



FAIRSFair
Fostering Fair Data Practices in Europe

Project Title	Fostering FAIR Data Practices in Europe
Project Acronym	FAIRsFAIR
Grant Agreement No	831558
Instrument	H2020-INFRAEOSC-2018-4
Topic	INFRAEOSC-05-2018-2019 Support to the EOSC Governance
Start Date of Project	1st March 2019
Duration of Project	36 months
Project Website	www.fairsfair.eu

D6.4 FINAL REPORT ON COMPETENCE CENTRE WITH KNOWLEDGE BASE

Work Package	WP6 - FAIR Competence Centre
Lead Author (Org)	Gabin Kayumbi (STFC)
Contributing Author(s) (Org)	Elizabeth Newbold, Katie Yates (STFC), Angus Whyte (DCC), Laura Molloy (CODATA), , Linas Cepinskas (DANS), Bridget Walker (RDA)
Due Date	28.02.2022
Date	23.02.2022
Version	1.0 DRAFT NOT YET APPROVED BY THE EUROPEAN COMMISSION
DOI	10.5281/zenodo.6119392

Dissemination Level

- PU: Public
 PP: Restricted to other programme participants (including the Commission)

- RE: Restricted to a group specified by the consortium (including the Commission)
- CO: Confidential, only for members of the consortium (including the Commission)

Versioning and contribution history

Version	Date	Authors	Notes
0.1	04.11.2021	Gabin Kayumbi	Initial structure
	16.12.2021 -31.01.2022	Gabin Kayumbi, Elizabeth Newbold, Angus Whyte, Katie Yates, Laura Molloy, Linas Cepinskas, Bridget Walker, Joy Davidson, Patricia Herterich, Laurence Horton	Contributions from work package partners
0.2	01.02.2022	Gabin Kayumbi	Version for internal review
	15.02.2022	Ellen Leenarts (DANS), Celia Van Gelder (DTL)	Internal Reviewers comments
1.0	23.02.2022.	Gabin Kayumbi, Elizabeth Newbold, Katie Yates	Content ready

Disclaimer

FAIRSF AIR has received funding from the European Commission's Horizon 2020 research and innovation programme under the Grant Agreement no. 831558 The content of this document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of such content.

Abbreviations and Acronyms

CODATA	Committee on Data of the International Science Council
D6.1	FAIRsFAIR Deliverable 6.1 Overview of needs of competence centres
D6.2	FAIRsFAIR Deliverable 6.2 Initial Core Competence Structures
D6.3	FAIRsFAIR Deliverable 6.3 Established Competence Centre for Variety of Communities
D6.5	FAIRsFAIR Deliverable 6.5 Report on three annual schools in core data skills for researchers
D6.6	FAIRsFAIR Deliverable 6.6 Report on five instructor training events
D6.7	FAIRsFAIR Deliverable 6.7 Report on schools run through franchings with local organisers
DCC	Digital Curation Centre
DC TAP	Dublin Core Tabular Application Profile
DMP(s)	Data Management Plan(s)
ENVRI-FAIR	Environmental Research Infrastructures building Fair services Accessible for society, Innovation and Research
EOSC	European Open Science Cloud
EOSC Future	Project funded from European Commission H2020 INFRAEOSC-03-2020. Platform for open science and FAIR data, resources and services in all scientific disciplines
EOSC Pillar	Coordination and Harmonisation of National Initiatives, Infrastructures and Data services in Central and Western Europe
EOSC Synergy	European Open Science Cloud - Expanding Capacities by building Capabilities
ESFRI	European Strategy Forum on Research Infrastructures

ExPaNDS	European Open Science Cloud (EOSC) Photon and Neutron Data Service
FAIR	Findable, Accessible, Interoperable, Reusable
ICTP	International Centre for Theoretical Physics
IG	Interest Group
INFRAEOSC-5b	Projects resulting from European Commission H2020 INFRAEOSC-05-2018-2019 sub topic b (EOSC Pillar, EOSC NORDIC, ExPaNDS, EOSC Synergy, Ni4OS)
HEIs	Higher Education Institutions
LOM	Learning Object Metadata
LRMI	Learning Resource Metadata Initiative
M6.9	FAIRsFAIR Milestone M6.9 Report on Final Review of Fair Competencies Landscape
NORF	National Open Research Forum
OER	Open Educational Resources
PaNoSC	Photon and Neutron Open Science Cloud
RDA	Research Data Alliance
RDA-ETHRD-IG	RDA Education and Training on Handling of Research Data Interest Group
RDM	Research Data Management
RPO	Research Performing Organisation
SRIA	Strategic Research and Innovation Agenda
SSHOC	Social Sciences & Humanities Open Cloud
T4FS	Terms4FAIRskills
TeSS	ELIXIR's Training e-Support System
WP3	FAIRsFAIR work package 3 - FAIR Data Policy and Practice
WP4	FAIRsFAIR work package 4 - FAIR Certification
WP6	FAIRsFAIR work package 6 - Competence Centre

Executive Summary

Emerging from the European Open Science Cloud (EOSC), the FAIRsFAIR project aimed at supplying practical solutions for the use of FAIR data principles throughout the research data life cycle, and in particular fostering FAIR data culture and the uptake of good practices in making data FAIR. Within that context of FAIRsFAIR project, WP6's main objective was particularly concerned with addressing the effective support of FAIR uptake by engaging with communities through the establishment of a FAIR Competence Centre whose activities were mapped along the trinary: *Advisory, Harmonisation, and Dissemination*.

The output of WP6 is contained in deliverables that report on the journey towards its main objective, including the tasks leading from the development of concepts to their implementation and practical operations. The first deliverable D6.1 (Herterich, P. *et al.*, 2019) explored the *Overview of needs for competence centres*; the second, D6.2 (Newbold, E. *et al.*, 2021) developed concepts in the *Initial core competence centre structures* while the third, D6.3 (Kayumbi, G. *et al.*, 2021) reported on practical operations in *Established competence centre for variety of communities*. The present deliverable D6.4 *Final report for competence centre and knowledge base* reports on the tasks accomplished by WP6 in the final stage of project and offers a comprehensive perspective on the work conducted.

Community engagement being at the core of its remit, WP6 put a particular focus on developing collaborations with internal and external stakeholders as well as creating synergies that informed the different stages of work in addressing the objective it was assigned. This approach ultimately led WP6 into revisiting in this document the set of minimal criteria that characterise a viable competence centre, previously presented in D6.2. This time, those criteria are also examined in the light of the collaborations developed with FAIR-related projects, in the context of a dynamic landscape of FAIR initiatives. Finally, this report concludes by formulating a set of key messages on community engagement through FAIR Competence Centre that focuses on *collaboration, community building, training material, and training & skills delivery*. Similarly, looking forward, the report proposes recommendations on the aforementioned themes for ongoing and future initiatives. The final section of this report also includes a supplement of details in Appendix 1 about the scale of the trainers network.

Table of contents

Executive Summary	5
1. Introduction	9
2. Competence Centre: Advisory - FAIRdata Forum	11
3. Competence Centre: Dissemination	16
3.1 Training Activities	16
3.2 Data Stewards online course	20
3.2.1 Description	20
3.2.2 Reusability considerations	21
3.3 Knowledge Base Development: Training Library	22
4. Competence Centre: Harmonisation	24
4.1 Findability: consensus on minimal metadata for materials	25
4.1.1 Testbed of training materials/resources	26
4.2 Accessibility: characterising collections and their content	27
4.3 Interoperability	28
4.3.1 Developing a metadata application profile to support semantic interoperability	28
4.3.2 Controlled vocabularies for the metadata properties	30
4.3.3 The Terms4FAIRskills terminology and browser	31
4.3.2.1 The T4FS Materials Browser	34
5. Evaluation of minimal viable criteria for a competence centre	35
6. Key Messages and Recommendations	37
6.1 Collaboration	37
6.2 Community building	38
6.3 FAIR training materials	38

6.4 Delivery of training and skills	39
7. Conclusion	41
8. Bibliography	42
Appendix 1 Scale of the trainers network	44
A.1. Core Skills in Data Science Summer Schools	44
A.2 Data Stewards Instructor Training	44
A.3. Collaboration	46
A.4 Promoting the training events	46
A.5 Blog posts	47
A.6 Feedback	47

Table of Figures

Figure 2.1a FAIRdata Forum visits (partial)	12
Figure 2.1b FAIRdata Forum visits (global)	12
Figure 2.2a Forum page views	13
Figure 2.2b Forum page views from anonymous users	13
Figure 2.2c Forum page views on low traffic	14
Figure 2.3 FAIR data competencies in universities	14
Figure 2.4 An example of users discussions	15
Figure 2.5 FAIR-Aware	16
Figure 3.1 Data Steward online course modules	20
Figure 3.2 Example of Data Steward online course content	21
Figure 3.3 Training Library	23
Figure 4.1 Course using schema from schema.org	29
Figure 4.2 Tabular Application Profile	30
Figure 4.3 Ten EOSC actors	32
Figure 4.4 T4FS Terminology	33
Figure 4.5 FAIRsFAIR Materials Browser	35
Figure A.1 Data Steward Instructor Training	45

Table of Tables

Table 3.1 Data Steward Training events	18
Table 4.1 Minimal metadata for learning resources	26
Table A.1 Data Science Summer Schools	44
Table A.2 Collaborations with external organisations	46

1. Introduction

The variety of domains, established practices, and culture have led to different levels in FAIR data uptake and compliance across the landscape of research communities. Thus, the FAIRsFAIR project acknowledged the need for an adaptive approach in engaging with communities whilst promoting and supporting FAIR data uptake. The creation of the Competence Centre addressed that need for community engagement and constituted the main objective of WP6.

