolova
v0.6	26/08/2025	Final draft	WP5 partners
V1.0	29/08/2025	Approved version	Sara Di Giorgio, Claudio Prandoni

TERMINOLOGY

Terminology	Definition
ABCD	Access to Biological Collection Data
AI	Artificial intelligence
CC	Competence Centres
CESSDA	Consortium of European Social Science Data Archives
CF	Core Facilities
CLARIN	Common Language Resources and Technology Infrastructure
DARIAH	Digital Research Infrastructure for the Arts and Humanities
DC	Dublin Core
DiSSCo	Distributed System of Scientific Collections
DMPs	Data Management Plans
DOI	Digital Object Identifier
ECR	Early Career Researchers
ECTS	European Credit Transfer and Accumulation System
EDM	Europeana Data Model
EIDM	Evidence-Informed Decision-Making
ELSI	Ethical, Legal, and Societal Issues
EMSO	European Multidisciplinary Seafloor and Water Column Observatory
ENES	European Network for Earth System Modelling
EPOS	European Plate Observing System
EQF	European Qualifications Framework
ESGF	Earth System Grid Federation
ESMRI	European School for Management of Research Infrastructures
FAIR	Findable, Accessible, Interoperable, Reusable
FBD	FAIR-by-Design methodology
ICOM	International Council of Museums
INGV	Instituto Nazionale di Geofisica e Vulcanologia
IPR	Intellectual Property Rights

LAM	Libraries, Archives, and Museums
LIDO	Lightweight Information Describing Objects
LP	Learning Path
MVS	Minimum Viable Skills (or Skillsets)
NFDI	National Research Data Infrastructure (Germany)
NGO	Non-Governmental Organisation
NS	National-Socialist
OER	Open Educational Resources
OS	Open Science
PIDs	Persistent Identifiers
RDA	Research Data Alliance
RDM	Research Data Management
RI	Research Infrastructures
SSH	Social Sciences and Humanities
TtT	Train-the-Trainer
WP	Work Package

Table of Contents

Executive summary

D5.6 "Recommendations for the Set-up of Open Science and Research Data Management Thematic Trainings" presents a strategic roadmap for embedding Open Science (OS) and FAIR Research Data Management (RDM) practices into research workflows. Developed under the Skills4EOSC project, this report offers guidance for designing effective, thematically tailored training programs that support collaboration between researchers, infrastructures, and experts across thematic communities.

The recommendations are based on an iterative co-creation process, including two dedicated workshops (May 2023 and February 2025) and lessons learned from pilot training activities.

Who is the report for?

The recommendations are directed at:

- Competence Centres: National and institutional hubs offering OS/RDM support.
- Six Thematic Communities: professionals in Research Infrastructures (RI), Social Sciences and Humanities (SSH), Solid Earth Science, Climate Science, and Open Scientific Collections. Recommendations relevant to experts in Ethical, Legal, and Social Issues (ELSI) and data governance are also included.
- External Training Providers: Research Infrastructures, Non Governmental Organisations (NGOs), and professional networks that can adopt and scale these materials.

The report presents two sets of recommendations, summarised as follows:

Cross-Thematic (Horizontal) Recommendations:

- Learning Paths should be modular and progressive (2–4 modules), blending theory with practice and tailored to disciplinary needs. Flexibility in delivery formats (online/on-site/hybrid) is encouraged, though challenges remain in balancing depth with accessibility.
- Minimum Viable Skillsets (MVS) are a valuable reference, but require local adaptation based on technical, institutional, and cultural contexts.
- Trainer profiles should embrace an inclusive Train-the-Trainer (TtT) model, promoting sustainability by empowering participants to become trainers themselves.

- ELSI Considerations must be addressed across disciplines, though legal aspects often require country-specific adaptations.
- FAIR-by-Design is a recommended method for training design, with implementation tailored to each community's expertise and infrastructure.
- Approaches to accreditation vary: while Solid Earth Science explores credit systems, other areas focus on content certification, highlighting the need for a modular, community-sensitive framework that integrates tools such as digital badges and European Digital Credentials for Learning. This ensures recognition and portability across disciplines and borders.
- Sustainability is essential: the report highlights the lack of formal incentives for trainers and dependence on project-based funding. Solutions include institutional partnerships, formal trainer roles, and certified TtT programs.

Thematic Recommendations:

- **Research Infrastructure Professionals:** Training should be modular and paced appropriately, informed by pre-course assessments, and offered in hybrid formats. Resources must remain openly licensed and aligned with the Minimum Viable Skillset (MVS) and FAIR-by-Design to enable reuse and scaling.
- **Social Sciences and Humanities (SSH)**: Recommendations focus on framing SSH data through methodologies rather than disciplines, leveraging existing data classifications (e.g., DANS, ICPSR), and clarifying the role of data, software, and analysis within digital research workflows.
- **Solid Earth Sciences:** Emphasis is placed on flexible delivery formats, modular training, and domain-specific applications. Practical skills (e.g., data handling, automation via Jupyter or Virtual Research Environments) should complement FAIR/OS content. Integration into curricula, adoption of OERs, and building trainer networks are key to long-term sustainability.
- **Climate Sciences:** Training should target FAIRification of data, workflows, and provenance, using real use cases and EOSC services. Dedicated modules on the practical use of Persistent Identifiers (PIDs) are recommended, alongside strong community-based practices.
- **Open Scientific Collections (Libraries, Archives, Museums)**: Guidance addresses digitisation (metadata standards, authenticity certification, stakeholder engagement), interoperability (globally unique PIDs, FAIR-compliant linking systems), and provenance (sensitive contexts, institutional policies, CARE/FAIR integration).
- **Ethical, Legal, and Societal Issues (ELSI) & Governance**: Training must combine awareness-raising with tailored content for policymakers and civil

servants, avoid curriculum overload, and ensure engagement through interactive and values-based discussions, both in-person and online.

The report calls for long-term integration of these recommendations into the European Open Science ecosystem. Priorities include multilingual training delivery, developing micro-credentials, adapting MVS profiles, incentivising trainers, supporting cross-centre collaboration, and staying informed on emerging technologies like Artificial Intelligence. The ultimate goal is to establish a durable, inclusive, and context-aware training framework that supports FAIR and Open Science transformation at all levels of the research system.

1.Introduction

The key objective of this deliverable is to advance the integration of OS and FAIR RDM practices[1] into research workflows by providing thematically grounded recommendations, developing tailored training courses for diverse research communities, and fostering collaboration between OS experts and researchers.

The report is linked to Work Package 5/Task 5.7 of Skills4EOSC. This WP brings together six thematic communities to ensure a broad and inclusive approach to embedding OS and FAIR RDM practices across disciplines through training activities.[2] The thematic areas covered include Research Infrastructures, Social Sciences and Humanities, Solid Earth Science, Climate Science, Open Scientific Collections, Ethical, Legal, and Societal Issues and Data Governance for Evidence-Informed Decision-Making (EIDM). By engaging these diverse domains, the project aims to develop tailored guidance and support the adoption of responsible and transparent research practices through the OS paradigm.

The recommendations are the result of a long-term, iterative process grounded in real-world testing and refinement. To develop and refine applicable recommendations, WP partners gathered twice (cf. Section 2 of the report). An online workshop took place on 3 May 2023, with an aim to discuss the approach for creating targeted training activities. Following the workshop, the core work of each Task involved designing pilot versions of their training modules, gathering feedback, incorporating improvements, and ultimately delivering a final Train-the-Trainer programme.

---

[1] Wilkinson, M. 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

[2] The following courses have been developed, and are also available on the Skills4EOSC learning platform (https://learning.skills4eosc.eu)

- Cocco et al. “Learning Path and Training Materials on "Open Science and Research Data Management in Solid Earth Sciences". Zenodo. https://doi.org/10.5281/zenodo.13684718
- D'Anca et al. “OS and RDM learning paths for Climate Change communities”. Zenodo. https://doi.org/10.5281/zenodo.14797326
- Lavitrano, M. “FAIR and RDM training modules for RI professionals”. Zenodo. https://doi.org/10.5281/zenodo.15731892
- Linés et al. “FAIR and RDM training modules for open collections”. Zenodo. https://doi.org/10.5281/zenodo.14537731
- Mystakopoulos et al. “OS and RDM learning paths for Social Sciences and Humanities communities”. Zenodo. https://doi.org/10.5281/zenodo.14797210

Each Task had sufficient time to engage deeply with the respective thematic community, assess specific needs, identify challenges, and explore potential solutions. This process was further enriched by a second workshop held in February 2025, where the final structure and content of the recommendations were discussed and validated. It was agreed that while some challenges and insights cut across all thematic areas – emerging as horizontal recommendations – each community would also articulate its own tailored, domain-specific recommendations.

This report therefore represents a consolidated synthesis of the work across all Tasks and captures the key lessons learned throughout this process of co-development and thematic reflection.

1.1 Links to Competence Centres

The recommendations are targeted for Competence Centres (CCs) that are becoming part of the CC network[3] and/or are involved in training activities. CCs were invited in the second workshop specifically to provide contextual information about their position in their national landscapes. As national or institutional hubs for OS and RDM expertise, they serve as intermediaries between policy, infrastructure, and research communities. Participating CCs from France (Recherche Data Gouv), North Macedonia (University of Saint Cyril and Methodius), Germany (Karlsruhe Institute of Technology), Greece (GRNET), Poland (Gdańsk University of Technology Library), Finland (CSC – IT Center for Science), and Sweden (Swedish National Data Service), and others represented in the pilot events and/or the Train-the-Trainer (TtT) sessions demonstrated how training efforts can be aligned with policy priorities, disciplinary structures, and research infrastructures. Their involvement ensured that training recommendations were grounded in existing support ecosystems and could potentially be scaled through national networks or federated models.

The CCs were also instrumental in identifying local challenges, such as trainer availability, policy gaps, or technical barriers, and facilitating peer learning across countries. In this regard, they are strategic partners and multipliers for the thematic training recommendations.

[3] Corleto, et. al. "Report on CCs and user support networks and recommendations for networks evolution". Zenodo. https://doi.org/10.5281/zenodo.15262091

1.2 Target Audiences

The development of thematic recommendations has revealed the importance of carefully identifying and engaging diverse, yet clearly defined, target audiences. Each Task focused on a distinct community or professional domain, underscoring the necessity of tailoring training approaches and materials to their unique operational environments. Below, a typology of those target audiences is provided.

