

Skills 4 eosc

D5.2 - OS and RDM learning paths for Social Sciences and Humanities communities

Lead Partner:	OPERAS
Version:	1.0
Status:	Submitted draft not yet approved by the EC
Dissemination Level:	PU - Public
Document Link:	https://zenodo.org/uploads/12805055
DOI	10.5281/zenodo.12805055

Deliverable Abstract

This deliverable outlines a training framework designed to help researchers in the Social Sciences and Humanities (SSH) adopt Open Science-compliant practices. The framework targets the research workflows within SSH and introduces the necessary skills and competencies to support an open and collaborative research methodology.



Skills4EOOSC 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 (<http://creativecommons.org/licenses/by/4.0/>). 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

Date	Name	Partner/Activity	Date
<i>From:</i>	OPERAS		27/08/2024
<i>Moderated by:</i>	Karla Avanço Fotis Mystakopoulos	OPERAS	
<i>Reviewed by:</i>	Iryna Kutchma [EIFL] Laurent Delvalle (AMU)		19/08/2024
<i>Approved by:</i>	Sara Di Giorgio	GARR	27/08/2024

DOCUMENT LOG

Issue	Date	Comment	Authors
v0.1	16/07/2024	Draft version	Karla Avanço (OPERAS), Carol Delmazo (OPERAS), Arnaud Gingold (OPERAS), Fotis Mystakopoulos (OPERAS), Elena Sokolova (OPERAS), Irakleitos Souyioultzoglou (OPERAS)
v0.2	19/08/2024	WP review	—
v0.3	27/08/2024	Final draft	Karla Avanço (OPERAS), Carol Delmazo (OPERAS), Arnaud Gingold (OPERAS), Fotis Mystakopoulos (OPERAS), Elena Sokolova (OPERAS), Irakleitos Souyioultzoglou (OPERAS)
v1.0	29/08/2024	Approved	Sara di Giorgio (GARR)

TERMINOLOGY

<i>Terminology/Acronym</i>	<i>Definition</i>
FAIR	Findable, Accessible, Interoperable, Reusable
MVS	Minimum Viable Skillset
OA	Open Access
OS	Open Science
RDM	Research Data Management
RI	Research Infrastructure
SSH	Social Sciences and Humanities
ToT	Train-of-Trainers

Contents

Executive summary	7
1.Introduction	8
2.Learning path and training material	11
2.1 Methodology.....	11
2.2 Compliance to the FAIR-by-Design methodology.....	17
2.3 Module Description.....	19
2.4 Overview of the training material.....	19
3.Training Pilots	21
3.1 Pilot Delivery.....	21
3.2 Feedback and Lessons Learned	21
4.Next Steps	23
5.References	24

Executive summary

This deliverable outlines a training framework designed to help researchers in the Social Sciences and Humanities (SSH) adopt Open Science-compliant practices. The framework targets the research workflows within SSH and introduces the necessary skills and competencies to support an open and collaborative research methodology.

The deliverable is the main output of T5.2 "OS and RDM in the Social Sciences and Humanities", which addresses current and emerging Open Science (OS) and FAIR RDM practices in this disciplinary area. Accordingly, the proposed training materials (available on a [github repository](#)¹) have been designed to build on the learning path created as part of T5.2, which aims to provide a sustainable framework for ongoing development by addressing challenges such as diverse data types and formats, limited adoption of standards, ethical and legal concerns, skills gaps and lack of awareness of OS and FAIR principles.

The methodological approach involves an in-depth analysis of requirements, standards, and a pilot training event to validate and refine the materials based on stakeholder feedback. Additionally, a key component is the Training of Trainers (ToT) approach, which aims to establish a network of skilled professionals within the network of collaborating Competence Centres. This ensures the continued adoption of necessary standards, tools, and best practices for effective OS and FAIR RDM workflows in the SSH fields.

¹ <https://ssh-tot-fair-by-design.github.io/OS-RDM-SSH/latest/>

1. Introduction

In the context of Open Science (OS), the FAIR (Findability, Accessibility, Interoperability, and Reusability) principles are an essential element for the curation and dissemination of research data. The implementation of FAIR-compliant Research Data Management (RDM) practices in the disciplinary fields of the Social Sciences and Humanities (SSH) fosters an open and collaborative research methodology. This deliverable outlines a comprehensive training program that addresses the specificities of research workflows in the SSH and introduces to the research community the necessary skills and knowledge to adopt FAIR practices.