The work conducted in WP6 is detailed in reports which are briefly summarised in this section. The first report, D6.1 *Overview of needs for competence centres* (Herterich, P. *et al.*, 2019)¹, provided an analysis of the landscape of available FAIR-related competence centres with a focus on gaps in disciplines and features. The report presented a summary of characterisation of 36 competence centres, based on the following aspects: type of initiative offering the tools and/or services; domain targeted; type of end user targeted (researcher, research support staff, policy makers, etc.); tools/services provided; how the initiative was resourced. The report included a synthesis of interviews conducted with FAIR stakeholders. Additionally, the report outlined expectations of the FAIRsFAIR project for the project's Competence Centre and needs raised by the community in surveys, interviews and workshops. Finally, D6.1 presented a series of recommendations on advisory services, harmonisation and dissemination of outputs that the Competence Centre could offer which will be the foundation for any future work. The initial explorations pointed out a very dynamic landscape of FAIR initiatives and the need to seek continuous input from communities as the project progressed. The following were highlighted as key functionalities when developing the core competence centre:

1. *a definition of a web-based communications platform and its functionality, which can be efficiently managed as a delivery platform for material on FAIR data practice and a communication hub for user queries and user generated content.*
2. *the definition and design of a knowledge base suitable for cataloguing and accessing information on FAIR data practices. This should allow for cross-referencing to information maintained in other external data sources.*
3. *the design of policies and practices of the competence centre, including the mediation of user queries and criteria for developing additional material for harmonisation and training purposes.*

The landscape analysis in D6.1 included a characterisation of competence centres which has now been updated in *M6.9 Report on Final Review of Fair Competencies Landscape* (Newbold, E. & Kayumbi, G., 2021)² to reflect the changing landscape of new FAIR initiatives and projects.

¹ Herterich, Patricia, Davidson, Joy, Whyte, Angus, Molloy, Laura, Matthews, Brian, & Kayumbi Kabeya, Gabin. (2019). D6.1 Overview of needs for competence centres (1.0 DRAFT). FAIRsFAIR. <https://doi.org/10.5281/zenodo.3549791>

² Newbold, Elizabeth, & Kayumbi-Kabeya, Gabin. (2021). M6.9 Report on Final Review of Fair Competencies Landscape (V1.0). Zenodo. <https://doi.org/10.5281/zenodo.5708089>

The second report, *D6.2 Initial core competence centre structures* (Newbold, E. *et al*, 2021)³, described the set-up of the core competence centre, including human resources, initial knowledge base design and tools, communications infrastructure, defined responsibilities, and expectations on service levels. This report focused on the design and functionality of a Competence Centre and how it would meet the needs of its user base. It defined the scope of a Competence Centre, and the characteristics for a minimum viable competence centre. Following the recommendations from the first report and based on the continuous interactions with stakeholders and user communities, the design of functionalities aimed as much as possible to avoid unnecessary duplications and kept the implementation under constant review while ensuring the approach met stakeholders' needs.

D6.3 Established competence centre for variety of communities (Kayumbi, G. *et al*, 2021)⁴ was concerned with the development of a competence centre as a model of engagement and support for research communities. The emphasis was put on the description of operations of the core competence centre, including initiatives aiming to identify synergies and areas of harmonisation that are required to support the development of a knowledge base. The main objective for the period covered by this report was to translate the structures described in D6.2 into operational mechanisms.

The operations described in D6.3 were developed around the three main roles of the FAIRsFAIR Competence Centre.

- **Advisory** (*networking space to enable stakeholder collaborations and routing people to expertise in thematic areas*);
- **Dissemination** (*training material and courses; enable sharing of best practices*);
- **Harmonisation** (*consistently describe resources relating to FAIR data stewardship; provide leadership in harmonisation of policies, tools and resources*).

D6.1 and D6.2 have analysed the landscape of competence centres and have included desktop research and surveys in their approach. An update of the landscape analysis in D6.1 is presented in *M6.9 Report on Final Review of Fair Competencies Landscape* (Newbold, E. & Kayumbi, G., 2021). In contrast, D6.3 was descriptive and reported on operations.

The structure for the final report D6.4 is outlined as follows:

i) reporting on WP6 activities during the period of time covering the final stage of FAIRsFAIR project, detailing tasks and operations contributing to the development of FAIR Competence Centre. Where external activities of relevance to the Competence Centre with the participation of, but not initiated within WP6 are reported that will be explicitly mentioned;

³ Newbold, Elizabeth, Kayumbi Kabeya, Gabin, Matthews, Brian, Davidson, Joy, Herterich, Patricia, Whyte, Angus, & Molloy, Laura. (2020). *D6.2 Initial Core Competence Centre Structures (1.0 DRAFT)*. Zenodo. <https://doi.org/10.5281/zenodo.3732889>

⁴ Kayumbi-Kabeya, Gabin, Newbold, Elizabeth, Whyte, Angus, Cepinskas, Linas, & Molloy, Laura. (2021). *D6.3 Established Competence Centre for Variety of Communities (1.0 Draft)*. Zenodo. <https://doi.org/10.5281/zenodo.4560474>

ii) presenting the work conducted within WP6 and also in conjunction with other partners during the project's final stage;

iii) finally, formulating key lessons learned and proposing recommendations drawn from the work contained in previous reports, and also drawing from the aforementioned i) and ii).

The report includes a supplement of information in Appendix 1 on the scale of the trainers network resulting from the training activities for both data science and data stewardship.

2. Competence Centre: Advisory - FAIRdata Forum

Conceptually, the FAIR Competence Centre is a shared hub of expertise that connects people and resources on FAIR related matters. We have materialised this concept by building elements that combine to form the FAIR Competence Centre. i) The FAIRdata Forum⁵ which is an online platform providing space for community engagement and supports knowledge and skill transfer through training. On the Forum, interactions users-users and users-experts on FAIR matters occur. This online space also provides support to training activities, offering space to FAIR issues in HEIs whilst engaging with general users on FAIR related topics. ii) The Training Library⁶, building elements of a knowledge base, mainly focusing on the provision of FAIR training material. All these elements of the FAIR Competence Centre⁷ are readily accessible through the FAIRsFAIR website⁸.

Since the Competence Centre advisory role has been previously defined and its operations through the FAIRdata Forum sufficiently described in D6.3 (Kayumbi, G. *et al*, 2021), herein we will limit ourselves to reporting on the community engagement activities that followed the publication of the aforementioned report. We would however remind readers that whilst the FAIRdata Forum is open to the public for browsing, only registered users can post or comment and only group members can access a restricted area of the Forum specifically dedicated to that group. Such restricted areas could be community-managed and serve the purpose of internal interactions.

A user-experience test of the FAIRdata Forum has been conducted by our partner the Digital Curation Centre. Its scope included Forum access, navigation and functionalities. The resulting suggestions were part of final improvements carried on the Forum.

To date, FAIRdata Forum counts 134 registered users, including 20 moderators. It has supported nine Data Steward Training events and created two additional community-managed spaces. One was dedicated to users from Higher Education Institutions (HEIs) to present and discuss specific needs and challenges of Research Data Management (RDM) competencies and skills intended for

⁵ <https://fairdataforum.org/>

⁶ <https://www.fairsfair.eu/competence-centre/training-library>

⁷ <https://www.fairsfair.eu/competence-centre>

⁸ <https://www.fairsfair.eu/>

doctoral level research⁹. The other aimed at supporting the work and activities of the FAIRsFAIR Synchronisation Force¹⁰.

Figure 2.1a reports the number of all users visits (February 2021 - January 2022, a period not covered by previous reports) with expected peaks corresponding to events supported by the forum. A similar pattern was observed during the previous reporting period of time (June 2020 - January 2021) as illustrated in Figure 2.1b.

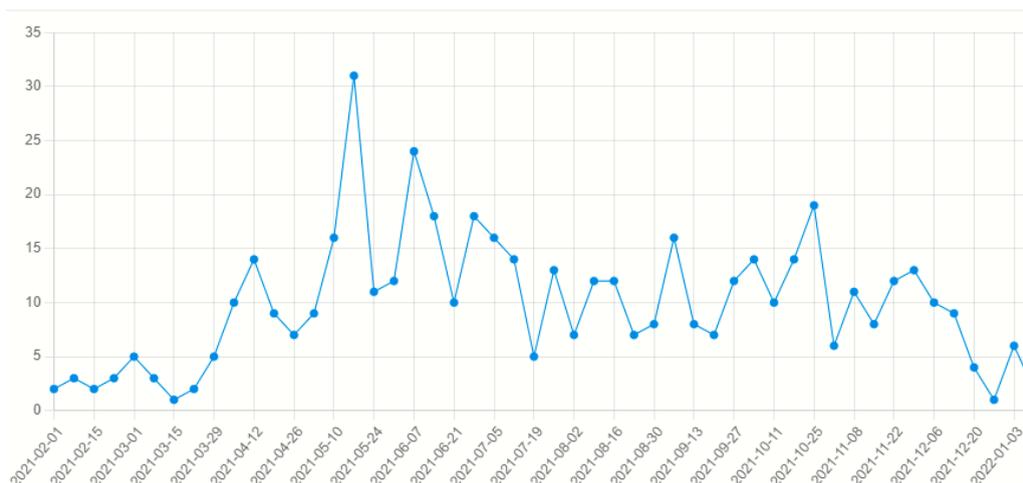


Figure 2.1a. FAIRdata Forum visits (partial). This figure shows the number of all user visits (Feb. 2021 - Jan. 2022). The numbers reported are accumulated visits during a week.



Figure 2.1b. FAIRdata Forum visits (global). This figure shows the number of all user visits (June 2020 - Jan 2021). The numbers reported on this figure are accumulated visits in a week. Peaks of visits signal the occurrence of events supported by the FAIRdata Forum.