1.2.1 Thematic Communities

- Research Infrastructure Professionals (T5.1): managers, operators, and professionals working in RIs and Core Facilities (CFs).
- Social Sciences and Humanities (T5.2): researchers, support staff, and data managers across a variety of SSH disciplines.
- Solid Earth Science (T5.3): sub-communities within the broader geoscientific field (e.g., Seismology, Volcanology, Geophysics). Training was developed for early career researchers (ECRs), data providers, and those involved in RIs like the European Plate Observing System (EPOS) and the European Multidisciplinary Seafloor and Water Column Observatory (EMSO), addressing their varied levels of digital and data literacy.[4]
- Climate Sciences (T5.4): researchers, data stewards, and infrastructure managers. The community, particularly within the European Network for Earth System Modelling (ENES) and the Earth System Grid Federation (ESGF),[5] handles large-scale data but faces systemic gaps in accessible, updated training. Training efforts thus focused on basic and intermediate levels, aiming to reach academic institutions and climate organisations more broadly.
- Open Scientific Collections (T5.5): curators, scientists, and RDM leaders within Libraries, Archives, and Museums (LAM institutions). These professionals operate at the intersection of heritage and research, managing digitised physical collections. The recommendations focused on equipping them with the skills to handle sensitive provenance data, apply licensing, and use persistent identifiers.
- ELSI and Data Governance (T5.6): policy makers, civil servants, and decision-makers involved in digital regulation and evidence-informed decision-

[4] EPOS ERIC: https://www.epos-eu.org ; EMSO: https://emso.eu/

[5] ENES: https://portal.enes.org/index.html ESGF: https://esgf.github.io/index.html

making. These actors occupy influential roles in shaping research policy and legislation, and the training aimed to strengthen their ethical reasoning and resilience to power asymmetries in data governance.

1.2.2 External Open Science Training Providers

Beyond the core project network and national CCs, there exists a growing ecosystem of external OS training providers, including professional networks, Non-Governmental Organisations (NGOs), and RIs. These organisations offer Open Educational Resources (OER), thematic workshops, certification schemes, and discipline-specific training. They serve as both collaborators and amplifiers, capable of integrating the thematic materials developed in T5.1–T5.6 into their own programming or co-hosting events that extend the reach of the recommendations. Therefore, the recommendations and needs identified here should serve the wider community as well.

2. Methodology

WP5 developed a set of resources and learning paths tailored to the specific needs and practices of each disciplinary area. These were initially tested during pilot training events, where a first iteration of the materials was presented to small groups of participants. The pilots were an important step of the process to evaluate the relevance of the content, identify challenges, and collect feedback on what worked well and what required improvement.

The feedback gathered during the pilot phase directly informed the adjustments made in the final version of the training resources and helped in identifying the conceptual framework for the recommendations.

The pilots were followed by Train-the-Trainers sessions with representatives of the thematic communities, which supported the creation of a community of trainers equipped to share knowledge and to strengthen capacity within their institutions and networks.

Two workshops were organised to support the designing of the training activities, and the formulation of actionable recommendations. The first workshop took place before the start of the pilot training sessions. Its main goal was to identify key challenges, discuss the objectives of the learning paths under preparation, and define the expected outputs, targeted communities, and issues to be addressed. It also helped align the work and planning across the different Tasks. The second workshop was held after completion of the pilots, with a focus on reviewing the feedback received, and using it to identify and refine the recommendations.

2.1 Workshop on the Designing of Thematic Trainings

The initial (online) workshop was organised on 3 May 2023, with an aim to explore the specificities of OS and RDM training within the respective disciplinary fields, and discuss strategies for addressing the diverse needs of stakeholders – including research communities and research infrastructures operating at both national and international levels.

The workshop comprised four main sessions, and a concluding round of collective discussion. Task leaders and community representatives outlined the current state of training in their thematic areas and discussed effective approaches for introducing training as a core practice.

The introductory session provided an overview of the disciplinary areas involved, with a focus on researchers' needs and emerging research workflows. The follow-up discussion underscored the limited adoption of OS and RDM practices across all thematic fields, and the need to develop training resources and services that support all stages of the research lifecycle.

The second session focused on current developments in OS and RDM within each thematic community. It offered an opportunity to exchange strategies and solutions used by research infrastructures to support the development of a sustainable OS ecosystem. Discussions also revealed common challenges, including the financial sustainability of OS and FAIR-oriented initiatives, the heterogeneity of data types, and the fragmentation of service provision.

The third session was dedicated to the planned training activities for each thematic area. Participants identified challenges related to coordination and alignment within each community, noting that existing OS and RDM training only partially meets stakeholder needs. Furthermore, it was emphasised that training design should correspond to the varied roles and diverse career paths that researchers often assume, as well as the emerging involvement of other professionals – such as data stewards and infrastructure staff- in research workflows.

The fourth session focused on the ethical and legal implications of access to scientific data throughout the research lifecycle. A detailed presentation outlined the EU legal framework on data, covering both personal and non-personal data, and addressed commonly cited legal barriers to OS, particularly concerning IPR and copyright.

As outlined during the final session, the main conclusions of the workshop helped shape the conceptual framework for the WP5 training activities, as follows:

- A balanced approach that addresses discipline-specific needs while also ensuring the inclusiveness of trainings and learning paths.
- Harmonisation across various initiatives and thematic fields emerged as a common foundation for effective coordination and progress.
- Accordingly, a priority for WP5 would be to ensure the availability of skilled trainers capable of designing and delivering tailor-made trainings.

These guiding principles were revisited and evaluated during the second WP workshop, which took place following the delivery of the training pilots.

2.2 Workshop on Developing Recommendations

After the completion of the pilot activities of all Tasks, T5.7 focused on coordinating partners in drafting thematic recommendations, which were discussed and validated in a workshop held on 26 and 27 February 2025, at the University of Milan. **Day 1** focused on reviewing the draft thematic recommendations, with presentations from Task representatives and discussions around key challenges. These exchanges led to a first round of revisions to the recommendations by the end of the day. **Day 2** shifted to the perspective of Competence Centres, examining the broader context in which the recommendations would be applied. Together, participants identified the directions the recommendations should take, distinguishing between two levels: cross-thematic recommendations and more detailed, community-specific recommendations.

2.2.1 Pre-workshop template

As a preparatory step, a structured template was created and shared with partners, with an aim to support each Task in drafting preliminary recommendations based on experiences and insights gathered during the pilot activities. These early drafts were intended to serve as a foundation for further discussion.

The template provided a clear structure to ensure consistency, while allowing flexibility to accommodate specific needs and contexts. It invited contributors to reflect on their target audiences, training content, stakeholder involvement, and key challenges encountered. This approach aimed to ground the recommendations in practical experience while aligning them with broader OS and RDM goals. The table below summarises the core structure of the template and the elements considered:

Section	**Purpose**	**Key Points**
Area for Recommendation	Define the scope and thematic context of the recommendations.	Target group profile, national context, specific needs and challenges.
Learning Path	Describe the structure and rationale behind the training program.	Number of modules, progression of skills, timing, thematic adjustments.

Learning Objectives	Highlight the unique learning outcomes tailored to the thematic area.	Relevance to OS/RDM, alignment with specific research disciplines.
Stakeholders	Identify all relevant actors involved in co-developing and delivering the recommendations.	Institutions, networks, infrastructures, and their contributions.
Training Content	Detail the curriculum, delivery formats, and methodological approaches.	FAIR-by-Design compliance, practical vs. theoretical focus, use of tools and case studies.
Key Challenges Identified	Summarise the obstacles and feedback that influenced or emerged from the pilot activities.	Barriers to OS/RDM uptake, feedback from participants, post-delivery insights.
Any Other Issues	Capture additional insights including certification, communication, and dissemination practices.	Optional topics based on pilot experience.

Table 1 : Overview of the Thematic Recommendations Template

2.2.2 Workshop activities

The workshop brought together a diverse range of institutions and CCs, illustrating the broad engagement across the Skills4EOSC project. Participants represented universities, research infrastructures, data service providers, and organisations focused on OS policy and training. Academic institutions were represented by the following consortium partners: University of Bologna, University of Trento, University of Turin, KU Leuven, the University of Edinburgh, and alongside research organisations such as the Instituto Nazionale di Geofisica e Vulcanologia (INGV) and the Natural History Museum of Vienna.

A strong presence of national CCs added valuable insights to the discussions. Competence Centres from France (Recherche Data Gouv), North Macedonia (University of Saint Cyril and Methodius), Germany (Karlsruhe Institute of Technology), Greece (GRNET), Poland (Gdańsk University of Technology Library), Finland (CSC – IT Center for Science), and Sweden (Swedish National

Data Service)[6] shared their approaches to implementing OS training and fostering community engagement at national and institutional levels. This diversity of perspectives ensured that the thematic recommendations would be grounded in both practical experience and a variety of policy and infrastructure contexts, strengthening their relevance and applicability across institutional contexts and domains.

Day 1 – Draft Recommendations, Thematic Communities

The first day was dedicated to presenting the initial drafts prepared by each thematic task group (see Annex 2). These were discussed in plenary sessions with feedback from all participants, followed by a mapping exercise to identify common themes and challenges. Breakout discussions allowed for deeper exploration of selected topics, and the day concluded with group work to revise and enrich the drafts based on the input received. Below we summarise the key reflections from each thematic community, addressing both the broader challenges of OS and the specific challenges of OS training. These reflections were developed prior to the workshop (using the template), with each community highlighting issues from its own perspective. In some cases, the insights reflect overarching theoretical challenges of OS; in others, they focus more directly on practical considerations with regards to training. The two dimensions are closely interlinked: while training requires a dedicated approach due to its unique challenges, understanding OS more broadly is essential for shaping effective training design.

T5.1 – Research Infrastructures

Due to scheduling conflicts T5.1 was not directly represented, creating a gap in terms of task-specific input. However, the discussion benefited from the strong representation of RI professionals (OPERAS, EPOS, GRNET), as well as from the contributions on behalf of CCs operating in the RI domain.

[6] Recherche Data Gouv: https://recherche.data.gouv.fr/en
University of Saint Cyril and Methodius https://www.ucm.sk/en/
Karlsruhe Institute of Technology: https://www.kit.edu/english/
GRNET: https://grnet.gr/en/
Gdańsk University of Technology: https://pg.edu.pl/en/gdansk-tech-library
CSC – IT Center for Science: https://csc.fi/en/
Swedish National Data Service: https://snd.se/en

As a result, the horizontal recommendations discussed during the workshop reflect the perspectives of RI professionals across the thematic areas.