The deliverable is the main output of T5.2 “OS and RDM in the Social Sciences and Humanities”. The task specifically addresses current and emerging practices in Open Science and FAIR RDM. Drawing on the experience of the participating institutions and through collaboration with SSH Research Infrastructures (RIs), it aims to create a consistent learning path along with relevant training materials that can be reused in future training initiatives. The overall objective is to build a sustainable framework for continuous development in these key areas, by addressing a spectrum of identified challenges, such as:

- **Diversity of data and formats:** SSH research generates a spectrum of data resources, including textual documents, images, audio and video recordings, survey responses, interview transcripts, etc. Each data type often requires specific methods for their curation, which complicates efforts to implement standardised methodologies. Additionally, the heterogeneity of data formats requires specialised tools and expertise, which significantly limits the potential of adopting FAIR RDM practices comprehensively.
- **Limited uptake of standards:** researchers in the SSH face challenges in adopting and implementing standards. The limited introduction of uniform standards applicable across the SSH disciplines reflects the

scarcity of robust frameworks that address the diverse needs of researchers and the differences in data formats and practices.

- **Ethical and legal aspects:** SSH research often involves sensitive and personal data, which are subject to ethical guidelines and data protection regulations. These frameworks complicate efforts to make data openly accessible and interoperable while maintaining compliance with legal requirements (like GDPR, for example).
- **Limited skills:** As training in digitally-oriented practices is not a standard practice in the field of SSH, the skills gap results in difficulties in ensuring data interoperability and reusability, as well as in using digital tools and services. Additionally, the evolution of data management technologies and practices requires that researchers continuously update their skills to effectively organise, document, and curate their data in ways that align with FAIR principles.
- **Lack of awareness** on FAIR RDM and OS principles: one of the main barriers toward the adoption of OS and FAIR practices in the SSH is the prevailing lack of awareness: many researchers are not adequately informed about the benefits of OS and the ways that FAIR practices can facilitate their research and enhance the visibility, reproducibility, and impact of their work. Consequently, targeted advocacy efforts for OS and FAIR and guidance on implementing workflows are instrumental in fostering the adoption of relevant practices in the disciplinary fields.

The methodological approach is based on a collaborative design process that leverages the diverse expertise of project partners and Open Science experts and is practically implemented as an in-depth analysis of the specific requirements and challenges, followed by a series of consultations. To validate and refine the training materials, the task designed and implemented a pilot training event, which was delivered to a selected group of stakeholders, including researchers, representatives of RIs, and scholarly communication professionals. The feedback from these participants resulted in the refined resources presented in this report. This iterative process

ensures that the final learning path is both practical and relevant, thus addressing the needs of the target audience.

A key component of this task is the Training of Trainers (ToT) approach. By training skilled professionals, it aims to establish a team of experts in the network of collaborating Competence Centres, ensuring the future uptake of OS practices within these institutions and their respective communities. A foreseen follow-up training event will be delivered to potential trainers, who will be equipped with an in-depth understanding of the standards, applications, tools, and best practices required for effective OS and FAIR RDM workflows in the SSH disciplinary fields.

2. Learning path and training material

The training design was based on a workflow that initiated with defining competencies and skills. We then proceeded to the instructional design phase, which was compliant to the FAIR-By-Design Methodology² - to ensure that the training materials were findable, accessible, interoperable and reusable. In this framework, each design step undergoes a FAIR compliance checking. As part of the co-design methodology, representatives of other SSH research infrastructures (DARIAH, CEESDA, and CLARIN) reviewed the final iteration of the learning path.

The structure of the training program corresponds to the various phases of the research and data lifecycle. Emphasis was placed on exploring the workflows involved in managing research data, and establishing connections between the RDM process and a publication strategy that aligns with Open Science practices.

2.1 Methodology

An essential part of training design is the identification of the target audience. In the case of this task, the audience consists of Social Sciences and Humanities (SSH) researchers and/or support staff who would like to become trainers on OS in the SSH domain. The SSH disciplines are multiple and heterogeneous in terms of practices, outputs, and methodological approaches. For this reason, this task targets SSH researchers and support staff in general, without specifying a disciplinary domain or level of proficiency.