⁹ <https://fairdataforum.org/t/about-the-fair-data-competences-in-universities-category/320>

¹⁰ <https://fairdataforum.org/t/about-the-third-synchronisation-force-workshop/260>

However, those figures do not discriminate between visits from registered users (including forum maintenance tasks by administrators and moderators) and anonymous users. Hence the difficulty in using these figures to measure visits from the wider members of the public.

A different way of measuring visits is by reporting figures on consolidated page views¹¹ instead (Figure 2.2a). When stripped of logged-users to retain anonymous users only, those analytics could represent a way of assessing the number of visits from the wider public, albeit only indirectly (Figure 2.2b).

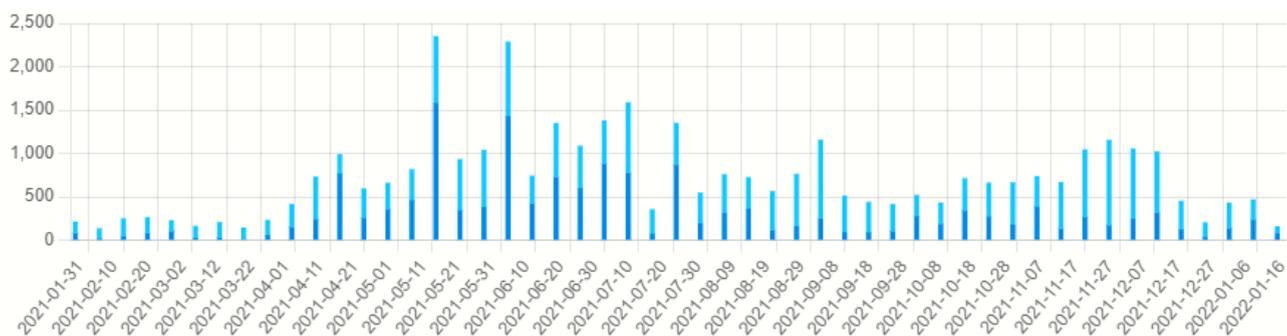


Figure 2.2a. Forum page views. This figure reports the number of pageviews (this metric factors in the page landing and reloading from a given user) from all users, including administrators and moderators, between Jan. 2021 and Jan. 2022. The dark blue represents logged-in users and light blue anonymous users. Those analytics constitute an indirect measure of the forum visits.



Figure 2.2b. Forum page views from anonymous users. These figures exclude all logged in users (including administrators and moderators), between Jan. 2021 and Jan. 2022. Light blue represents anonymous users. Those analytics constitute an indirect measure of the forum visits from the wider public.

Additionally, in an attempt to maximise the likelihood of capturing only page views from non-registered users whilst minimising the page views from the forum staff, we select a window of observation that includes a period where staff is normally on leave. The underlying assumption is that registered users would normally log in to access the forum and that the forum administrators and moderators did not access the forum when on leave. Figure 2.2c shows some level of access

¹¹ Please note that the pageview measure counts every time a user lands on a page and factors in page reloading. One user’s visit could therefore result in multiple pageviews.

even during expected low traffic times which are likely to originate from the anonymous, wider members of the public.

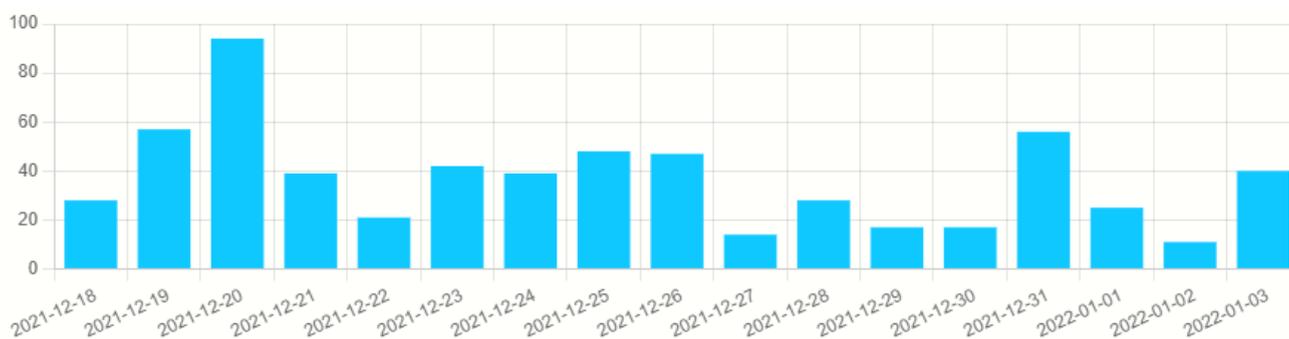


Figure 2.2c. Forum page views at low traffic. Page views from anonymous users during forum staff leave, between 2021-12-18 and 2022-01-03. These show the forum was still visited and those page views were unlikely to originate from the staff.

Looking at the forum content, Figure 2.3 is an example of a community-managed space. This is a space managed by HEIs aiming to open discussions about the specific needs and challenges of RDM competences and skills in research at doctoral level. Figure 2.4 shows an example of discussions between users. The forum has also provided users with *FAIR News*¹², a space for users to share and discuss news and blog posts relating to FAIR, and a *Feedback*¹³ area for suggestions on improving this space destined for communities.

Welcome to FAIR data competences in Universities forum

FAIR data competences in Universities



andre

3 May '21

May 2021

This forum is intended to be a place where participants have the opportunity to present and discuss the specific needs and challenges of RDM competences and skills at the doctoral level, as well to share experiences about the practical adoption in doctoral education programmes, such as doctoral education courses and doctoral schools.

Select one of the following categories in which you would like to share your contributions.

- RDM & FAIR data skills throughout educational portfolio - use cases ⁸
- Training programmes for researchers and professional staff - educational practices ⁷
- Practical guidance on the application of FAIR principles - good practices ⁵



1 / 3

May 2021

Figure 2.3 *FAIR data competencies in universities*¹⁴ is an example of a community space. This space is dedicated to HEIs and aims to open discussions about the specific needs and challenges of RDM competences and skills at the doctoral level.

¹² <https://fairdataforum.org/c/news/21>

¹³ <https://fairdataforum.org/c/site-feedback/2>

¹⁴ <https://fairdataforum.org/c/fairdatacomp/31>

FAIR and Closed Systems

User and Implementation Stories



millie

Hi all,

Does anyone have any advice or experience implementing FAIR principles in the context of closed systems i.e secure platforms, or secure research environments. To reach the highest levels of FAIR, it appears one must make data and metadata publicly available (for the most part).

In closed systems, it is not feasible nor practical to adopt many of the FAIR facets, as they are defined in much of the existing literature. Instead, can we think about interpreting the 'world' of a secure system, as the 'world in which the system lives', and 'FAIR usage' as FAIR usage within the bounds of the system?

Dec '21

Dec 2021

1 / 7

Dec 2021

29d ago

created	last reply	6	111	3	1			
 Dec '21	 29d	replies	views	users	link	 3	 3	



maaike

Hi Millie,

Thank you for opening this interesting topic!

One question for clarification, the secure research environment you mention, is this an environment used during research only or also for archiving and long term preservation?

Accessibility is indeed an important part of FAIR. In case your data cannot be publicly available, maybe you could look into publishing only the metadata?

Dec '21

1								
---	--	--	--	--	--	--	--	--

Figure 2.4 An example of user discussions about FAIR. FAIRdata Forum is a space for dissemination of expertise and knowledge.

In addition to providing a space for networking and access to the data stewardship training programmes, the FAIRdata Forum has been used for signposting and highlighting resources and activities that support data stewardship and areas of relevance for the data stewards' professional community, as well as training activities delivered in the project. Using the forum in this way provides a dynamic and relatively simple way to provide information and resources to a specific community and can bring related resources and activities together in context. For example, the forum has provided a space for discussion, frequently asked questions and information sharing relating to the [FAIR-Aware tool](https://fairaware.dans.knaw.nl/)¹⁵ from FAIRsFAIR WP4. The FAIR-Aware tool was also featured in the data stewardship instructor training programme. In this way, the FAIRdata Forum has provided a space where the linkages between these different activities and outputs can be realised.

¹⁵ FAIR-Aware <https://fairaware.dans.knaw.nl/>

FAIR-Aware tool: FAQ and User Discussion

FAIR-Aware is an online tool which helps to assess how much you know about the steps for making data(sets) findable, accessible, interoperable, and reusable (FAIR) before uploading them into a data repository. By answering 10 simple questions and reading the practical guidance, you will be ready to make your data FAIRer.

FAIRdata Forum

FAIRdata Forum supports the FAIR Competence Centre as an open space for users (individuals, institutions, organisations) to post queries, answers, comments and events notifications pertaining to making data FAIR. The FAIRdata Forum was initiated by the FAIRsFAIR "Fostering FAIR Data Practices in Europe" project which has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558.

FAIR-Aware tool: FAQ and User Discussion
all
Latest
Top

+ New Topic

FAIR-Aware events

This topic is created to inform about upcoming events related to [the FAIR-Aware tool](#).

Figure 2.5 FAIRdata forum hosts a space dedicated to FAIR-Aware, an online tool to assess users' knowledge about making data FAIR.

A close examination of the figures on FAIRdata Forum usage, browsing the content of its spaces (managed, closed and open) have brought to light some observations about the challenges in engaging with users. Most traffic was concentrated in correspondence with training events and remained low otherwise. Participants to training events were incentivised to access the Forum as it provided them with material and information pertaining to their training. Despite efforts in publicising the Forum both within and outside FAIRsFAIR, the traffic generated by other users remained low. So did the interactions between generic users and participants of the training events after the event. The volume of exchanges and discussions did not materialise at the levels that would have translated into a substantial Forum traffic. One cause was locality: users tend to stay within their domain-specific established communities. The tension between providing an approach that is domain-specific versus domain-agnostic played a role in the quite modest traffic observed on the Forum. Users seek information on FAIR matter primarily within spaces that belong to their own domain. A Forum that targets users from any communities regardless of their domains faces the challenge of gathering traction.