T5.2 – Social Sciences and Humanities

The SSH thematic recommendations were shaped by the pilot implementation in Spring 2024 and focused on the need to adapt OS and RDM training to the diversity and complexity of SSH disciplines. MVS profiles were a useful reference point but required adaptation to reflect varying roles and disciplinary contexts. Defining "data" in SSH emerged as a central challenge, especially for research involving monographs, archives, or non-traditional outputs. The learning path and objectives emphasised flexibility, alignment with varying familiarity levels, and the inclusion of domain-specific materials. Stakeholder engagement was seen as critical, with infrastructures such as the Consortium of European Social Science Data Archives (CESSDA), the Common Language Resources and Technology Infrastructure (CLARIN) and the Digital Research Infrastructure for the Arts and Humanities (DARIAH) having a central role in supporting the creation of a community of practice around OS and RDM.[7] Training content needed to balance theoretical and practical elements, with attention to licensing, multilingualism, and national policy contexts. Identified challenges included limited recognition of OS practices, disciplinary silos, and the need for improved integration into academic curricula.

T5.3: Solid Earth Sciences

The thematic recommendations addressed the heterogeneous nature of the community, which encompasses multiple subdomains with distinct data standards and practices. While MVS profiles were a helpful checklist, they could benefit from thematic specificity. Training was designed modularly to accommodate time constraints and varying skill levels, with core learning outcomes focused on understanding OS/FAIR principles, applying them in context (e.g. EPOS infrastructure), and using data portals effectively. Content design needed simplification of methodologies like FAIR-by-Design (FBD)[8] to match the community's technical familiarity. Stakeholders included Early Career Researchers, data providers, and RIs. Key challenges included limited trainer capacity, varied technical literacy, fragmentation in standards

[7] CESSDA: https://www.cessda.eu/ ; CLARIN ERIC: https://www.clarin.eu/ ; DARIAH: https://www.dariah.eu/

[8] Filiposka et al. (2023). D2.2 Methodology for FAIR-by-Design Training Materials (1.4). Zenodo. https://doi.org/10.5281/zenodo.8305540

adoption, and the need to embed OS/RDM into curricula and funding frameworks.

T5.4 Climate Sciences

The thematic recommendations for Climate Sciences addressed the challenges of delivering consistent and accessible OS and RDM training within a data-intensive but fragmented research community. While the MVS profiles were used as general reference, they lacked sufficient specificity for the needs of climate researchers, data stewards, and infrastructure managers. The community, particularly within ENES, faces a lack of reusable and systematically updated training materials, with most initiatives tied to temporary project funding and implemented at single institutions. To address this, a structured training framework was proposed, consisting of two core modules complemented by optional thematic units, all featuring integrated hands-on components.

Three clear learning objectives were defined: a general introduction to OS and FAIR; thematic application of FAIR principles in climate modelling; and planning for FAIR/RDM within projects. Training delivery emphasised flexibility (online and on-site), practical engagement, and modular design. Key infrastructures such as ENES and ESGF were central stakeholders, with universities and climate research organisations identified as important targets. Although certification was not prioritised, communication and long-term dissemination strategies were seen as essential for embedding a culture of openness. The main challenges included community heterogeneity, the absence of a central training repository, and a lack of sustainable, scalable training infrastructure.

T5.5 Open Scientific Collections

The thematic recommendations for Open Scientific Collections focused on the unique needs of curators and research data managers working with physical collections undergoing digital transformation within Libraries, Archives, and Museums (LAM institutions). A tailored MVS profile was developed for the role of digital collection curators, refining the broader data steward profile to address the specific challenges faced in digitisation, metadata standardisation, and ethical handling of provenance information.[9]

[9] The complete description of the MVS "Digital Collections Curator" can be found in Appendix 1 of: Linés, C., et al. (2024). D5.5 FAIR and RDM training modules for open collections. Zenodo. https://doi.org/10.5281/zenodo.14537731

The training path comprised four modules, combining theoretical lessons with case studies. Learning objectives were closely aligned with working with collections as data, emphasising OS, FAIR, and CARE (Collective Benefit, Authority to Control, Responsibility, Ethics) principles. Key topics included the application of metadata standards (e.g. ABCD, DarwinCore, LIDO),[10] the use of Persistent Identifiers (PIDs), licensing for reusability, and the management of sensitive provenance data. The FAIR-by-Design methodology was particularly useful in supporting reproducibility by enabling the citation of physical objects and their digital surrogates.

Stakeholders included LAM institutions, academic and research organisations, funders, professional networks (e.g. International Council of Museums -ICOM; Research Data Alliance-RDA), and RIs such as the Distributed System of Scientific Collections (DiSSCo).[11] Major challenges identified included a lack of authenticity certification for digitised objects, insufficient resources for large-scale digitisation, gaps in metadata for historical materials, and the need for clear institutional policies addressing colonial or ethically sensitive content. The pilot training confirmed a high demand for practical, technical guidance and peer exchange. Lessons learned informed recommendations to establish a community network of data curators, promote reuse of OERs, and adapt training to fast-evolving technologies like 3D digitisation and Artificial Intelligence.

T5.6 ELSI and Data Governance

The ELSI-related recommendations addressed the need to support decision-makers in digital regulation, particularly policy makers and civil servants, by equipping them with foundational OS and RDM knowledge. These individuals operate at the intersection of science and policy, where asymmetries of knowledge and influence frequently favour industry actors. To address this, the training sought to foster critical engagement with legal and ethical frameworks and equip participants with practical tools for evidence-informed decision-making.

---

[10] ABCD: http://www.tdwg.org/standards/115
DarwinCore: https://dwc.tdwg.org/
LIDO: https://cidoc.mini.icom.museum/working-groups/lido/lido-overview/

[11] ICOM: https://icom.museum/en/
RDA: https://www.rd-alliance.org/
DiSSCo: https://www.dissco.eu/

The learning path followed a general OS/FAIR introduction and then introduced three targeted modules covering legal and ethical frameworks, OS under EU regulatory frameworks, and data governance strategies aligned with FAIR. The FAIR-by-Design methodology was found particularly useful in easing syllabus preparation and supporting access to training materials. Key challenges included the breadth and complexity of ELSI topics, time constraints, limited space for discussion in the pilot, and the difficulty of keeping non-expert audiences engaged. Stakeholders were identified at both EU and national levels, with training applicable to public institutions, funders, NGOs, and ministries. The feedback highlighted a need for more time and interactivity in future versions of the course.

Day 2 – Competence Centres

On the second day the focus shifted to the adoption and dissemination of the recommendations. CCs from various countries presented their perspectives on OS training, which informed a collective discussion on action planning. Further breakout work supported the drafting of concrete steps for implementation, and the workshop concluded with a collaborative session to outline the final deliverable and plan the next stages of the work.

A key strength of the workshop was the broad participation of CCs from across Europe, each offering insights grounded in distinct institutional settings, national policy environments, and disciplinary needs. The diversity of these contributions significantly enriched the collective process. The presentations delivered by representatives from the CCs reflected a spectrum of models—from government-led national initiatives to university-anchored infrastructures and collaborative networks.

Each CC shared its own organisational structure, training strategy, and operational challenges. Some, such as the French centre, operate within large, nationally funded federated networks with a dual focus on training and infrastructure provision. Others, like the centre in North Macedonia, are more compact in scale but strategically embedded in major national universities, reaching large proportions of the country's research community through targeted thematic training and hands-on support. The German and Polish centres highlighted the role of university-based infrastructures and cross-institutional alliances, including their engagement in consortia such as the National Research Data Infrastructure (NFDI) and EOSC-related networks. Finland and Sweden presented national-level centres that rely heavily on collaboration and decentralised networks to scale training provision and

address the needs of diverse disciplines. Notably, many centres combine practical support services (e.g. repository hosting, DMP guidance, legal advice) with training programmes, often using thematic case studies or domain-specific examples.

Despite the variation in structure and scale, several shared challenges emerged across the CCs: engaging thematic communities outside core networks, identifying and retaining qualified trainers, ensuring sustainability beyond project-based funding, and translating OS principles into discipline-sensitive, actionable practices. Strategies to address these issues ranged from master trainer programmes (e.g. Greece, Sweden) and national competence networks (e.g. Poland, France) to participatory workshops and use-case driven training (e.g. Finland, North Macedonia). The importance of aligning national efforts with European initiatives such as EOSC, and strengthening mutual learning among centres, was emphasised throughout the discussions.

Overall, the breadth of experience presented offered a dynamic and grounded understanding of what it means to support OS at scale. Their contributions helped shape the emerging recommendations, offering practical and strategic input to ensure that future actions remain context-sensitive, community-driven, and aligned with evolving European research infrastructures.

2.3 Mapping of OS/RDM Topics Across Disciplines

The table below summarises the OS and RDM topics covered by each Task. The training pilots delivered by T5.1 to T5.5 were intentionally diverse: each thematic community designed its own course structure, selected domain-specific tools, and addressed community needs in a targeted way. Domain-specific ELSI aspects were integrated into the Learning Paths, while T5.6 produced a report that elaborates on issues relevant across all thematic communities, as well as task-specific aspects raised during the drafting of the recommendations.[12]

To present both the disciplinary variations and the underlying commonalities consistently, a landscape exercise was undertaken:

Identification of topics:

[12] Drążewski, K. (2025). ELSI requirements for RI professionals & thematic communities. Zenodo. https://doi.org/10.5281/zenodo.15299344

- Horizontal elements highlighted in Section 3 of this report (e.g. Learning Paths, MVS profiles, ELSI) were tagged first.
- Additional topics that recurred in pilot documentation, such as metadata standards or hands-on tools, were added to reflect actual practice.

Scoring by thematic task:

- Each topic was traced in the pilot reports, workshop minutes and training materials.
- Fully covered: integrated into learning objectives and practised.
- Partially covered: mentioned or addressed briefly.
- Not covered: absent from the task.

The table offers an at-a-glance overview using discipline-specific examples taken from this mapping. It can serve as an initial reference for institutions that plan to reuse or adapt the training materials: green cells point to topics already well developed across the trainings, orange cells to areas that would benefit from further refinement, and blue cells to topics that may require additional development in future training cycles.