With the audience identified, the first step in training design is determining the competencies and skills to be developed. These were summarised in the form of learning outcomes, which were compiled in alignment with the Catalogue of Open Science Career Profiles (or Minimum Viable Skillset - MVS)³, which is one of the main outputs of the project. The Minimum Viable

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

³ Whyte et. al. (2023). D2.1 Catalogue of Open Science Career Profiles - Minimum Viable Skillsets

Skillset "describes essential skills and concepts required to deliver Open Science outcomes for communities and organizations" (Whyte et al., 2023). Based on two MVSs (those for early career researchers and senior researchers), the training designers identified a new set of skills and competencies for SSH researchers - outlined as follows:

- Knowledge of policies, opportunities, and OS practices in SSH.
- Ability to recognize SSH-specific OS principles and identify relevant practices at every stage of the research workflow.
- Knowledge of the research and data life cycle in the SSH.
- Ability to make OS-compliant choices to collect, annotate, and document data and software, create metadata, use relevant taxonomies, and make data available for publication and preservation in repositories.
- Adequate capacity to implement SSH-specific FAIR principles.
- Ability to assess the FAIRness of existing resources.
- Capacity to produce a DMP to plan data management at the beginning of a research project.
- Ability to apply open publication practices, such as publishing preprints, publishing in open access journals and platforms, and ensuring data and code availability in open repositories -to the extent possible.

The next step in the designing of the learning path was to support the accomplishment of these competencies and skills. The analysis of the MVS for SSH researchers showcased that all skills and knowledge correspond to the phases of the research life cycle; thus, the learning outcomes were mapped to these stages and subsequently categorised into learning modules. As this training focuses on SSH researchers in general, (i.e. without addressing the specificities of each discipline) the learning path is based on a simplified version of the research lifecycle, one that includes three phases: a) planning research; b) active research, and c) dissemination and preservation of research.

(v1.2). Zenodo. <https://doi.org/10.5281/zenodo.8101903>



Fig.1 – Representation of the research life cycle tailored for the trainings

Each module of the training focuses on one of the three phases of the life cycle. It is noteworthy that, as it is a cycle, it starts and ends at the same point: planning research. At this particular phase, it is crucial for researchers to be able to assess the quality and FAIRness of training resources, research materials/ data, and other research outputs and know how to reuse them. However, this topic was addressed during Module 3, for simplicity's sake. The skills and competencies addressed in each module and their interpretation into learning outcomes are presented in Figures 2, 3, and 4.

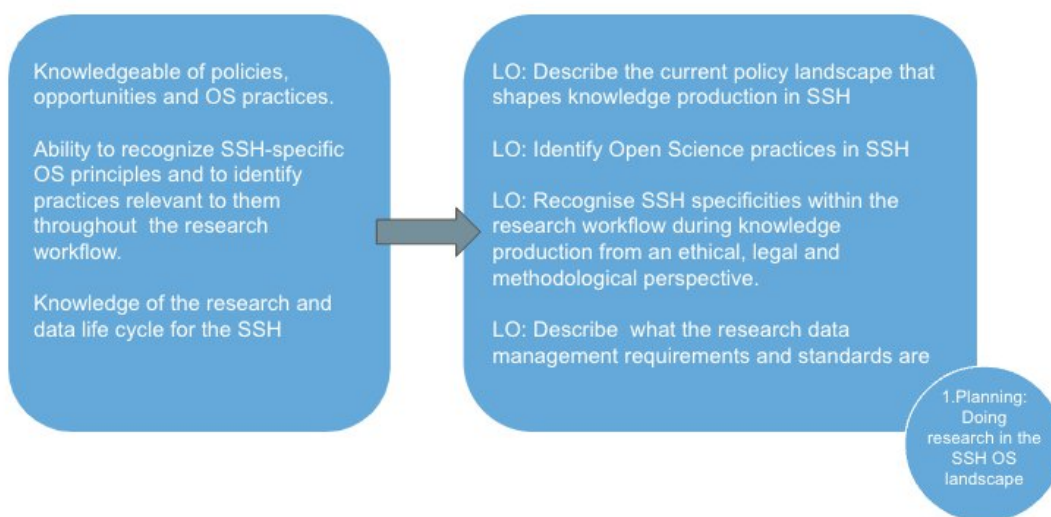


Figure 2. From competencies and skills to desired learning outcomes (Module 1)