3. Competence Centre: Dissemination

3.1 Training Activities

A key activity and output of the Competence Centre has been the development and delivery of training and skills through a series of events and courses. FAIRsFAIR has supported the CODATA-RDA Research Data Science Summer Schools which deliver training in research data science skills to early career researchers predominantly from low and middle income countries. Three summer schools were supported during the course of the project (2019 in Trieste, Italy, in person, and 2020 and

2021 online) and will be reported on in the deliverable “D6.5 Report on three annual schools in core data skills for researchers”¹⁶.

Additionally, the concept of franchising schools through partnerships with local organisers was explored and will be reported in “D6.7 Report on schools run through franchising with local organisers”¹⁷.

A further strand of training was focused on Data Stewards Instructor training; a detailed report on the activities will be provided in “D6.6 Data Steward Instructor Training”¹⁸. The initial pilot event in Trieste (Italy) was delivered in partnership with the International Centre for Theoretical Physics (ICTP) and ran in parallel with the research data science summer school. After this event and the move to online delivery, the research data science and data steward events were split and run as separate activities. In total, nine instructor training events (see Table 3.1) were delivered and one further event is expected to take place at the time of submission of the present report:

Training event location	Year	Modality	Notes
CODATA Data Science School, ICTP, Trieste, Italy	2019	In-person	Part of the FAIRsFAIR Initiative, the CODATA-RDA Summer School had a strand dedicated to data stewardship and was held in Trieste, Italy in August 2019.
University of Manchester, UK	2020	Online	The University of Manchester is an advanced institution in terms of having well established RDM policies and support services in place. As such, the sessions focused more on building up awareness about the FAIR principles and responsible research and innovation than on introducing research data management and sharing concepts.
Universidad Nacional de Costa Rica, Costa Rica	2020	Online	This event included the standard course content but also included an additional session on developing and implementing data policies.
Botswana Open University, Botswana	2020	Online	FAIRsFAIR RDA-CODATA in collaboration with the Botswana Open University. This event (likewise the Costa Rica event above) was outside the geographic perimeter of the project activities but was part of international

¹⁶Bezuidenhout, Louise, & Shanahan, Hugh. (2022). D6.5 Report on three annual schools in core data skills for researchers. Zenodo. <https://doi.org/10.5281/zenodo.6074589>

¹⁷ Shanahan, Hugh, Davidson, Joy, & Newbold, Elizabeth. (2022). D6.7 Report on schools run through franchising with local organisers. Zenodo. <https://doi.org/10.5281/zenodo.6043907>

¹⁸ Shanahan, Hugh, Bezuidenhout, Louise, Davidson, Joy, Newbold, Elizabeth & Yates, Katie. (2022). D6.6 Data Steward Instructor Training. Zenodo. <https://doi.org/10.5281/zenodo.6074459>

			outreach. The workshop introduced participants to research data management (RDM), Open Science, and the FAIR Principles and showed how these can support open and responsible research practices. The workshop also provided recommendations for developing and implementing RDM policies and support services
Ghent University, Belgium	2021	Online	FAIRsFAIR, EOSC-Pillar in collaboration with Ghent University. This regional event covered Belgium Flanders
National Open Research Forum (NORF), Ireland	2021	Online	NORF/University College Cork. NORF Data Stewards Instruction Training. FAIRsFAIR, EOSC Synergy and Ireland's National Open Research Forum (NORF) delivered a three day instructor training workshop to support the development of data stewardship skills among staff in higher education institutions and other research performing organisations in Ireland. A key aim of the workshop was to empower a network of peers where best practices are exchanged and where those with more experience can share their knowledge with those just getting started
Gdańsk University of Technology, Poland	2021	Online	Gdańsk Data Stewards Instructors Training . This event was structured in the same fashion as NORF
Open Data Steward Instructor Training Workshop(no geographic restriction)	2021	Online	Two events (November, December). Different format: Open Data Steward instructors training. Those training sessions were open to worldwide participants with no restriction on location.
Teesside University, UK	2022	Online	This 2 hour introductory course will take place on February 28th and make use of a cut-down version of the Data Steward training materials.

Table 3.1 FAIRsFAIR - CODATA RDA collaboration: Data Steward Training events.

Five events were delivered as national or regional events working with a host organisation, and the final two (Open Data Steward Instructors Training Workshops) had open calls for attendees to allow us to reach a broader range of participants from across Europe and beyond. One additional event is planned for late February 2022 for Teesside University to pilot a very cut-down version of the standard programme. After the three training activities in 2020, a further five training events were

delivered during 2021 supported by the FAIRdata Forum as described above. Following feedback from participants, the spaces for the training events within the Forum were set up as open “public” groups rather than closed “private” groups so that the programme and logistics could be viewed without the need to register on the Forum prior to the event. During the Open Data Steward Instructors Training Workshops, supported by FAIRdata Forum, trainers have collected feedback from a sample of participants on the FAIR-Aware tool. Users were mainly employed in research and research support roles, mostly from universities and to a lesser extent research performing organisation (RPO) and research infrastructure / e-Infrastructure. The most representative domains were: social science & humanities, life and natural science. Overall, among the surveyed users these were the general trends:

- High levels of FAIR awareness
- Most users were likely to comply with FAIR
- They generally recognised that FAIR-Aware helped raise their awareness of FAIR principles

However, some topics remained challenging to many:

- community-endorsed standards
- provenance information
- controlled vocabularies
- digital curation and preservation.
- Interoperability and Reusability, perceived as difficult to attain.

After the delivery of the in-person data science summer school and pilot data steward event in Trieste in summer 2019, the impact of the COVID-19 pandemic required the training activities to be re-developed and moved to virtual online settings. Whilst the changes in the structure and delivery were necessitated by restrictions on travel and in-person activities due to the pandemic, the move to online provided an opportunity to extend the reach of the training in terms of the numbers of individuals trained and the geographical coverage achieved, for both the research data science summer schools and the data stewards events. A detailed breakdown of the extent of the training network is provided in Appendix 1.

Another benefit of the change of direction and the redesign of the training was the opportunity to collaborate on the delivery of the events with a range of different partners in a variety of models. The CODATA-RDA research data science summer schools have operated in collaboration with the ICTP since the inception of the schools, and as a result have benefitted from the infrastructure provided. Given the long standing relationship and support the CODATA-RDA research data science schools didn't use the FAIRdata Forum for supporting the schools. For the delivery of the online schools hosting the 2021 online summer schools on the established ICTP moodle site.

For the data stewards instructor training, FAIRsFAIR was able to draw on the relationships established within the INFRAEOSC-5b projects¹⁹ and collaborate with both EOSC Pillar and EOSC Synergy to address gaps in our content (e.g., pedagogy) and deliver the training events bringing different perspectives and expertise into the design and delivery. The partnership with EOSC

¹⁹ <https://www.eoscsecretariat.eu/communities/eosc-regional-projects>

Synergy resulted in a further collaboration opportunity and the development of a self-directed data steward online course as described below in section 3.

3.2 Data Stewards online course

3.2.1 Description

The data steward instructor training events described above were successful in terms of attendance numbers and range of institutions participating, but the demand for training exceeds what is practically possible to deliver in the timeframe of the project and resources available. Feedback showed that there is a continuing need for data stewardship training across Europe and more widely. With this in mind, it was decided that the development of a self-directed [online course](#)²⁰ would help to fill this demand and provide ongoing training opportunities following the end of the project. Rather than setting up an online learning platform/learning management system in the FAIRsFAIR project, we were able to leverage the connections with the EOSC Synergy project that had been established through our training activities and make use of their moodle platform to host the course. This not only reduced duplication of effort but also amplified the work of both projects and allowed for the development of a functional online course that guides learners through the resources in the correct order.

The majority of the training material has been reused from previous online data steward instructor training workshops that were delivered in 2020 and 2021. The purpose of these workshops was to support the development of data stewardship skills among staff in Higher Education Institutions (HEIs) and other research performing organisations.



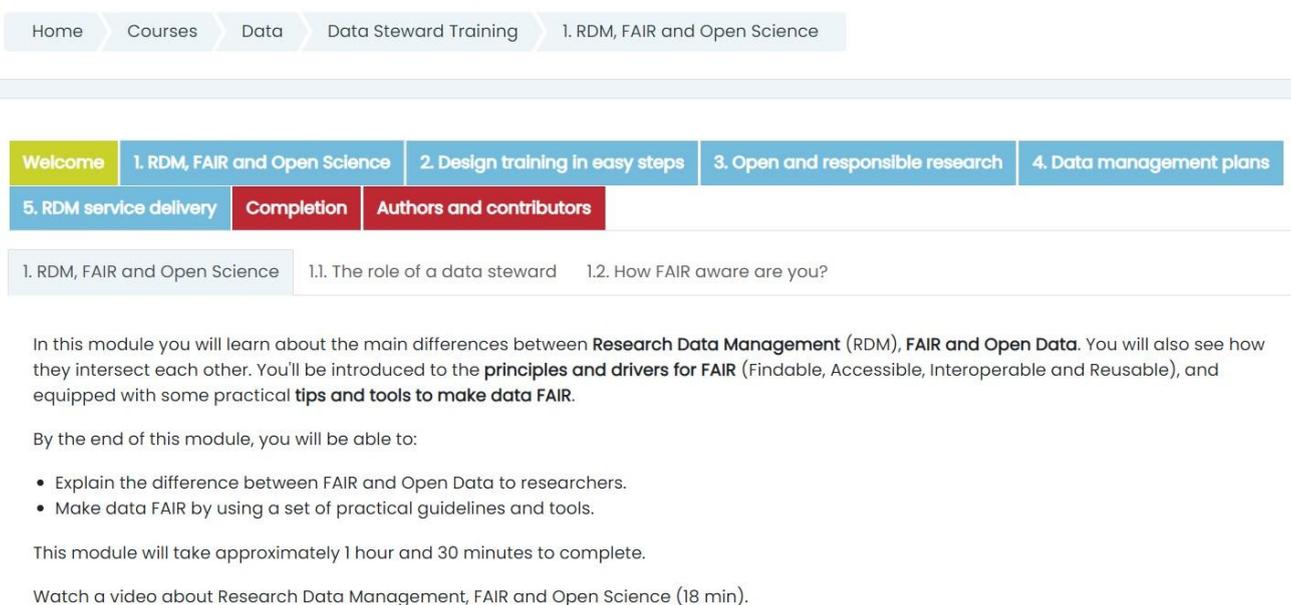
Figure 3.1 Data Steward online course modules.