Topic Group	Topic (●/○)	T5.1	T5.2	T5.3	T5.4	T5.5	T5.6
Training Architecture	● Learning Paths	✓	✓	✓	✓	✓	✓
	● MVS Profiles	✓	✓	~	✓	✓	✓
	● Trainer Profiles	~	~	~	~	~	~
	○ Delivery Formats	✓	~	✓	✓	✓	~
	● Accreditation	~	×	~	×	×	×
Core OS/RDM Principles	● FAIR-by-Design	✓	✓	✓	✓	✓	✓
	○ Licensing & Reuse	✓	✓	✓	~	✓	✓
	○ Metadata Standards	~	~	✓	✓	✓	~
	○ PIDs	~	~	~	✓	✓	×
	○ DMPs	✓	~	✓	✓	✓	✓
Topic Group	**Topic (●/○)**	**T5.1**	**T5.2**	**T5.3**	**T5.4**	**T5.5**	**T5.6**
Legal / Ethical / Policy	● Legal & Governance	~	~	×	~	✓	✓
	● ELSI Considerations	~	~	×	~	✓	✓
Infrastructure & Tools	○ Use of RIs	✓	✓	✓	✓	✓	~
	○ Hands-on Tools	×	✓	✓	✓	✓	~
Ecosystem & Support	○ Stakeholder Involvement	✓	✓	✓	✓	✓	✓
	● Sustainability	✓	~	~	~	~	~

Visual encoding:

● Horizontal topic from Section 3 (expanded in Chapter 3)
○ Topic derived from task-level implementation
✓ Fully covered ~ Partially covered × Not covered

T5.1: Research Infrastructures
T5.2: Social Sciences and Humanities
T5.3: Solid Earth Science
T5.4: Climate Science
T5.5: Open Scientific Collections
T5.6: Ethical, Legal and Societal Issues

Table 2: Mapping of OS/RDM Topics Across Thematic Tasks

2.4 Scoping limitations

At this stage, it is important to acknowledge certain limitations in the scope of the thematic recommendations. First and foremost, while the recommendations are rooted in specific disciplinary contexts, some horizontal elements may be applicable across various fields. However, the level of granularity in the final recommendations should be interpreted within the boundaries of the thematic and disciplinary areas they address.

A more significant limitation concerns the long-term applicability of these recommendations—whether horizontal or discipline-specific. These outputs were developed within the timeframe of the project and are intended to serve as a foundation rather than a fixed framework. Their continued relevance depends on regular review and adaptation by the institutions that adopt them, based on evolving local and disciplinary needs.

Two recent developments highlight the need for flexibility and responsiveness. First, the COVID-19 pandemic transformed the delivery of training, necessitating rapid shifts to fully online formats. Although many in-person and hybrid options have since returned, this shift underscored the importance of tailoring training approaches to diverse learning environments. Our current recommendations focus on thematic content and do not prescribe modality-specific training strategies.

Second, the emergence of AI has significantly reshaped how research is conducted, disseminated, and reused. The speed and scale of these changes illustrate the challenge of producing future-proof guidance. As such, these

recommendations should not be viewed as static or definitive, but rather as a dynamic resource, intended to evolve alongside technological, institutional, and societal developments.

3. Horizontal Recommendations (Cross-Thematic)

This section presents a set of horizontal recommendations for training, which aim to support capacity-building efforts that are adaptable across disciplines, institutions, and career stages.

3.1 Learning Paths

The concept of a Learning Path (LP) indicates a pedagogical framework for structuring training initiatives in OS and RDM. Across various disciplinary contexts, LPs should serve to organise educational content in a progressive manner, calibrated to participants' prior knowledge, disciplinary requirements, and anticipated skill development trajectories. The following overview synthesises, in the form of actionable recommendations, the approaches adopted in the WP5 thematic areas, with particular attention to modularity, pedagogical sequencing, and responsiveness to learner profiles and institutional constraints.

All LPs adopt a modular structure, typically comprising 2 to 4 core modules that progress from foundational OS/RDM principles to domain-specific applications. This structure allows for flexible delivery, accommodating both short sessions and extended formats, and supports adaptation to diverse audiences.

There is a strong emphasis on audience calibration, with content tailored to varying levels of prior knowledge and disciplinary maturity. Frameworks such as the MVS are used to define entry points and learning outcomes, especially in disciplines where OS concepts are less embedded.

Training designing consistently aims to balance theory and practice, combining conceptual instruction with hands-on components like case studies and exercises. However, several areas note challenges in offering adequate depth within limited timeframes, particularly for technical or regulatory topics.

Time and logistical constraints are a recurring concern. Shorter formats are preferred by established professionals, while longer workshops suit Early Career Researchers and are often embedded in structured programmes.

Blended and hybrid formats are commonly adopted to increase accessibility and resilience.

Differences primarily arise from domain-specific needs. For example, Solid Earth Sciences training aligns with the EPOS infrastructure, Climate Sciences focuses on FAIR data in modelling, and ELSI training engages with complex legal frameworks. Each requires tailored examples, tools, and pacing.

Challenges include:

- Addressing heterogeneous skill levels without diluting content
- Avoiding superficial treatment of complex subjects
- Meeting demand for more technical and implementation-focused guidance
- Ensuring long-term sustainability and integration into curricula

Recommendation

LPs across disciplines should reflect a shared commitment to modular, learner-centred, and practice-oriented training in OS and RDM. Their effectiveness depends on the ability to adapt to specific audiences, integrate domain-relevant content, and balance flexibility with depth. Future efforts should prioritise curriculum integration, deeper technical content, and mechanisms for sustained delivery.

3.2 Minimum Viable Skills/Skillsets

The concept of Minimum Viable Skills profiles functions as a foundational tool for identifying the core competencies needed to engage with OS and RDM. Across thematic areas, MVS are regarded as a useful starting point, particularly for researchers and Early Career Researchers, for introducing key concepts such as FAIR principles and RDM practices.

However, a recurring observation is the generic nature of MVS, which limits their capacity to address the nuanced needs of specific research communities. While general profiles are valuable for outlining shared competencies, meaningful application requires community-level adaptation to reflect disciplinary, technical, and organisational particularities. This was especially evident in areas like SSH and Open Scientific Collections, where profiles were adapted during the training design process, and refined iteratively.

MVS also support needs analysis and learning path design, helping to identify which skills are most relevant to a given audience and guiding training

structure accordingly. In ELSI and Data Governance, profiles such as Policy Maker, Civil Servant, and Honest Broker were found particularly pertinent, with training aimed at helping participants map and develop those competencies.

Recommendation

MVS should be adopted not as static templates, but as evolving frameworks. Their scope is to serve as a conceptual framework that can be further co-developed with the communities they aim to serve.

3.3 Trainer Profiles

During the second workshop, participants discussed the expected trainer profiles for the courses. At their core, the courses are designed to equip Trainers-of-Trainers, making it essential to understand the kind of audience each course is intended for.

It was generally agreed that some familiarity with OS and RDM is necessary, as the courses are not meant to teach the fundamentals but to build confidence in delivering training. However, there was also a clear recommendation to avoid being too prescriptive or exclusive when selecting participants. Being overly restrictive may discourage individuals who are enthusiastic and motivated but lack formal credentials. Flexibility in accepting applications is encouraged.

A more fundamental issue raised was the definition of a *Master Trainer*. This term prompted discussion, particularly among Competence Centres, regarding whether it refers to a single person overseeing programme development or to multiple individuals. The recommendation is that a Master Trainer should be understood as anyone who has completed the necessary training modules and can train others, ideally resulting in multiple Master Trainers within an institution or centre. This approach strengthens sustainability and capacity-building.

Recommendations

Maintain flexibility in participant selection to include motivated individuals who may not fully meet predefined criteria but show potential as future trainers.

Adopt a broad and inclusive definition of Master Trainer, allowing multiple individuals to attain this role to support sustainability and institutional capacity.

3.4 Ethical, Legal, Societal Issues

Ethical, Legal, and Societal Issues are increasingly recognised as a cross-cutting, horizontal dimension that affects all tasks and scientific disciplines involved in OS. A key challenge is determining whether ELSI-related training should be designed at a generic level, applicable across domains, or tailored more specifically to the needs and contexts of each discipline. One important consideration is that legal issues are highly nuanced and often vary between countries due to differing national regulations. This makes it difficult to provide a one-size-fits-all training solution. In this round of recommendations, ELSI is therefore approached from a horizontal perspective, aiming to establish a shared baseline of understanding across disciplines. For example, within the SSH training course, no dedicated ELSI module was included. Instead, the course incorporated practical activities, such as understanding the use of Creative Commons licences and how these tools can empower researchers to participate more confidently in the OS landscape.

Crucially, the ELSI team has produced a range of valuable outputs, which we strongly encourage all contributors and trainers to read. These include use cases, guidelines, and practical insights that help clarify how ELSI considerations can be integrated into training programmes across different thematic areas. These resources are documented in the relevant project milestone and represent an important reference point for anyone involved in OS training design.

Recommendation

Integrate ELSI as a horizontal dimension within OS training programmes. Rather than creating standalone ELSI modules, training could embed ELSI considerations – such as licensing, data protection, and societal impact – within practical, discipline-relevant activities.

3.5 Fair-by-Design Methodology

This section summarises feedback on the FAIR-by-Design (FBD) methodology as applied to pilot training across several thematic areas. Overall, the methodology was viewed as helpful for guiding training design and promoting FAIR principles such as citation and reusability. However, experiences varied:

- **SSH** respected the FBD methodology, especially during the pilot phase, where the community produced (and reproduced) the output for the pilot training in GitHub using Markdown and Obsidian, as suggested by the WP2 team. However, this iteration was not repeated for the actual TtT and some clarity is needed to what extent all the tools can or should be used. For example, when it comes to creating learning objectives, the FBD was helpful, and it was followed to the letter.
- **Solid Earth Sciences** found the full methodology too complex, particularly the use of markdown, which posed an accessibility barrier. They adopted a simplified, adapted approach to ensure substantial compliance.
- **Open Scientific Collections** applied most parts of the methodology, and –as part of the training- highlighted its role in enabling citation of physical and digital objects using PIDs, aligning with reproducible research practices.
- **Climate Sciences** applied the methodology as fully as possible, noting its flexibility for future additions.
- **ELSI** found it very helpful, especially for reducing syllabus preparation time and improving access to recommended readings.

In essence, while the FBD methodology was positively received for promoting FAIR principles, its implementation required adaptation to community-specific needs and technical capacities, particularly in domains less familiar with digital tools. A clear recommendation for future training based on this feedback would be:

Recommendation

Adopt a flexible and context-sensitive approach to implementing the FBD methodology. Future training should apply the FBD methodology flexibly, adapting it to each community's needs and technical skills. Emphasis

should be on core FAIR principles like citation and reusability, rather than strict technical formats. Support and alternatives should be offered for complex tools, making it easier for all communities to engage meaningfully with FAIR practices.

3.6 Accreditation Framework

While the term "Accreditation Framework" is not explicitly mentioned in the sources, the thematic areas offer insights into how such a framework is implicitly understood or approached.

In Solid Earth Sciences, the idea of accreditation is most developed. Although formal certification is not currently provided beyond proof of attendance or completion, there is active discussion about introducing grading systems (e.g. participation points, final projects) and integrating training into university curricula with the potential for academic credits. This is closely tied to the broader goal of recognising OS and RDM competencies in ways that could support career development for researchers and support staff. However, progress at scale would require alignment with university structures and European education associations.