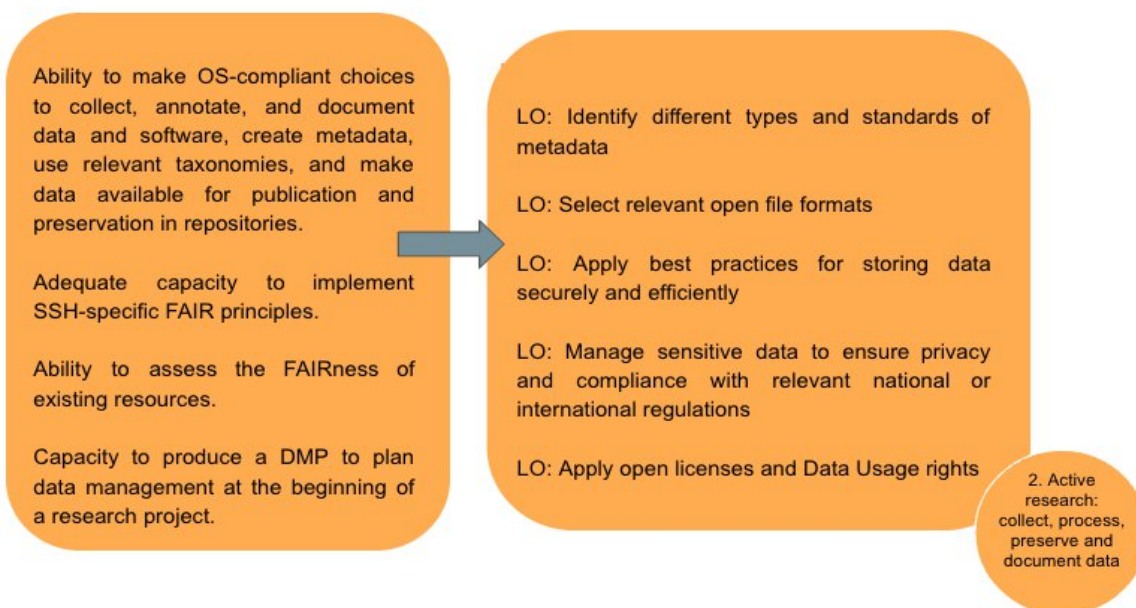


Figure 3. From competencies and skills to desired learning outcomes (Module 2)

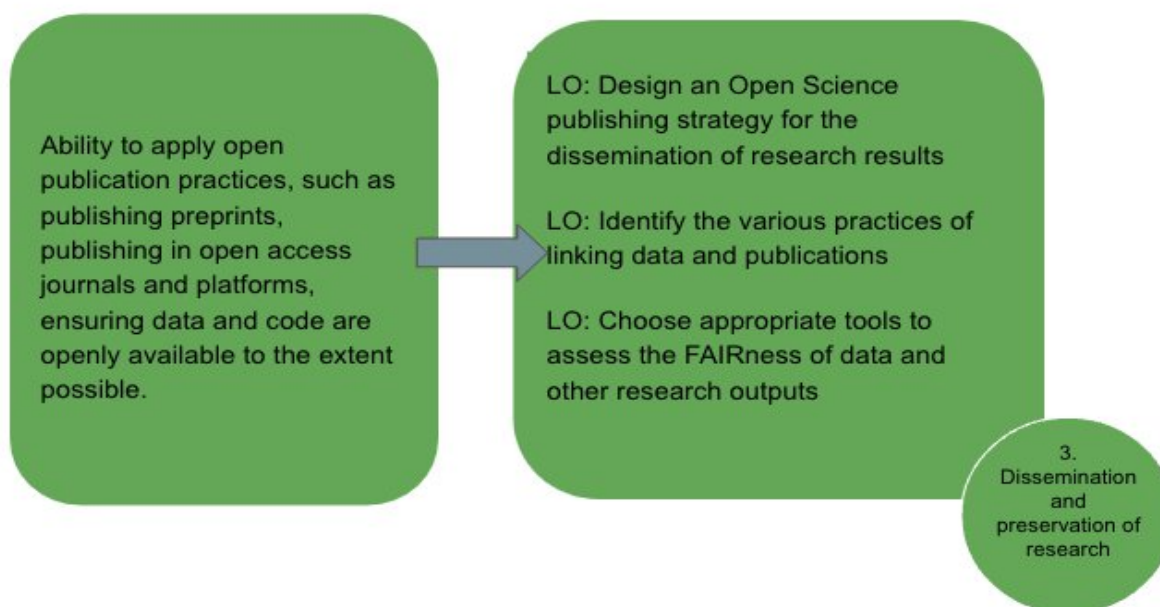


Figure 4. From competencies and skills to desired learning outcomes (Module 3)

Both the set of competencies and skills, as well as the set of learning outcomes, can be further developed to meet the needs of different SSH disciplines and subdomains, or adapted to different proficiency levels

An additional module with guidelines for trainers was included. This module focuses on the methodology of training design - description of learning outcomes, preparation of a learning path, and planning of the learning experience.