²⁰ <https://moodle.learn.eosc-synergy.eu/course/view.php?id=132§ion=0#tabs-tree-start>

In this course, learners gain the skills and knowledge on how to be a successful data steward. Five self-paced modules (Figure 3.1) prepared by a team of international experts will guide them step by step through relevant topics for their daily work. By watching a series of recorded videos and completing assignments, learners will be up to speed with data stewardship skills to help them guide researchers.

All five modules are introductory. Learners can choose to follow them all or pick the ones that are most relevant to them. Each module takes on average one hour to complete, requiring around five to six hours to complete the entire course. Recorded video presentations, full transcripts, PowerPoint presentations and a number of learning activities and online resources are available to guide learning. Figure 3.2 illustrates an example of content.

Data Steward Training



The screenshot shows a breadcrumb trail: Home > Courses > Data > Data Steward Training > 1. RDM, FAIR and Open Science. Below this is a horizontal menu with five items: Welcome (yellow), 1. RDM, FAIR and Open Science (blue), 2. Design training in easy steps (blue), 3. Open and responsible research (blue), and 4. Data management plans (blue). Below the menu are two red buttons: 'Completion' and 'Authors and contributors'. A sub-menu for '1. RDM, FAIR and Open Science' is open, showing '1.1. The role of a data steward' and '1.2. How FAIR aware are you?'. The main content area contains text about the module's focus on RDM, FAIR, and Open Data, followed by learning objectives, completion time, and a video link.

Figure 3.2 Example of content from the Data Steward online course

3.2.2 Reusability considerations

In creating the online course the concept of reusability was considered and, where practical, was incorporated into course development. The course has been licensed as CC-BY and all materials deposited in [Zenodo](https://doi.org/10.5281/zenodo.6108288)²¹ for long term preservation, to make it citable and accessible beyond the life of the project. Materials are included in the training library (see 3.3) to aid discovery. If required the

²¹ <https://doi.org/10.5281/zenodo.6108288>

whole Data Steward course could be downloaded and reused within another learning platform, should someone wish to reuse the training materials within their own organisation.

The materials used to create the online Data Steward course were largely adapted from the Data Steward training events outlined in section 3.1 of this report (Training and skills: Data Science and Data Stewardship training events), with the exception of the DMP module which was reused from the [EOSC Bringing synergy to better data management and research in Europe](#)²² course. As such, all materials required adjustments and edits in order to fit within the context of a self-paced moodle course.

Editing course materials proved time consuming, particularly in relation to creating transcripts and correcting video closed captions, which might have been reduced had accessibility (in terms of disabilities) been considered from the start. There was also the need to consider whether it was better to re-use the presentation recordings from the National Open Research Forum (NORF) Data Steward training event (June 2021) or re-record each presentation. Had time allowed, it may have been better to re-record each presentation to allow the videos to be more concise and tailored to the context of the moodle course.

As already noted in D6.7 Report on schools run through franchising with local organisers, small teaching components are more easily adopted for re-use by instructors or trainers other than those who originally created the materials. This principle has been applied in the development of the moodle course and therefore each module has been separately uploaded to Zenodo as well as depositing an image of the the entire course in the native Moodle format. Following the creation of the Data Steward online course, options to widen interoperability and re-use were considered, for example exporting the course as a SCORM package²³ compatible with multiple e-learning platforms. However, it should be noted that the Moodle platform does not support exporting courses as SCORM packages. To convert the course into a SCORM package would have required the course to be re-built from the start using a SCORM conversion tool, which the project timeline did not allow.

3.3 Knowledge Base Development: Training Library

The approach adopted towards knowledge base developments detailed in D6.3 has been concentrated on issues relating to harmonisation as described below in Section 4. The initial work in D6.1 has shown a dynamic and fast-changing FAIR landscape and with it initiatives and projects whose remits presented some overlap. Accordingly, we aimed to avoid duplication and focused our work on areas where we can add value, as well as taking account of the developments in the EOSC Future project to provide a central Training and Skills knowledge hub²⁴, and how the materials produced within the project may feed into these developments. The development of the Training Library²⁵ as a simple knowledge base, has kept that aim in consideration while filling gaps in the

²² <https://moodle.learn.eosc-synergy.eu/course/view.php?id=7>

²³ <https://scorm.com/>

²⁴ [Training and Skills - Knowledge Hub - EOSC Future Public - Wiki EOSC Future](#)

²⁵ <https://www.fairsfair.eu/competence-centre/training-library>

provision of appropriate training material and creating synergies with other initiatives, including INFRAEOSC-5b projects.

The training library is a knowledge base of materials and resources that can be used to support training and development of FAIR related topics. It represents a collection and presentation of training materials developed across the FAIRSFair project and related activities including materials that have been explicitly developed for training as well as those that are available for re-use and inspiration but have not been explicitly designed for training. It includes guides, webinar recordings as well as training course materials. It is structured and searchable (by keywords derived where applicable from Terms4FAIRskills) and facets for audience (group) adapted from the EOSC actors in figure 4.3, material type, and format. All relevant information on the item is displayed in the search results which link directly to the external resource.

Training Library

[Home](#) / [Competence Centre](#) / [Training Library](#)

The training library contains records to **materials and resources** that can be used to **support training and the development of a harmonised understanding on FAIR** related topics.

Items in the library include:

- Materials that have been explicitly produced/developed to be reused and shared beyond the initial target audience e.g. train-the trainer related training materials.
- Materials that are available for re-use/inspiration but have not been explicitly designed for training or re-use beyond the initial purpose e.g. webinars.

The initial selection of materials has been drawn from the FAIRSFair project and partners and related FAIR activities.

OBJECT	DATE	MATERIAL TYPE	FORMAT
What are Persistent Identifiers and why to use them? The FAIR data principles call for the manged use of persistent identifiers. This short video explains what persistent identifiers are and why to use them. YouTube	16/02/2022	Guide	video
Module 8. Data management plans (quiz and activity) This quiz, learning activity and resources are a part of the introductory online course "Bringing synergy to better data management and research in Europe". The course helps learners get up to speed with innovative services and research data management approaches and do better research. This module was reused in the introductory course "Data Steward Training" as Module 4. Zenodo	20/09/2021	Course	pdf, docx
Module 5. Research data management service delivery This module is a part of the introductory online course "Data Steward Training". In this course learners will develop the skills and gain knowledge on how to be a successful data steward. By the end of this module, learners will: Understand different types of RDM services which their institution or organisation may need to provide. Be able to use the RISE model to plan and develop your RDM services. Zenodo	12/01/2022	Course	pptx, docx, video
Module 3. Open and responsible research This module is a part of the introductory online course "Data Steward Training". In this course learners will	12/01/2022	Course	pptx, docx, video

Keywords

Certification
 Costs
 Cross Domain Integration
 Data Documentation Initiative
 Data Management Plan
 Data Management Plans
 Data Science
 Data Stewardship
 Datasets

Target Group

Data Curator

Data Research Infrastructure Support Professional

Data Scientist/Data Analyst

Data Steward/Data Librarian

Educator/Trainer

EOSC Enabler

Policy Maker

Research Software Engineer

Researcher

Material Type

Case Study

Course

Exercise

Guide

Presentation slides

Presentation recording

Reference material

Workshop

Other

Format

pdf

pptx

docx

video

Other

Figure 3.3 Training Library: records of materials and resources to support training and a harmonised understanding of FAIR related topics.

There are a number of training catalogues within the EOSC community, in a fragmented landscape, where there is a need for better discoverability of and access to training resources to support FAIR data stewardship. With this aim in mind and the objective of the Competence Centre to work in collaboration to work on harmonisation activities, we have concentrated our knowledge base activities progression community approaches to the discovery, harmonisation and interoperability of training catalogues and resources which are discussed in the next section.

4. Competence Centre: Harmonisation

The third role of the FAIRsFAIR Competence Centre was to promote harmonisation of tools and techniques for information exchange about enabling FAIR data. This required collaborating with, and contributing to, several initiatives involved in implementing the vision set out by the EOSC Executive Working Group on Training and Skills in its report *Digital Skills for FAIR and Open Science*²⁶ and in particular, we worked closely with:

- A sub-group of RDA Interest Group on Education and Training in Handling Research Data (RDA-ETHRD-IG)²⁷. This focuses on defining minimal metadata for training resources, through plenary workshops and bi-weekly task group calls.
- Representatives of the ESFRI Clusters, EOSC-5 projects, and EOSC Future project, through two workshops²⁸ co-organised with the EOSC-5 Task Force on Training and Skills, one with the [OpenAIRE Community of Practice on Training Coordination](#)²⁹ and a fourth³⁰ with the EOSC Future project WP9.
- The [terms4FAIRskills initiative](#)³¹, through their participatory hack series [workshops](#)³² supported by FAIRsFAIR and by EOSC Co-creation funding. First, three core team hacks took place where terms4FAIRskills terminology was developed and then five groups hack sessions and a “annotate-a-hack” week session that included demos, discussion and review (Molloy, L. *et al*, 2021).