In Climate Sciences, accreditation is not a current focus. Training is seen primarily as a way to build thematic knowledge, and there have been no requests or expectations for certification or credits, even when training is delivered in academic settings. Open Scientific Collections interpret accreditation differently, focusing on the certification of digitised objects rather than the training of individuals. They recommend building on existing models like the Core Trust Seal[13] to develop a framework for certifying the authenticity of digital materials within the community.

Overall, the concept of an Accreditation Framework is interpreted differently across thematic areas—ranging from training recognition through grading and credits (Solid Earth Sciences), to non-prioritisation (Climate Sciences), and even to object-level certification (Open Scientific Collections). There is no unified approach, but some fields are beginning to explore pathways for more formal recognition of OS training.

[13] https://www.coretrustseal.org/

Recommendation

Develop a flexible and modular accreditation framework that can be tailored to the specific needs, maturity levels, and institutional contexts of different communities. This approach aligns with the broader goals of the Skills4EOSC project and supports global efforts to create meaningful mechanisms for recognising participation in OS training.[14]

3.7 Sustainability

Sustainability emerged as a central concern, intimately connected to accreditation. Sustainability encompasses both human and institutional dimensions: at present, becoming a trainer is largely a voluntary activity, with no formal incentives or compensation structures in place. While OS is grounded in a strong sense of moral responsibility and commitment to good scientific practices, this ethos alone cannot sustain long-term engagement. Expecting individuals to consistently contribute their time and expertise, especially in developing and delivering training, without adequate recognition or reward is neither equitable nor scalable.

A broader issue is that employers may not prioritise or support staff involvement in OS training unless it is clearly linked to funded activities or institutional mandates. This can limit participation, particularly for individuals whose roles do not formally include OS responsibilities or who are not directly affiliated with OS projects. As a result, engagement with the training materials and their continued use may decline once project funding concludes.

In addition to the human resource challenge, there is also a structural sustainability gap. As previously noted in the report, the end of project funding raises concerns about the long-term uptake of the recommendations, training content, and practices developed. Relying solely on volunteer trainers after the project's completion introduces a real risk that the envisioned impact will not be fully realised.

To support sustainability, it is essential to strengthen the connections between thematic communities, CCs, and the broader user support networks

11. For an outline of how accreditation is implemented in the context of Skills4EOSC, see Paolini (2025).

established through the project. Embedding training activities within institutional frameworks, exploring models for compensated or credited participation, and ensuring that training resources are maintained and updated beyond the project's lifetime will be key to achieving lasting impact.

Recommendations

Recognise and formalise trainer roles within institutions and research organisations, either through compensated time, professional development credits, or formal workload allocation. This helps ensure that OS training is not dependent solely on voluntary effort.

Establish mechanisms for post-project sustainability, such as:

- Partnerships with competence centres, universities, or research infrastructures that can host and maintain training resources.
- A shared support framework or coordination hub to connect thematic communities with ongoing user needs and facilitate the continued delivery of training
- Train-the-trainer programmes linked to certification or accreditation systems to enable a self-sustaining network of trainers.

4. Thematic Recommendations

4.1 Research Infrastructure Professionals

The pilot course on data management for RI professionals addressed the specific needs of managers, operators, and other professionals working in RIs and Core Facilities. It aimed to build capacity in implementing OS and FAIR Research practices. The pilot training, held in September 2024, engaged 30 RI professionals through a blend of online lectures, group discussions, role plays, and interactive tools. It was followed by a Train-the-Trainer course, designed to address current needs and challenges, identified as follows:

Need

There is a critical need to upskill RI professionals in OS and FAIR data practices. As data becomes increasingly central to RI operations, professionals must be equipped to develop and implement robust data management policies, ensure compliance with FAIR principles, and align with evolving OS mandates.

Challenges

- Varying needs and RI operational frameworks.
- Modules such as "FAIR principles" and "Data Management Plans" might be overly complex or content-heavy.
- Balancing foundational and advanced material, given the diverse professional audience and individuals' backgrounds.

Recommendations

Restructure complex modules into smaller sessions (e.g., splitting DMP topics into multiple parts for better absorption).

Adjust pacing and depth of content, especially in modules dealing with policy and FAIR principles as part of RI management procedures.

Introduce pre-course surveys or assessments to better understand the operational framework of Ris and participant needs; tailor the course accordingly.

Explore hybrid formats combining theoretical sessions with hand-on interactive modules.

Examples

- Indicative structure of a module on DMP:
  - Introduction to DMPs (purpose, funder requirements, templates).
  - Writing practical DMP sections (data collection, metadata, storage).
  - Case studies and reviewing real DMPs from RI projects.
- Staged approach on FAIR principles:
  - What FAIR means in theory.
  - FAIR in practice –how an RI implements FAIR data and metadata.
  - Use of FAIRification tools and procedures
  - Application of a policy framework

Module alignment

- The final version of the course incorporates advanced content modules to better meet the needs of experienced professionals.
- Participants from the pilot expressed the benefit of seeing real-world application of data workflows in RI settings, which were integrated into revised modules.
- The Skills4EOSC Learning Platform will host the course for broader access and integration with institutional training efforts, including those of the European School for Management of Research Infrastructures.

4.2 Social Sciences and Humanities

One of the persistent challenges in Open Science training for the SSH is the absence of a clear and widely accepted definition of what constitutes “data”. This ambiguity complicates efforts to develop effective training, share research materials, and apply FAIR principles consistently across SSH disciplines. Additionally, limited familiarity with Data Management Plans (DMPs) further hinders the ability to introduce structured approaches to data

or research material management. While existing DMP tools and templates can serve as valuable starting points for these conversations, they often require adaptation to reflect the specific practices, methodologies, and types of outputs common in SSH research.

Need

There is an urgent need to clarify and standardise the definition of SSH data to support shared understanding across disciplines and stakeholder groups. This clarity is crucial for effective data management planning, repository use, metadata documentation, and compliance with OS mandates.

Challenges

- The concept of "data" varies significantly across SSH disciplines due to diverse research traditions.
- There is no common typology or classification of SSH data to guide training and repository practices.
- The role of digital tools and software in SSH research adds complexity in identifying and categorising data (e.g., distinguishing between raw data, processed data, and analytical outputs).

Recommendations

Frame the definition of SSH data through a methodological lens, rather than by disciplinary boundaries, to accommodate the diversity of practices.

Promote and build upon existing classifications of SSH data, for example those used by SSH-focused data repositories.

Clarify the relationship between data, software, and analysis in digital research workflows by distinguishing between inputs, processes, and outputs.

Examples

- Qualitative methodology, common in SSH, offers a clear pathway to defining data, such as interview transcripts, observational notes, and coded responses.

- Data repositories like DANS and ICPSR[15] provide submission guidelines that implicitly define acceptable types of SSH data, offering a starting point for training discussions.
- In digital humanities, researchers use software to process and analyse texts or images; identifying the role of software helps clarify what constitutes data throughout the research process.

Module alignment

Within various modules, specific lectures and activities have been created to reflect the need for discussions on topics such as Data Management Plans, Open-Source Software, and data definition through methodological lens.

4.3 Solid Earth Sciences

The Solid Earth Sciences community is diverse, comprising multiple subdomains with varying data management practices, standards, and familiarity with FAIR, OS, and RDM principles. Despite efforts toward harmonization, challenges remain in standardisation, vocabulary alignment, and training availability. More advanced subdomains are often less adaptable due to entrenched standards. Overall IT skills are moderate to low, with limited support staff and training resources. RDM is typically handled at the individual project level and seen as overhead, lacking institutional dedicated tools, integration and long-term sustainability planning.

There is a high demand for OS/RDM and FAIR training within the community, but the time it takes to learn is often expected to be unrealistically short, partly due to the specific traits of the community, and to the trend toward microlearning (an aspect that has to be mitigated as it may result in superficial understanding and poor practical application).

Need

A key need for this community is to address time constraints and heterogeneous starting levels. Theoretical understanding is often necessary for beginners, while practical application requires familiarity with geoscience-specific data standards and legal frameworks. The EPOS Data Portal exemplifies these principles and serves as a central training tool.

[15] DANS: https://dans.knaw.nl/en/ ; ICPSR: https://www.icpsr.umich.edu/web/pages/

Training must be tailored to diverse audiences—ranging from theory-focused beginners (e.g. students) to practice-oriented Early Career Researchers.There's significant training potential due to the lack of structured, geoscience-specific OS/FAIR resources, with a broad target audience across EPOS ERIC, European universities, and research institutions.

Challenges

- Time constraints: it is especially difficult for professionals and senior researchers to commit to long courses due to specificities such as emergency response, shift work for surveillance monitoring and time off for sample collection and on-site campaigns with little to no connectivity.
- Varied expertise: wide range of starting levels of expertise within a community. The challenge lies in finding the right level to improve beginners' skills without causing experienced individuals to lose interest.
- Heterogeneous community: what may appear to be a single community from the outside is, in fact, a collection of diverse sub-domain communities, each with their own specific characteristics, standards and habits.
- Shortage of reusable, standardised materials: there is no single body of knowledge on how to apply OS, FAIR and RDM in Solid Earth Sciences.
- Declining university enrolment: Declining enrolment in Solid Earth Sciences directly impacts training audience. Courses are predominantly attended by senior, highly trained professionals, posing the challenge of evolving established practices. Notably, some courses have also seen participation from non-European students, adding a further dimension to the audience's diversity and training needs.

Recommendations

Flexibility in online, on-site, and hybrid formats is essential, particularly when integrating training into larger thematic events.

Short attention spans necessitate modular training strategies. Developing modules on domain-specific applications and tailoring the training on the scientific interests/needs of a specific subdomain helps to bring the training nearer to the user and make it more relevant for them.

Complement basic/advanced knowledge on OS/FAIR/RDM with handling data basics/automation of scientific data processing workflows (e.g. using Jupyter notebooks or Virtual Research Environments) that showcase novel reuses of FAIR/Open Data/Scientific products.

Integrate OS/FAIR training into university curricula to address university students and early career professional researchers.

Ensure long-term upskilling: encouraging Open Educational Resources and the adoption of a (simplified) FAIR-by-design methodology for their creation to maximise educational materials reuse and adaptation and creating a network of trainers for sustainable learning within the RI, also ensuring alignment of OS and RDM education with evolving research needs.