The elements of the learning path are summarised in the table below:

Target audience	Researchers in the Social Sciences and Humanities
Type of training	Train of Trainers / Virtual
Granularity	Adapted for beginners, addressing the research phase through general terms and concepts
Competencies and skills	<ul style="list-style-type: none"> • Knowledgeable of policies, opportunities and OS practices. • Knowledge of the research and data life cycle for the SSH • Ability to recognize SSH-specific OS principles and to identify practices relevant to them throughout the research workflow. • Ability to make OS-compliant choices to collect, annotate and document data and software, create metadata, use relevant taxonomies, and preserve data • Adequate capacity to implement SSH-specific FAIR principles. • Ability to assess FAIRness of existing resources. • Capacity to compile a basic DMP. • Ability to apply open publication practices, such as publishing preprints, publishing in open access journals and platforms, ensuring data and code are openly available to the extent possible.

Learning outcomes	<p>Module 1</p> <p>LO: Describe the current policy landscape that shapes knowledge production in SSH</p> <p>LO: Identify Open Science practices in SSH</p> <p>LO: Recognise SSH specificities within the research workflow during knowledge production from an ethical, legal and methodological perspective</p> <p>LO: Describe what the research data management requirements and standards are</p> <p>Module 2</p> <p>LO: Identify different types and standards of metadata</p> <p>LO: Select relevant open file formats</p> <p>LO: Apply best practices for storing data securely and efficiently</p> <p>LO: Manage sensitive data to ensure privacy and compliance with relevant national or international regulations</p> <p>LO: Apply open licenses and Data Usage rights</p> <p>Module 3</p> <p>LO: Design an Open Science publishing strategy for the dissemination of research results</p> <p>LO: Identify the various practices of linking data and publications</p> <p>LO: Choose appropriate tools to assess the FAIRness of data and other research outputs</p> <p>Module 4</p> <p>LO: Recognise the three stages of backward design</p> <p>LO: Write learning outcomes</p> <p>LO: Design a learning path</p>
Composition	3 modules on OS and RDM + 1 with Guidelines for Trainers

Table 1. Elements that constitute the learning path

Instructional design is an iterative process. Accordingly, the learning path is the result of subsequent iterations that took place during meetings to conceptualise the training, after the delivery of a pilot training (where the learning path was presented to a group of participants), and during the finalisation of the presented version of the training material.

2.2 Compliance to the FAIR-by-Design methodology

The process of the designing and development of the training followed the FAIR-by-Design methodology (Filiposka, 2023). The steps implemented are presented in Table2, based on the FAIR-by-Design Microlearning framework⁴.

<i>Step</i>	<i>Implementation</i>	<i>Comments</i>
1. Prepare		
Do you understand FAIR implications?	Yes	
Define purpose, learning objectives, target audience	Yes	
2. Discover		
Finding existing resources	Yes	
Identify potential reuse	Yes	
3. Design		
Define syllabus and structure	Yes	
Identify granularity and facilitation materials	Yes	
4. Produce		

⁴Filiposka, S., Mishev, A. (2024). FAIR-by-Design Microlearning (1.0.0). <https://doi.org/10.5281/zenodo.11548062>

Develop content using common apps and formats	Yes	Formats and tools suggested in the internal FAIR-by-Design Train of Trainers
Define machine readable metadata	Yes	
5. Publish		
Define license and other related info	Yes	Resources available under CC BY 4.0
Release to public for instructors and learners	In Progress	Training material currently under formatting
Enable feedback gathering	Yes	Feedback form provided
6. Verify		
Final Quality Assessment check	In Progress	Feedback collected after the pilot. Waiting for final QA framework to be developed.
Use gathered feedback for continuous improvement	In Progress	Feedback gathered during the pilot was taken into account. Participants in the ToT sessions will also be asked to provide feedback.
7. Continuous Improvement		
Create a list of potential improvements	In Progress	
Choose a set of improvements to be implemented and start a new release cycle	Yet to start	To be implemented in collaboration with the Competence Centres

Table 2. FAIR by-Design methodology checklist based on Microlearning.

2.3 Module Description

The first module of the training (Planning) is an introduction of the Open Science environment to SSH researchers. This module has four learning outcomes and eight learning units.