Much of this work was based on the notion that efforts to build competences and capacity for FAIR data stewardship will be supported by making the training and learning materials about this topic FAIR themselves. Doing so should address potential duplication of resources in producing the materials, and improve the discoverability of these materials across project websites and catalogues.

²⁶ <https://data.europa.eu/doi/10.2777/59065>

²⁷ <https://www.rd-alliance.org/groups/education-and-training-handling-research-data.html>

²⁸ <https://doi.org/10.5281/zenodo.4769468>

²⁹ <https://www.openaire.eu/cop-training>

³⁰ <https://doi.org/10.5281/zenodo.6225226>

³¹ <https://terms4fairskills.github.io/>

³² <https://doi.org/10.5281/zenodo.4772741>

Project activity in this area covered three main areas described further in this section:

- **Findability** - building consensus on minimal metadata terms for learning materials
- **Accessibility** - characterisation of learning resources and collections, and approaches to curating these
- **Interoperability** - exchange between EOSC project catalogues of metadata about materials, and terminology for describing learning outcomes of those materials

Reusability of learning materials was identified as a challenging area, requiring further work by RDA-ETHRD-IG on extended metadata to encapsulate the learning context. Nevertheless, the materials produced and delivered through the CODATA-RDA schools of research data science were reused by different instructors leading to some reflections on the challenges of reuse, as described in Section 3.2.2.

FAIRSFAR activity in this area was informed by previous work especially in the EOSC pilot project (Kühn and Streit, 2017³³; Whyte et al, 2019³⁴) and in life sciences (Garcia et al. 2020)³⁵.

4.1 Findability: consensus on minimal metadata for materials

Metadata for describing learning and training resources are already fully defined in a number of standards that have been developed in the Open Educational Resources (OER) community. Catalogues established for professional development in research data management and related areas have used a variety of OER standards, each identifying different metadata properties to describe training resources and learning materials. As the EOSC WG report *Digital Skills for FAIR and Open Science* (2021) noted, mapping between these standards will be essential if an EOSC federated learning and training catalogue is to be interoperable with these community-specific catalogues.

This situation parallels that of research data catalogues, and some options for canonical schema representation considered in that context (in FAIRSFAR WP3 for example) are also relevant here (e.g. DCAT-2, schema.org). There is a similar need to map each variant standard to a superset, in preference to mapping each variant to every other variant.

The problem is not unique to EOSC, and in the wider international context the RDA-ETHRD-IG addresses this issue of standards fragmentation by seeking to establish a minimum metadata application profile for learning and training resources. The approach is to identify those properties of learning and training resources that are considered necessary to fulfil requirements derived from user stories, expressing the discoverability needs of different groups of data management specialists and professionals. Starting from a broad list of terms, the IG narrows down to terms used by existing

³³Kühn, Eileen, & Streit, Achim. (2017). D7.2: Interim report and catalogue of EOSC skills training and educational materials (1.0). Zenodo. <https://doi.org/10.5281/zenodo.3395959>

³⁴Whyte, Angus, Leenarts, Ellen, De Vries, Jerry, Huigen, Frans, Kuehn, Eileen, Sipos, Gergely, Kalaitzi, Vasso, Dijk, Elly, Jones, Sarah, & Ashley, Kevin. (2019). D7.5: Strategy for Sustainable Development of Skills and Capabilities (1.1). Zenodo. <https://doi.org/10.5281/zenodo.5095052>

³⁵Garcia L, Batut B, Burke ML, Kuzak M, Psomopoulos F, Arcila R, et al. (2020) Ten simple rules for making training materials FAIR. *PLoS Comput Biol* 16(5): e1007854. <https://doi.org/10.1371/journal.pcbi.1007854>

catalogue services that describe learning and training resources, identifying commonalities and potential mappings to the OER standards used.

On this basis FAIRSFAR participated in bi-weekly meetings of a task group set up in the RDA-ETHRD IG, addressing ‘Minimal Metadata for Learning Resources’. Between June 2020 and December 2021 the group arrived at a consensus set of recommended elements³⁶. There are 14 elements (Table 4.1) which can be broadly seen to refer to descriptive information, access information and educational information.

The aim of the minimal set is that it can be used in a variety of settings to satisfy different use cases, the learner, the trainer and the service provider as well as accommodating the description of materials from different perspectives: formal education, professional development and informal education. In Section 4.3 below we return to the interoperability challenges in mapping this set of terms to those of existing EOSC catalogues.

Descriptive Information:	Access Information:	Educational Information:
<ul style="list-style-type: none"> ● Title ● Abstract/Description ● Author(s) ● Primary Language ● Keyword(s) ● Version Date 	<ul style="list-style-type: none"> ● URL ● Resource URL Type ● Access Cost ● License 	<ul style="list-style-type: none"> ● Target Group ● Expertise Level ● Learning Outcome(s) ● Learning Resource Type

Table 4.1 Minimal metadata for Learning resources.

4.1.1 Testbed of training materials/resources

FAIRSFAR set up a ‘testbed’ of materials to consider the applicability of the RDA-ETHRD-IG Minimal Metadata Set specifically in the EOSC context. The testbed was compiled cooperatively with contributions from across projects³⁷. The aim of the testbed was to:

- identify issues regarding harmonisation to inform future work e.g EOSC Future catalogue developments,
- explore how the minimal metadata set may be used to demonstrate improved discoverability, and
- determine whether further explanation, guidance and/or recommendation are warranted for different use cases based on analysis.

As the RDA minimal metadata recommendations are intended to be applicable to multiple use cases, the testbed contributions represented a cross-section of materials including formal and informal training materials from multiple sources such as training resource catalogues, learning

³⁶ [Recommendations for minimal metadata for learning resources](#)

³⁷ Contributions to the testbed from: Ni4OS, EOSC-Pillar, SSHOC, DMT Clearing House, ELIXR/TeSS, ENVRI-FAIR, OpenAire, FAIRSFAR, PaNoSC/ExPaNDS, EOSC Synergy

management systems and the Zenodo repository, as well as resources from websites and video sharing platforms.

The results of the testbed show that the application of the minimal metadata set would result in a richer description and greater consistency in the description of learning resources and training materials. However there are still some aspects that need further consideration especially in the EOSC landscape and context.

The main areas that need further work in the EOSC context relate to whether all the metadata elements should be required as specified by the RDA-ETHRD-IG, and guidance on the selection and use of controlled vocabularies.

Most of the metadata elements as defined by the RDA work are commonly used already across training catalogues to some degree, but not all are mandatory. The two notable exceptions where use was not common was Access Cost and Learning Outcomes. During a workshop³⁸ to discuss the testbed analysis in January 2022 there was a strong opinion that a) all the elements are necessary to describe training materials and learning resources and that this should be the desired goal but b) to require all elements to be mandatory could be a limiting factor in participation in the EOSC Future training catalogue.

4.2 Accessibility: characterising collections and their content

FAIRSFAR also participated in a further sub-group of RDA-ETHRD-IG to identify ‘Core characteristics of learning resource collecting and servicing organisations (such as catalogues and registries)’³⁹.

The work of this group looked at the evolving landscape of learning resources in research data management and identified the need for a set of core characteristics that describe these organisations and resources. This aims to aid users in understanding why these resources have been made accessible, It also aims to help organisations that are building services structure their learning resource collections for sustained access. The characteristics are not an exhaustive list and are intended as a guide covering five issues relating to: content, content descriptions, governance, services and operations.

It is not intended that all learning resource collecting and servicing organisations would be expected to apply all of the criteria as characteristics may or may not apply in a given situation. But the characteristics identified can also play a role in further work on “rules of participation⁴⁰” for training material providers in a federated EOSC environment.

³⁸ [Draft: Training resources metadata testbed - discussion - Google Docs](#)

³⁹ [Core Characteristics of Learning Resource Collectors](#)

⁴⁰ [RoP for Onboarding EOSC Training Resources and EOSC Marketplace - EOSC Future Public - Wiki EOSC Future](#)

4.3 Interoperability

4.3.1 Developing a metadata application profile to support semantic interoperability

The challenges in addressing semantic interoperability between catalogues of learning materials emerged from a series of workshops FAIRsFAIR co-organised with the INFRAEOSC-5b Training and Skills taskforce, to address the issues relating to harmonising of training resource catalogues and materials. The first workshop took place in October 2020, and focused on sharing information between projects to understand what short term actions could be taken to demonstrate interoperability between training catalogues and to inform further work. An overview was included in the D6.3 report, and this work fed into both the EOSC WG Training and Skills report, and the RDA ETHRD-IG minimal metadata work described in section 4.1 of the current report.

A second workshop in April 2021 focused on Harmonising Training Resource Metadata for EOSC Communities⁴¹. The goal of the workshop was to plan joint action by current EOSC projects to address recommendations in the EOSC Working Group report, and identify issues and decisions to be taken on priority areas. The discussion focused on minimal metadata for learning resources and fed into the ‘testbed’ work. The workshop considered the steps needed to enable interoperability between existing catalogues of training materials, and ranked these to identify the most challenging areas.

Participants identified metadata classification and exposure as the most challenging step, considering the diversity of approaches to this adopted by existing catalogues. For example, the ESFRI cluster project ENVRI-FAIR uses IEEE LOM (Learning Object Metadata)⁴², ELIXIR’s Training e-Support System (TeSS) uses the Bioschemas standard, while the US-based Data Management Training Clearing House⁴³ uses Learning Resource Metadata Initiative (LRMI)⁴⁴ from the Dublin Core Metadata Initiative (DCMI).

The ‘findability’ challenges around the variety of descriptive terms used in these standards were addressed over the course of the project through the collaboration with RDA ETHRD-IG. The interoperability issues similarly arise from the divergent standards, but relate to the semantic representation of the metadata and the syntax used to expose it in machine-actionable form. This would be needed both for exchanges between catalogues and for harvesting of metadata from them, to populate the EOSC training catalogue that was beginning to become a realistic prospect as the EC funded the EOSC Future project in 2021.