Design training for a multi-faceted audience and for outreach: It is recommended to adopt a dual approach. First, to counter declining enrolment, courses must be promoted as a strategic showcase for new talent, highlighting Solid Earth Sciences as a modern, high-impact career. Second, the curriculum must be tailored for the actual audience. For senior professionals, training should focus on high-value case studies that evolve established practices, rather than on basics. For the international participants, materials should be culturally aware and accessible, leveraging this diversity to foster global collaboration. A modular, flexible format is advised to suit the needs of these distinct groups.

Examples

- Importance of identifying data and the ethical implications of citing it.
- Importance of data provenance and establishing relations between datasets (who uses what and how).
- Importance of establishing relations between data and scientific literature.
- Include hands-on sessions using practical examples of tools:
  - EPOS ERIC Data Portal (https://www.ics-c.epos-eu.org/);
  - Assess the FAIRness of a published dataset using F-UJI (https://www.f-uji.net/)

- DMP Online (https://dmponline.dcc.ac.uk/), Argos (https://argos.openaire.eu/), or Data Stewardship Wizard (https://ds-wizard.org/)
- DataCite Commons (https://commons.datacite.org/) and APIs (https://support.datacite.org/docs/api)
- Identification of the most suitable data archive for publishing data using dedicated search engines such as re3data (https://www.re3data.org/)

Module alignment

The Learning Path serves as a guideline on OS, FAIR and RDM principles in the context of Solid Earth Sciences and uses the EPOS platform as an example of practical implementation. The modules foster skills in accessing and using data hosted by EPOS, with specific sections on interoperability, data management, and hands-on exercises.

4.4 Climate Sciences

Early Career Researchers working on Climate Sciences have specific needs to be addressed in training with respect to the practical application of the FAIR principles to datasets, workflows and provenance. Indeed, while the FAIR principles and theory are already integral part of the current modules, dealing with them in real cases introduces additional challenges which deserve dedicated and practical training. These are described more in detail in this section.

Need

Practical insights on data FAIRification and the application of FAIR principles to workflows and provenance jointly with the use of tools and platforms that support FAIR practices.

Challenges

- Gap between FAIR theory and concrete implementation.
- Low adoption of provenance practices in the community.
- Provenance tools still not established in the community.

- Low adoption of PIDs and PID practices which affects FAIRification of data, workflows and provenance.
- Heterogeneous landscape in terms of workflow management systems and provenance tools.

Recommendations

Develop dedicated training modules on FAIRification of data, workflows and provenance applied to concrete use cases, by making stronger use of tools and services from the research infrastructures/EOSC ecosystem.

Develop training modules on practical application of PID in concrete use cases, as a foundational component for the FAIRification of data, workflows and provenance in the climate domain.

Develop and promote community-based practice on FAIRification of data, workflows and provenance.

Establish a clear link between a theoretical introduction on the FAIR principles and more vertical modules on the FAIRification practices.

Examples

- Training at major community-based events (e.g., of the European Geosciences Union).
- Training at major RIs and cross-community events (e.g. EOSC conferences).

Module alignment

The current structure of the course introduces OS and FAIR, outlines FAIR principles in practice, and guides participants through the stages of planning for FAIR climate data. In future iterations, a new module titled “FAIRification in action!” could be integrated to address identified need with a comprehensive set of exercises/hands-on based on real use cases in the climate domain. The new module could include two sub-modules, respectively on:

- Specific aspects (e.g. application of PIDs).
- Data, Workflows and Provenance.

4.5 Open Scientific Collections

The thematic community working on Open Scientific Collections has specific needs with regards to interoperability through the uniform application of standardised PIDs and metadata schemas, the digitisation of physical objects and the consideration of ethical issues in historical collections.

Needs

- Professionals working with object-based LAM collections, have a particular need of implementing a FAIR digitisation of physical objects.
- Encourage interoperability of collections through Persistent Identifiers (PIDs) and metadata standards.
- Considering the historical contexts in which many physical objects in LAM institutions were collected, there is an urgent need to consider ethical aspects around their provenance in trainings for FAIR collections data management. This concerns mostly collections with colonial contexts, as well as from national-socialist (NS) times and/or other violent historical events.

Challenges

Digitisation:

- Identify relevant (domain-specific) metadata standards.
- Lack of certification systems to prove the authenticity of digitised material.
- Lack of resources to digitise collections, which are frequently of large to very large sizes.

Interoperability:

- Missing metadata entries because of lack of information regarding e.g. historical collections.
- The standardisation of PID implementation in data management planning.

Provenance:

- Raising personal and institutional awareness, as well as fostering sensitisation of colonial and NS contexts of collections.
- Capture and expose these contexts in standardised metadata entries.

Recommendations

Digitisation:

Capture a minimum set of information on the object according to domain-specific metadata standards.

Adapt a methodological approach for optimal digitisation (digital representation/twin) and cataloguing (metadata capture) for internal and external re-use of the data.

Encourage discussions and eventually a broader agreement in the domain specific community **on an authenticity certifications system at object-level**.

Increase stakeholder engagement (e.g. Science-Policy) to achieve more financial support for digital strategies and efforts in LAM institutions and increase the efficiency of processes by providing guidelines.

Interoperability:

Implement the concept for "globally unique persistent identifier", independent of internal or external re-use. Find a system for generating unambiguous PIDs according to collection management that comply with data re-use policies, using the FAIR-by-Design methodology.

Prepare in advance systems and/or platforms for linking data (e.g., complementary materials in archives and/or libraries) to tackle missing metadata.

Provenance:

Recognise –at institutional level- sensitive topics and assert a clear position towards the handling of these.

Encourage active engagement and discussion on multiple institutional levels through e.g. awareness-raising campaigns and/or trainings for employees.

Establish institutional and collection-specific policies for data publication and re-use that follow the CARE principles and data protection legal frameworks alongside the FAIR concept.

Examples

Digitisation:

- Domain-specific metadata standards like ABCD (Access to Biological Collection Data), DarwinCore, DC (Dublin Core), EDM (Europeana Data Model), LIDO (Lightweight Information Describing Objects).
- The Core Trust Seal certification system for repositories can be used as an example to build on an existing certification framework for digital repositories.

Interoperability:

- Domain-specific guidelines for biodiversity data: GBIF Guide to PIDs for biodiversity data .

Provenance:

- Clear positioning and recognition of institutional colonial background by the Museum of US in San Diego, California, USA: https://museumofus.org/decolonizing-initiatives
- Guidelines to work with objects of colonial contexts have been published before (e.g., Guidelines for German Museums. Care of Collections from Colonial Contexts - https://d-nb.info/122810221X/34), and can be used as an example to establish collection-specific policies for data publication and re-use.

Module Alignment

Need #1 is addressed in Module 2: Data digitisation, processing and preservation. Need #2 is covered in both Module 1: Planning and strategy, and Module 3: Sharing - making library, archive & museum collections FAIR. Need #3 is addressed in Module 4: Data governance – legal and ethical consideration.

4.6 Ethical, Legal, Societal Issues

Among the multiple stakeholder roles identified in the project, policy makers and civil servants engaging in evidence-aware / evidence-informed decision-making are in a unique position affording them opportunities for direct impact upon regulatory frameworks which are then binding upon multiple actors, or even at societal level. This position places such decision-makers in the crosshairs of advocacy groups representing various lobbying interests, typically with a strong asymmetry of resources in favour of industry stakeholders.

To further increase the challenge, civil servants and policy makers acting in digital regulation must necessarily deal with significant asymmetries of resources which again favour industry actors and may render policy makers susceptible to relying on policy choices favouring industry narratives and the interests of established commercial actors.

A variant of this dynamic has also been identified in the Belgium national training tackling academic and research policymaking through the lens of the interrelation between using and promoting OS practices, and the existing understanding of research integrity requirements.[16]

Need

A demand has been identified for cross-cutting upskilling of stakeholders, framed using the applicable MVS profiles, towards developing an awareness and independent ability to harness the potential of OS approaches and frameworks in mitigating and overcoming the resource asymmetries that ELSI experts and practitioners face in their daily decision-making. In particular, this needs to entail:

- Developing an understanding of the alignment of OS frameworks and approaches with the fundamental goals of applicable ELSI frameworks.
- Building an ability to map the required level of knowledge and skills that could be used by addressees of the relevant MVS profiles in developing a substantive independence including an increased resilience to narratives of industry stakeholders.
- Developing and strengthening an ambition for pursuing a more ethical and fundamental rights-aligned policymaking along with an understanding that

[16] For a report from the Belgian research integrity workshop, please see Deliverable 3.2 Evidence—based Policy and Public Administration training – final report [forthcoming].

even under the circumstances of strong power and knowledge asymmetries, the values of democratic societies can and should still be protected and preserved, and that supporting and fostering OS remains a major pillar of such protection.

Challenges

- Highly cross-cutting topics: Due to their broad context and a wide array of implications, courses on ELSI matters can be particularly content-heavy. When discussed in a training environment, they require a pre-existing broad perspective on society and its policy-related challenges in the digital age. During the pilot phase, this has also been identified as a blocking issue for some stakeholders to understand the relevance of the course to their needs.
- Weaknesses of indirect training formats: While the preparation phase can be fully self-paced, an ELSI OS course should ideally feature a values-based discussion to ensure an understanding of the deeper underlying foundation of the discussed approach. While easier to perform in a classroom context, this has been found to be quite challenging in an online setting with varying levels of engagement on the part of learners.
- Heterogeneity of capacity among stakeholders: Due to the broad nature of ELSI topics, the EIDM courses launched under Skills4EOSC identified significant differences among participants in a course in terms of their substantive preparation and levels of familiarity with the diverse topics, rendering the discussion more difficult to follow for those less well prepared.

Recommendations

Combine training with awareness-raising efforts: To tackle the relatively high barriers of entry that may preclude potential learners from understanding the relevance of OS ELSI training to their daily work, training opportunities should be advertised while coupled with awareness-raising campaigns to demonstrate the added value in the given community.

Compartmentalisation of content: as demonstrated in the Train-the-Trainer deployments of ELSI training under WP3, learners find broadly formulated ELSI courses quite challenging. Attention should be paid to the individual backgrounds of the training groups, and their particular training needs to avoid overloading the curriculum.

Maintain a momentum during live courses: due to a strong values-based component, it is helpful to tackle the broad content in a face-to-face setting. The activities envisaged in the course outline should be devised to maximise engagement within subgroups and a lively discussion with emotional engagement. When organising an in-person session is not feasible, learners should be required to keep their cameras and microphones on.