The second module (Active Research) focuses on what we call “active research”. It presents elements that correspond to different stages of the research: data collection, data processing, preservation and documentation. In this section, RDM and more technical aspects related to the FAIR Principles are detailed. This module has five learning outcomes and five learning units.

The third module (Dissemination and Preservation) focuses on the evaluation of the research outcomes and their posterior preservation and dissemination. The module also tackles access and reuse of research outcomes. This item is actually part of the first step of the research life cycle. It represents the beginning of a new cycle that draws from research outcomes properly preserved and disseminated. This module has three learning outcomes and three learning units.

The fourth module (Guidelines for Trainers) contains three learning outcomes and two learning units.

2.4 Overview of the training material

The training material can be found in the following link: <https://ssh-tot-fair-by-design.github.io/OS-RDM-SSH/latest/>

It consists of a Gitbook with the resources that the task produced. The Gitbook comprises four modules and 18 learning units. For each learning unit, there is textual content with the following elements:

- Description of the unit with learning outcome.
- Learning tools and resources necessary for each part of the training: it consists of a slide presentation to be used during the training. Some

units may also present exercises and other types of supplementary material.

- The content of the unit.
- References.
- Further Reading.

The textual content of the unit and the presentations that accompany it are complementary to each other.

3. Training Pilots

3.1 Pilot Delivery

The pilot was a milestone achieved in month 18 of the project (February 2024). It was delivered as a series of three subsequent sessions (with a total duration of 9h), which included presentations by the Skills4EOSC trainers, discussions with the participants, and breakout sessions that enabled the trainees to discuss in more detail the practices and workflows presented in the training materials. A total of 21 trainees, with a background in different SSH disciplines, attended the webinars.

The pilot focused on the three modules that address the skills and competencies described in the Minimum Viable Skillset for Researchers. At this stage, the training materials consisted of slide presentations that provided an overview of the practices and the fundamental aspects of Open Science, FAIR data, and RDM in SSH. Additional training materials that address pedagogical considerations were integrated into the third session.

3.2 Feedback and Lessons Learned

The participants in the pilot were invited to give their feedback about the training. The suggestions we received can be grouped in the three categories presented below:

- Improve the pace of training:
 - spend less time presenting basic concepts and more time giving advice for trainers on how to address the different topics
 - dedicate more time to exercises
- Improve delivery:
 - speak slower
 - include subtitles
- Improve the material:
 - add more references to relevant resources

The training materials were updated on the basis of the relevant suggestions. The alterations include (a) production of textual material to accompany the slide presentations; (b) inclusion of a “Further Reading” session at the end of the learning units, which lists extra resources; (c) improvement of the visual elements of the slides. A supplementary module (module 4) was produced to address specifically the pedagogical aspects of training.

To improve the delivery, we will extend the duration of the training. This will allow a more extensive coverage of the pedagogical aspects and provide more time for practical activities.

Finally, we will improve training delivery at the best of our capacity, including the addition of technical features such as subtitles.

4. Next Steps

The delivered learning path serves as a basic framework that can be reused and further adapted to meet specific needs and practices in diverse SSH research areas. Thus, the main focus of the task in the final year of the project will be on designing key actions to ensure the dissemination and adoption of the training resources. This will involve scheduling training sessions, such as webinars and consultation sessions, to engage in further collaboration with the targeted stakeholders.

Additionally, to sustain the training efforts and facilitate continuous learning, the task will focus on setting up a team of master trainers affiliated with the network of Competence Centres. The master trainers will undergo a training program based on the piloted training and will be consulted on the relevance of the suggested topics to the contexts in which their institutions operate. As part of the process, workflows will be identified to provide ongoing support and guidance to other stakeholders and ensure that the training resources remain current beyond the project timeframe.

5. References

<i>No</i>	<i>Description/Link</i>
R1	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
R2	Filiposka, S., Mishev, A. (2024). FAIR-by-Design Microlearning (1.0.0). https://doi.org/10.5281/zenodo.11548062
R3	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., Sharma, C., Méndez, E., & Lazzeri, E. (2023). D2.1 Catalogue of Open Science Career Profiles - Minimum Viable Skillsets (v1.2). Zenodo. https://doi.org/10.5281/zenodo.8101903

Documents and slides relevant to the learning paths for SSH Communities are also available in Zenodo:

https://zenodo.org/records/12805056/files/OSandRDMforSSH_material.zip