To inform this work, WP6 commissioned an expert report to consider options for creating a metadata application profile, to build on the ongoing collaboration in ETHRD-IG. The report briefly reviews abstract models for data about training materials, and the specification of their semantic aspects as standard schemas. It then suggests an approach to harmonisation, first through

⁴¹ Newbold, Elizabeth, Kayumbi, Gabin, Whyte, Angus, & Lazzeri, Emma. (2021). Summary Report: Workshop on Harmonising Training Resource Metadata for EOSC Communities. Zenodo. <https://doi.org/10.5281/zenodo.4769468>

⁴² <https://ieeexplore.ieee.org/document/9262118>

⁴³ <https://dmtclearinghouse.esipfed.org/>

⁴⁴ <https://dublincore.org/about/lrmi/>

establishing a minimal metadata application profile, and then by setting up a metadata store that would exploit linked data standards to make this metadata available in machine-actionable form⁴⁵.

The first step was to establish a metadata application profile, an approach defined as a “schema consisting of data elements drawn from one or more namespaces, combined together by implementers, and optimised for a particular local application”⁴⁶. This allows for some “mixing and matching” of terms from different namespaces.

This recommendation led to further work in the context of the RDA ETHRD-IG to engage with the Dublin Core community on the feasibility of using the Dublin Core Tabular Application Profile (DC TAP) format⁴⁷. The DC TAP is a useful format for developing such an application profile, as it allows spreadsheets to be used to document the properties used. It also allows for constraints to be specified, in the form of rules for value spaces such as the data type or class that may be used to provide a value for an element. Properties may be based on a broad vocabulary such as schema.org or Dublin Core Metadata Terms, and mapped to the more specialised schemas that EOSC catalogues employ (e.g. BioSchemas). This information can then be used to create more formal representations, for example in ShEx or SHACL, useful for validating whether the metadata held is sufficient. An example provided by the DC TAP Working Group is illustrated in Fig.4.1, describing a course using a schema derived from schema.org.

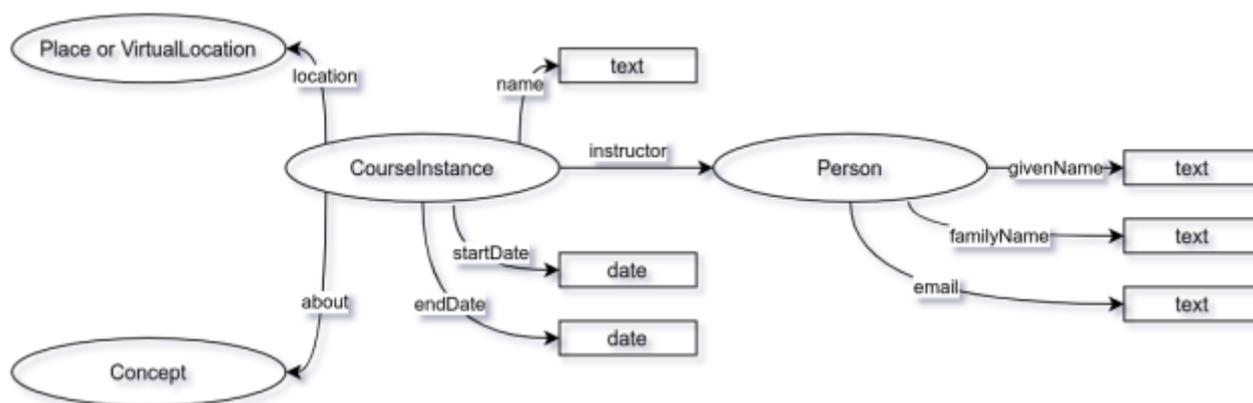


Figure 4.1: Example of course using a schema from schema.org.

Figure 4.2 shows a simple conceptual model followed by a TAP for Course Instances (offerings of a course) with specific start and end dates, at a virtual or physical location, about some concept, and with a person as the instructor.

⁴⁵ Barker, Phil, and Whyte, Angus. (2020). Harmonizing Metadata for Exchange of FAIR Training Materials (1.1). Zenodo. <https://doi.org/10.5281/zenodo.4434615>

⁴⁶ Heery, Rachel, and Manjula Patel. "Application profiles: mixing and matching metadata schemas." *Ariadne* 25 (2000) <http://www.ariadne.ac.uk/issue/25/app-profiles/>

⁴⁷ [GitHub - dcmi/dctap: DC Tabular Application Profile](#)

In the EOSC training landscape the DC TAP metadata application profile could help with the exchange of metadata among service providers, repositories and portals in this diverse landscape. The work is ongoing with an initial first draft⁴⁸ available. Following discussion within the EOSC projects we would recommend that this approach is taken forward by the EOSC Future WP9.

shapeID	propertyID	propertyLabel	mandatory	repeatable	valueNodeType	valueDatatype	valueShape	valueConstraint	Note
CourseInstance	rdf:type	instance of	y	n	IRI			sdo:CourseInstance	
	sdo:name	course name	y	n	Literal	xsd:string			
	sdo:startDate	start date	y	n	Literal	xsd:date			date must be in the future
	sdo:endDate	end date	y	n	Literal	xsd:date			must be after start date
	sdo:about	subject	y	y	IRI		Subject		
	sdo:instructor	primary instructor	y	n	IRI		Instructor		list only the instructor responsible for
Subject	sdo:location	location	y	y	IRI		Location		repeat only if a single course offering
	rdf:type	instance of	y	n	IRI			skos:Concept	
	rdf:type	instance of	y	n	IRI			sdo:Place,sdo:VirtualLocation	
Instructor	rdf:type	instance of	y	n	IRI			sdo:Person	
	sdo:givenName	given name	y	n	Literal	xsd:string			
	sdo:familyName	family name	n	n	Literal	xsd:string			
	sdo:email	email address	n	y	Literal	xsd:string			must be format [.*]@[.*].[*]

Figure 4.2 Tabular Application Profile example for a course⁴⁹

Further development of the metadata application profile needs to be taken forward in parallel with work to address further challenges identified by the community. These were discussed in a third workshop, co-organised with the OpenAIRE Community of Practice for Training Coordinators⁵⁰. The topics discussed included:

- Selecting controlled vocabularies
- Catalogue curation processes
- Sustainability of project catalogues

FAIRsFAIR pursued the first point in the RDA ETHRD-IG, as discussed below, although much further is needed on all three points and this is being taken forward in the EOSC Future project.

4.3.2 Controlled vocabularies for the metadata properties

The RDA-ETHRD-IG profile provides some examples of controlled vocabularies but these need to be determined by the context and landscape of the catalogue. In the context of EOSC and the development of the EOSC Future training catalogue guidance and recommendations would be beneficial. The issue however is complex as the training catalogues represent different domains and audiences and in some cases already use controlled vocabularies. Recommendations on controlled vocabularies is an area where there is a need for further community contribution and should be taken forward by the EOSC related projects.

⁴⁸ [rdaTest - Google Sheets](#)

⁴⁹ Phil Barker 2022 <https://blogs.pjkk.net/phil/dublin-core-tabular-application-profile>

⁵⁰ <https://www.openaire.eu/blogs/the-rdm-training-support-catalogue-landscape>

4.3.3 The Terms4FAIRskills terminology and browser

Liaising with the Terms4FAIRskills initiative⁵¹ has been a fruitful collaborative effort in the Competence Centre harmonisation role, yielding tangible outcomes and offering potential for further development. The Terms4FAIRskills terminology, and its associated annotation browser, are tools that can be used in annotating and searching for training materials related to the skills and knowledge needed to make and keep data FAIR. The development maturity achieved during the FAIRsFAIR project has brought these tools to a proof-of-concept stage, which has been useful for testing with the communities of stakeholders set out below. Ultimately, a mature version of the terminology could be integrated into a Knowledge Base-type environment to form a framework that allows well-described, open and interoperable training materials to be found and accessed. The browser demonstrates how the terminology can be deployed in practice to help a user effectively retrieve a specific training resource from a larger set of materials.

The Terms4FAIRskills (T4FS) terminology aims to provide a stable and community-accepted set of definitions for use across contexts to describe the skills required to make and keep data FAIR. A group of European experts first convened in January 2019 to discuss how to take forward this idea. It is clear that the aims of this initiative are confluent with the aims of the FAIRsFAIR project as a whole, and WP6 in particular. As such, WP6 colleagues have taken an active role in the development and testing of the current version of the terminology. A disbursement of funding from the EOSC Co-creation fund (2019-2020) also contributed to the production of a series of hack sessions during the life of the FAIRsFAIR project in order to test and develop the T4FS terminology to ensure that it is not overly shaped by the requirements of any particular research domain or sector, whilst retaining its focus on meeting the needs of European research.

The T4FS terminology can be applied to a variety of use cases by a range of potential actors. The completed terminology will be of use to trainers teaching FAIR data skills, researchers who wish to identify skill gaps in their teams, and managers who need to recruit individuals with specific skills and competencies, and to develop FAIR competency-based reward structures and training plans.

On a strategic level, these aims are confluent with the overall ambitions of the EOSC. *Digital Skills for FAIR and Open Science* (2021) identifies ten actors (including their roles and related skills) in the EOSC ecosystem for whom skills and training is relevant (see Figure 4.3).

The mature terminology will be an asset for these audiences in the use cases listed in the paragraph above, as EOSC implements its vision of a web of FAIR data and related services for science, making research data interoperable and machine actionable following the FAIR guiding principles.