Examples

- Example 1: in the pilot phase, some learners have reported their initial reluctance to join the training due to a perceived lack of relevance of OS training to their line of work. Only after discussing the topic with representatives of Skills4EOSC, they would understand the added value and join the training. In practice, this role should be played by a more organised awareness-raising effort built into the ELSI-themed OS training.
- Example 2: during deployment of the EIDM ELSI TtT courses for civil servants and policy makers, some participants reported that inclusion of the theoretical foundations of legal ethics was overly theoretical and not useful to ELSI practitioners. As a remedy, it was proposed to avoid overlaps by focusing on ethical rules already embedded in the law and thus focus on the relevant regulatory framework while ensuring a slightly less content-laden curriculum.
- Example 3: The values-based discussions during ELSI-themed OS live sessions have been found to suffer due to uneven engagement and difficulty to maintain rapport with some of the learners who would prefer to keep their cameras off or would not enable their microphones to take the floor. In the later deployment, participants were asked to keep their cameras on as a minimum requirement throughout the session to improve the engagement in the discussion.

Module Alignment

The Needs identified above are addressed jointly in the following Modules of EIDM Train of Trainers Course 2: "Ethical, Legal and Societal Implications (ELSI) and Data Governance", part of the Learning Path for OS and Evidence-informed Decision-making:

- Module 06: Legal and Ethical Frameworks and Considerations in OS
- Module 07: OS under the EU Data Regulatory Framework
- Module 08: Data Governance and Legislative Strategies for FAIR Research.

Each Module comprises two lectures and is designed to be as stand-alone as possible, to enable tailoring content to the needs of the specific audience.

5. Vision for the Future

The thematic trainings confirmed that discipline-sensitive Train-the-Trainer programmes can accelerate the uptake of OS and RDM practices. Workshop exchanges in Milan and subsequent Competence Centres' briefings highlighted that, to maximise their long-term benefit, the materials produced during the pilots would benefit from mechanisms that keep them up to date and widely accessible beyond the project lifetime. The points below highlight areas that could be addressed in subsequent phases of Skills4EOSC or in complementary EOSC initiatives.

Curating training assets – a light governance workflow covering version control, scheduled peer review and DOI assignment would help keep slide decks, exercises, MVS profiles current and citable after project funding ends.

Accreditation and micro-credentials – developing a micro-credential aligned with European qualifications frameworks could facilitate formal recognition of Skills4EOSC courses by universities and CC.

Adaptation toolkit for MVS profiles – a customise-and-deploy package, with editable templates and illustrative thematic examples, would enable task groups and CC to refine generic MVS profiles for additional disciplines and for centres at different stages of maturity.

Multilingual and inclusive delivery – a coordinated translation workflow and shared glossaries would enhance accessibility for centres operating in non-English environments and promote wider participation.

Incentives for sustained trainer engagement – modest grants, teaching-load recognition or service contracts could help maintain an active pool of master trainers and reduce reliance on voluntary effort.

Horizon scanning and rapid updates – an annual review of emerging technologies and policy developments such as AI-enabled research workflows or EU data spaces, would support updates to the curriculum.

Cross-centre collaboration platform – a lightweight registry that lists accredited trainers, hosts updated modules and records local adaptations could foster a distributed yet coherent training ecosystem.

Addressing these aspects – governance, accreditation, adaptation support, linguistic accessibility, trainer incentives and forward-looking monitoring –

would strengthen the durability of the thematic recommendations and support their integration into the evolving EOSC skills landscape.

6. Conclusions

D5.6 of the Skills4EOSC project outlines a forward-looking and inclusive strategy for advancing OS and FAIR-aligned RDM practices across research communities. Drawing on extensive pilot training activities, expert workshops, and collaboration with thematic partners and Competence Centres, this document provided both horizontal (cross-disciplinary) and thematic (domain-specific) recommendations aimed at embedding OS/RDM principles in diverse scientific and policy-making settings.

The recommendations collectively emphasise the importance of:

- Modular, progressive learning paths that are adaptable to different knowledge levels, institutional contexts, and learner roles.
- Use of Minimum Viable Skillsets and Train-the-Trainer models to enable scalable, community-driven capacity building.
- Deepening attention to Ethical, Legal, and Societal Issues, particularly in relation to data governance and regulatory frameworks.
- Applying and adapting the FAIR-by-Design methodology to ensure transparency, reusability, and integration in training development.
- Exploring accreditation options and micro-credential systems to incentivise and formally recognise OS training.
- Ensuring sustainability through institutional integration, trainer incentives, reuse of Open Educational Resources, and partnerships beyond project funding cycles.

The report also emphasises that successful OS/RDM training requires deep engagement with each thematic community's operational realities, from the complexity of defining "data" in the SSH, to the infrastructural demands of FAIR implementation in Climate and Solid Earth Sciences, to provenance-sensitive digitisation in cultural heritage collections.

Across disciplines, persistent structural challenges emerged: under-resourced training environments, fragmentation in technical standards, a scarcity of qualified trainers, and varying levels of policy readiness. However, the broad engagement in this work, including contributions from research infrastructures, government bodies, NGOs, and pan-European networks, points to a shared readiness to address these issues collaboratively.

Looking ahead, sustained impact will hinge on formal mechanisms that support:

- Continued evolution and localisation of Minimum Viable Skillsets profiles.
- Shared access to adaptable training content.
- Clearer frameworks for certification and trainer recognition.
- Ongoing alignment with the EOSC vision and updated technological and legal landscapes.

Ultimately, D5.6 reflects the belief that expanding access to open, FAIR-aligned research is not solely a technical challenge; it is a cultural and systemic transformation that requires inclusive pedagogical frameworks, ethical awareness, and committed institutional leadership.

7. References

No	Description/Link
R1	Access to Biological Collection Data task group. 2007. Access to Biological Collection Data (ABCD), Version 2.06. Biodiversity Information Standards (TDWG) http://www.tdwg.org/standards/115. Last accessed: 5 August 2025.
R2	Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., … Hudson, M. (2023). The CARE Principles for Indigenous Data Governance. In Open Scholarship Press Curated Volumes: Policy. Retrieved from https://openscholarshippress.pubpub.org/pub/xx3kj9rv
R3	CESSDA. (n.d.) Main website. https://www.cessda.eu/ Last accessed: 5 August 2025.
R4	CLARIN ERIC. (n.d.) Main website. https://www.clarin.eu/. Last accessed: 5 August 2025.
R5	CSC – IT Center for Science(n.d.) Main website. https://csc.fi/en/. Last accessed: 5 August 2025.
R6	Cocco, M., Michalek, J., Nedrebø, H., Stuebe, K. C., & Tanlongo, F. (2024). D5.3 - Learning Path and Training Materials on "Open Science and Research Data Management in Solid Earth Sciences" (1.0.0). Zenodo. https://doi.org/10.5281/zenodo.13684718
R7	Core Trust Seal. (n.d.) Main website. https://www.coretrustseal.org/ . Last accessed: 5 August 2025.
R8	Corleto, A., Di Giorgio, S., Paolini, G., Berberi, L., Candela, L., Costantini, A., Gaido, L., Green, D., JANIK, J., Tuminauskas, R., Whyte, A., Lazzeri, E., Prandoni, C., & Evangelinou, B. (2025). D7.3 Report on CCs and user support networks and recommendations for networks evolution. Zenodo. https://doi.org/10.5281/zenodo.15262091
R9	D'Anca, A., & Maltese, V. (2024). D5.4 OS and RDM learning paths for Climate change communities. Zenodo. https://doi.org/10.5281/zenodo.14797326
R10	DANS. (n.d.) Main website. https://dans.knaw.nl/en/. Last accessed: 5 August 2025.
R11	DARIAH. (n.d.). Main website. https://www.dariah.eu/. Last accessed: 5 August 2025.
R12	DarwinCore. (n.d.).https://dwc.tdwg.org/. Last accessed: 5 August 2025.
R13	DataCite. (n.d.). Datacite Commons. https://commons.datacite.org. Last accessed: 5 August 2025.

R14	DataCite. (n.d.). Datacite APIs. https://support.datacite.org/docs/api. Last accessed: 5 August 2025.
R15	Data Stewardship Wizard. (n.d.). Main page. https://ds-wizard.org/. Last accessed: 5 August 2025.
R16	Digital Curation Centre (n.d.) DMP Online. https://dmponline.dcc.ac.uk . Last accessed: 5 August 2025.
R17	DiSSCo. (n.d.) Main website. https://www.dissco.eu/. Last accessed: 5 August 2025.
R18	Drążewski, K. (2025). MS5.6 Report: ELSI requirements for RI professionals & thematic communities. Zenodo. https://doi.org/10.5281/zenodo.15299344
R19	EMSO. (n.d.) Main website. https://emso.eu/. Last accessed: 5 August 2025.
R20	ENES. (n.d.) Main website. https://portal.enes.org/index.html. Last accessed: 5 August 2025.
R21	EPOS ERIC. (n.d.) Main website. https://www.epos-eu.org. Last accessed: 5 August 2025.
R22	EPOS ERIC. (n.d.). Data Portal. **https://www.ics-c.epos-eu.org/**. Last accessed: 5 August 2025.
R23	ESGF. (n.d.). Main website. https://esgf.github.io/index.html Last accessed: 5 August 2025.
R24	Filiposka, S., Green, D., Mishev, A., Kjorveziroski, V., Corleto, A., Napolitano, E., Paolini, G., Di Giorgio, S., Janik, J., Schirru, L., Gingold, A., Hadrossek, C., Souyioultzoglou, I., Leister, C., Pavone, G., Sharma, S., Mendez Rodriguez, E. M., & Lazzeri, E. (2023). D2.2 Methodology for FAIR-by-Design Training Materials (1.4). Zenodo. https://doi.org/10.5281/zenodo.8305540
R25	F-UJI – Automated Research Data Assessment Tool. (n.d.). Main website. https://www.f-uji.net/. Last accessed: 5 August 2025.
R26	Gdańsk University of Technology. (n.d.). Main website. https://pg.edu.pl/en/gdansk-tech-library. Last accessed: 5 August 2025.
R27	German Museums Association (2021). Guidelines for German Museums. Care of Collections from Colonial Contexts, 3d Edition. **https://d-nb.info/122810221X/34**.
R28	GRNET. (n.d.). Main website. https://grnet.gr/en/ . Last accessed: 5 August 2025.

R29	ICOM. (n.d.). Main website. https://icom.museum/en/. Last accessed: 5 August 2025.
R30	ICPSR. (n.d.). Main website. https://www.icpsr.umich.edu/web/pages/. Last accessed: 5 August 2025.
R31	Karlsruhe Institute of Technology. (n.d.). Main website. https://www.kit.edu/english/ . Last accessed: 5 August 2025.
R32	Lavitrano, M. (2025). D5.1 FAIR and RDM training modules for RI professionals. Zenodo. https://doi.org/10.5281/zenodo.15731892
R33	LIDO working group. (n.d.). LIDO overview. https://cidoc.mini.icom.museum/working-groups/lido/lido-overview/. Last accessed: 5 August 2025.
R34	Linés, C., Whyte, A., Gänsdorfer, N., Ritschard, E., Gottwald, I., Kiesel, M., Rainer, H., Gothlin Illsley, W., Souyioultzoglou, I., Gingold, A., Di Giorgio, S., & Prandoni, C. (2024). D5.5 FAIR and RDM training modules for open collections. Zenodo. https://doi.org/10.5281/zenodo.14537731
R35	Mystakopoulos, F., Avanço, K., Souyioultzoglou, I., Gingold, A., Sokolova, E., & Delmazo, C. (2024). D5.2 OS and RDM learning paths for Social Sciences and Humanities communities. Zenodo. https://doi.org/10.5281/zenodo.14797210
R36	Museum of US (n.d.): Decolonizing Initiatives https://museumofus.org/decolonizing-initiatives. Last accessed: 5 August 2025.
R37	OpenAIRE. (n.d.). Argos. https://argos.openaire.eu/ . Last accessed: 5 August 2025.
R38	Paoling, G (2025). European Digital Credentials for Learning (EDC) vs Open Badges: Benefits, limitations and our experience. https://www.skills4eosc.eu/resources/events/paris-10-june-2025/49-10-european-digital-credentials-for-learning-edc-vs-open-badges
R39	Recherche Data Gouv. (n.d.). Main website. https://recherche.data.gouv.fr/en. Last accessed: 5 August 2025.
R40	re3data - Registry of Research Data Repositories. (n.d.). Main website. https://www.re3data.org. Last accessed: 5 August 2025.
R41	RDA. (n.d.). Main website. https://www.rd-alliance.org/. Last accessed: 5 August 2025.
R42	Skills4EOSC. (n.d.). Learning platform. https://learning.skills4eosc.eu Last accessed: 5 August 2025.
R43	Swedish National Data Service. (n.d.). Main website. https://snd.se/en. Last

	accessed: 5 August 2025.
R44	University of Saint Cyril and Methodius. (n.d.). Main website. https://www.ucm.sk/en/. Last accessed: 5 August 2025.
R45	Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. T*he FAIR Guiding Principles for scientific data management and stewardship*. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18
R46	Whyte, A., Green, D., Avanço, K., Di Giorgio, S., Gingold, A., Horton, L., Koteska, B., Kyprianou, K., Prnjat, O., Rauste, P., Schirru, L., Sowinski, C., Torres Ramos, G., van Leersum, N., & Lazzeri, E. (2024, June 4). Minimum Viable Skillsets - Catalogue of Open Science Career Profiles. Zenodo. https://doi.org/10.5281/zenodo.11469300

Annex 1 – Glossary of Key Terms

Accreditation – A structured approach for recognising and validating learning outcomes, potentially linked to the European Qualifications Framework (EQF) and micro-credentials.

Adaptation toolkit (for MVS profiles) – A set of editable templates and examples designed to help Competence Centres tailor Minimum Viable Skills (MVS) profiles and learning objects to specific disciplines and maturity levels.

CARE principles – Guidelines that complement FAIR for Indigenous data governance: Collective Benefit, Authority to Control, Responsibility and Ethics.

Competence Centre (CC) – A national or institutional hub providing expertise, training and support on Open Science and Research Data Management.

Creative Commons (CC) licence – A standardised licence that enables creators to grant permissions for reuse of their work under specified conditions.

Data Management Plan (DMP) – A formal document that outlines how research data will be handled during and after a project, covering collection, storage, sharing and preservation.

DOI (Digital Object Identifier) – A persistent identifier that provides a stable, resolvable link to a digital resource.

EIDM – Evidence-Informed Decision-Making, the use of the best available evidence to guide policy and regulatory choices, particularly in digital governance contexts.

ELSI – Ethical, Legal and Societal Issues, cross-cutting considerations addressing moral, regulatory and social dimensions of research and data practices.

FAIR principles – Foundational guidelines for data management: Findable, Accessible, Interoperable, Reusable.

FAIR-by-Design methodology – A design approach that embeds FAIR principles into training materials from the outset, facilitating reuse and citation.

Hands-on tools – Practical software, platforms or services used in training sessions to demonstrate OS/RDM workflows (e.g. NetCDF utilities, GitHub).

Learning Path – A sequenced set of modules that guides learners from foundational to advanced competencies in Open Science and RDM.

Master trainer – An individual who has completed Train-the-Trainer courses and is qualified to train others, thereby multiplying capacity.

Metadata standard – A formal specification (e.g. ABCD, Darwin Core, LIDO, NetCDF-CF) that defines structured elements for describing data.

Micro-credential – A small, stackable certification of learning outcomes, often aligned with EQF levels and expressed in ECTS.

Minimum Viable Skills (MVS) – The core set of competencies required to engage meaningfully with Open Science and RDM; used as a baseline for curriculum design.

NetCDF – Network Common Data Form, software libraries and data formats for array-oriented scientific data, widely used in climate science.

Open Educational Resource (OER) – A teaching, learning or research resource that is freely available for adaptation and reuse under an open licence.

Open Science (OS) – A scholarly communication model that aims for transparency, reproducibility and broad accessibility of research outputs.

Persistent Identifier (PID) – A long-lasting reference (e.g. DOI, Handle, ORCID) that reliably points to a digital or physical object.

Research Infrastructure (RI) – Facilities, resources or services that support research communities in their data generation, analysis and preservation efforts.

Train-the-Trainer (TtT) – A pedagogical approach in which participants learn both subject content and didactic skills, enabling them to train others.

Annex 2 – Preliminary recommendations

This annex presents the recommendations discussed and refined during the 2nd co-design workshop (Milan, 26-27 February 2025).

Social Sciences and Humanities

Section	T5.2 – SSH Summary
Area for Recommendation	Focused on diverse SSH disciplines; noted the difficulty in defining data, the limitations of MVS profiles, and the need for tailored approaches across domains.
Learning Path	Flexible and adaptable across disciplines; learning sequence shaped by prior OS knowledge and research roles.
Learning Objectives	Inclusion of data definition, specific outputs (e.g. monographs, transcripts, datasets), and consideration for licensing and language differences.
Stakeholders	Included researchers, educators, data managers, and European infrastructures like OPERAS, DARIAH, CLARIN, CESSDA.
Training Content	Balanced theoretical and practical content; OER emphasis; delivery shaped by language, licensing, and national maturity.
Key Challenges Identified	Key barriers included definitional ambiguity, weak incentives, lack of formal structures for OS uptake and training.
Any Other Issues	Explored certification, dissemination strategies, and integration into broader academic frameworks.

Solid Earth Sciences

Section	T5.3 – Solid Earth Sciences Summary
Area for Recommendation	Highlighted the fragmentation of the community into subdomains; noted need for domain-specific adaptation of MVS profiles and lack of reusable materials.
Learning Path	Structured around a two-module core with optional additions; addressed constraints like time availability and skill diversity.
Learning Objectives	Three core LOs: understanding OS/FAIR,

	applying principles in context (EPOS), and using data platforms effectively.
Stakeholders	Engaged ECRs, data providers, infrastructures (EPOS, EMSO), universities, and funders.
Training Content	Adapted FAIR-by-Design to practical levels; included EPOS portal use, NetCDF tools; supported online and blended formats.
Key Challenges Identified	Challenges included time limitations, disciplinary fragmentation, trainer shortages, and varying technical competencies.
Any Other Issues	Certification not prioritised; dissemination and curriculum integration were long-term goals, including OER expansion.

Climate Sciences

Section	**T5.4 – Climate Sciences Summary**
Area for Recommendation	Community handles massive data volumes but lacks consistent, updated training materials. Training is fragmented and often project-based. MVS useful but insufficiently detailed.
Learning Path	Two core modules with optional additions tailored to event type and audience. Hands-on components are embedded and adaptable.
Learning Objectives	Three modules: general OS/FAIR intro, thematic FAIR in modelling, and planning for FAIR/RDM. Theory-practice split is 2+1 hours per module.
Stakeholders	Researchers, data providers, and data managers. ENES and ESGF infrastructures are central. Universities and climate orgs are key targets.
Training Content	FAIR-by-Design applied; content includes ESGF use, NetCDF, visualisation tools. Tools for feedback include Google Forms and Mentimeter.
Key Challenges Identified	Main issues: lack of central training repository, systemic fragmentation, absence of sustainable training strategy. Survey insights pending.
Any Other Issues	Certification not prioritised; emphasis on

	knowledge transfer. Communication and dissemination require a coordinated long-term strategy.

Open Scientific Collections

Section	**T5.5 – Open Scientific Collections Summary**
Area for Recommendation	Targeted curators and RDM leads in LAM institutions. Developed a custom MVS for digital collection curators; focused on digital transformation of physical collections.
Learning Path	Developed a 4-module learning path including case studies; focus was on RDM, Open Science, FAIR, and CARE principles for digital collections.
Learning Objectives	Learning objectives tied to working with collections as data. Included digitisation, metadata, licensing, and sensitive provenance issues.
Stakeholders	LAM institutions, funding bodies, academic and professional networks (e.g., ICOM, RDA), and infrastructures like DiSSCo were key stakeholders.
Training Content	Training covered metadata standards, licensing, use of PIDs, and sensitive data management. FAIR-by-Design supported citation of physical objects.
Key Challenges Identified	Challenges included lack of authenticity certification, digitisation standards, metadata gaps, provenance sensitivities, and evolving technologies (incl. AI)
Any Other Issues	Lessons learned focused on the demand for practical skills, importance of PIDs, and establishing a curators network. Survey results collected.

Ethical, Legal, Societal Issues

Section	**T5.6 – ELSI and Data Governance Summary**
Area for Recommendation	Targeted policy makers and civil servants in digital regulation, including those influencing research-based policy through

	funding mechanisms.
Learning Path	Three-module course follows OS/FAIR introduction: legal/ethical frameworks, EU data regulation, and FAIR-aligned governance strategies.
Learning Objectives	Aimed at building resilience to industry narratives, promoting values-based policy, and reinforcing democracy-supporting Open Science.
Stakeholders	Mapped at EU and national levels, including European Commission DGs, ministries, policy NGOs, funders, and research-performing organisations.
Training Content	Used FAIR-by-Design to streamline preparation and improve access to core readings; training balanced law, ethics, and governance topics.
Key Challenges Identified	Challenges included information asymmetry, broad topic scope, superficial coverage due to limited time, and difficulty sustaining engagement.
Any Other Issues	Feedback stressed the need for more time for practical work and open discussion; course length needs to grow without losing accessibility.