⁵¹ <https://terms4fairskills.github.io/>

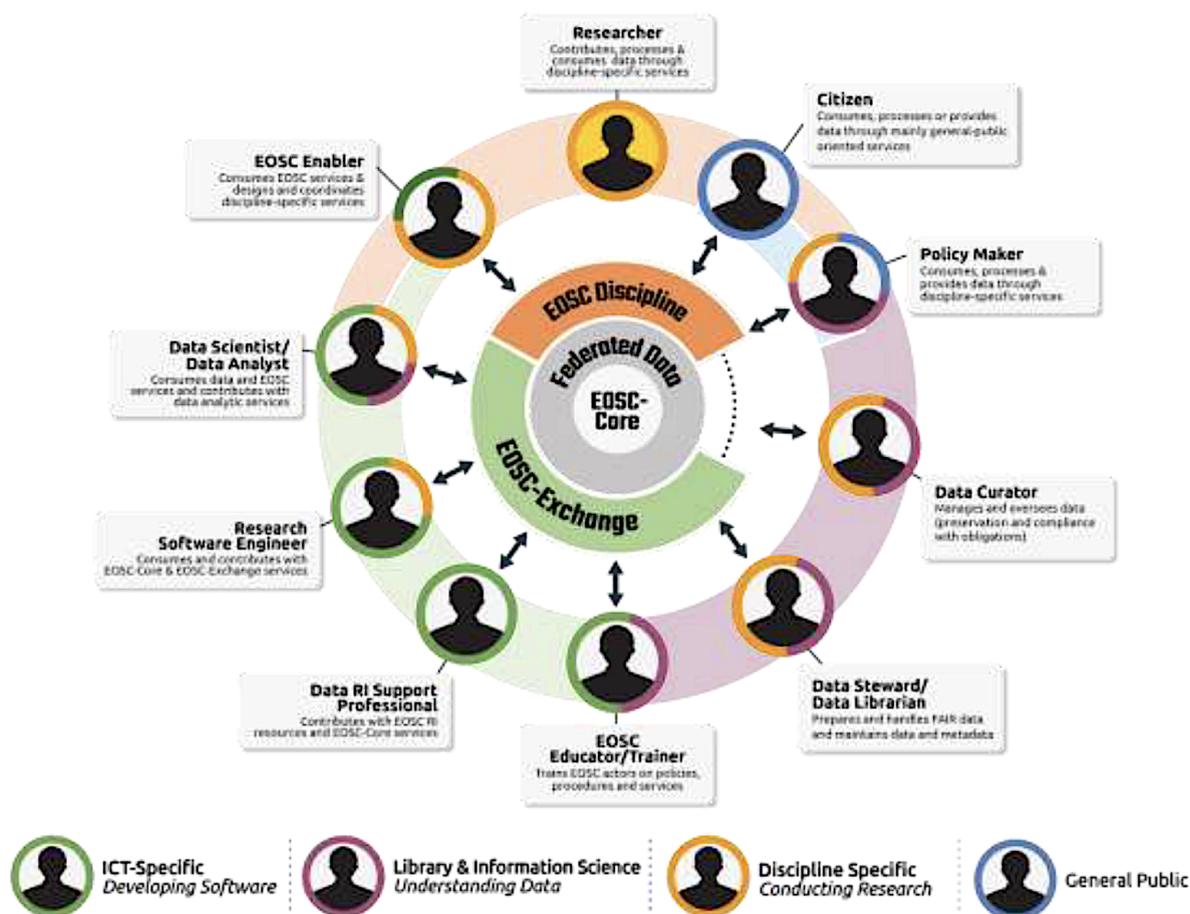


Figure 4.3: Ten EOSC actors. From Digital skills for FAIR and Open Science: report from the EOSC Executive Board Skills and Training Working Group, p. 17. Used under licence CC-BY.

Terms4FAIRskills is of potential use to all ten actors in their engagement with FAIR data, particularly:

- Educators/Trainers in the annotation of their training materials and for discovery of their training materials.
- Researchers in their skills development and, when mentoring or managing other researchers, in the creation and assessment of training plans.
- Data Curators, Data Stewards/Data Librarians, Data Scientists/Data Analysts in the annotation of their skills sets, the presentation of these skills in CVs or otherwise, and the assessment of their Continuing Professional Development (CPD) activities.

Collaboration between the T4FS initiative and FAIRsFAIR was important for testing the practical utility of the terminology. One of three key use cases in the development of the terminology during the life of the FAIRsFAIR project is the description of a range of training materials developed within

WP6 for early career researchers and data stewards (see section 2.2 and 3.2 above)⁵². Progress made on the development and testing of the terminology in this period is recapped in the terms4FAIRskills technical report⁵³.

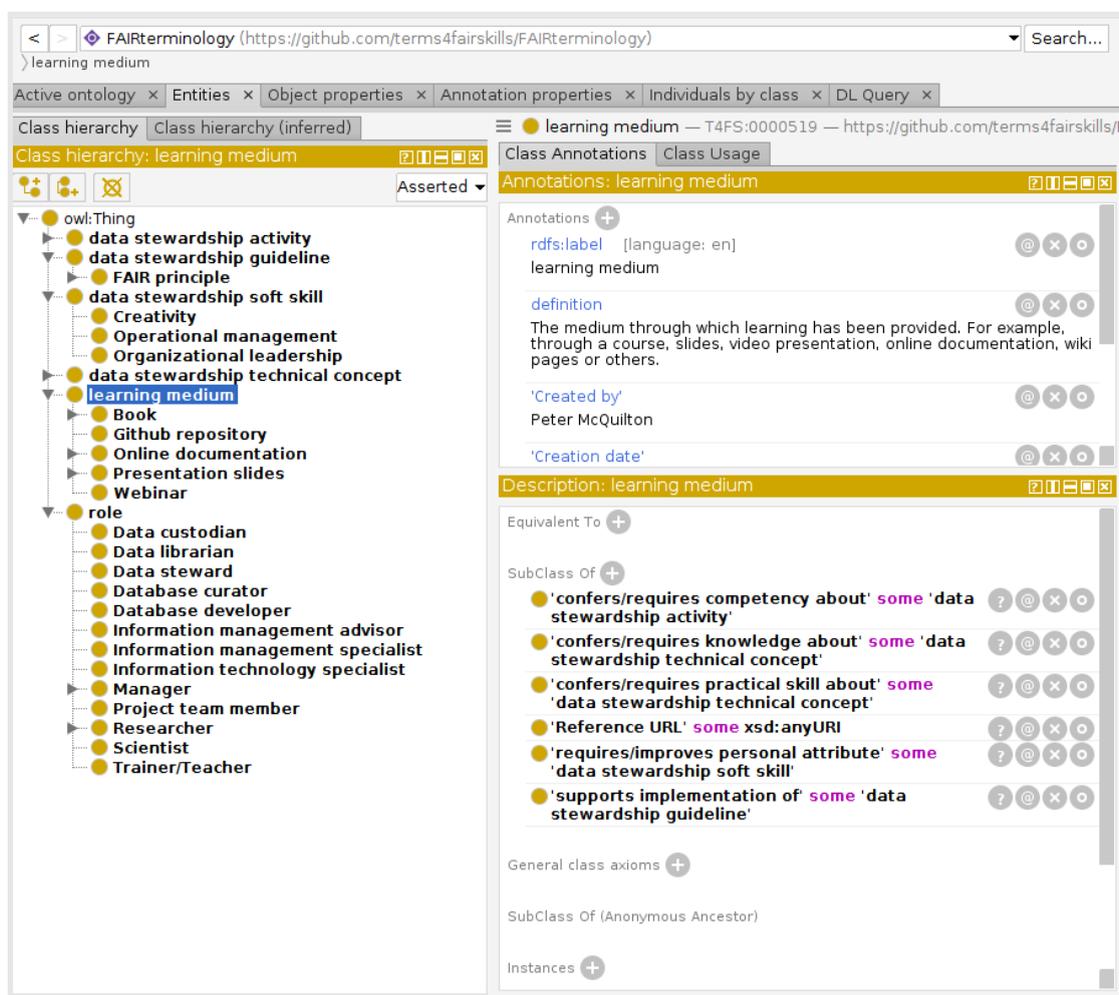


Figure 4.4: A view of the T4FS terminology being edited in the Github environment.

In short, the terminology was considered for use in the WP6 competence centre due to the close fit between the scope of the terminology and the scope of the materials held and described by the Competence Centre, but the sheer size of the terminology (currently around 600 terms) made it too large for effective use for the relatively small collection of objects in the Competence Centre

⁵² The other two use cases were the ELIXIR Training Platform Training and Events Portal (TeSS) and the EOSC-Pillar RDM Training and Support Catalogue

⁵³ Molloy, Laura, McQuilton, Peter, & Le Franc, Yann. (2021). EOSC Co-creation funded project 074: Delivery of a proof of concept for terms4FAIRskills: Technical report (1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.4772741>

environment. In practice, it would have been likely to suggest a large number of terms to the user for which there would be no corresponding objects to retrieve and would have required considerable development effort to deploy effectively in the Competence Centre environment as it is currently configured. It has instead been found useful as a controlled vocabulary for a description of assets deposited in the Competence Centre (see Figure 3.3 above), where terms have been selected from T4FS and made available as filters for the user to select, in order to find the objects for which they are looking.

The latest version of the terminology is openly [available on Github](#)⁵⁴ and can be forked and repurposed by other publicly-funded projects, thus demonstrating an example of a sustainable and reusable output from FAIRSFAR activity. As well as being a practical tool to help those developing and using training materials in the future, which in turn will support researchers and others in making and keeping data FAIR.

4.3.2.1 The T4FS Materials Browser

Further to the development and testing of the T4FS terminology, described above, and reflecting on our experiences using it as a controlled vocabulary for the Competence Centre training library, it became clear that there is demand for open-source tools that can interrogate or redeploy the T4FS terminology in new and useful ways. Accordingly, WP6 colleagues collaborated to develop a lightweight, proof-of-concept browser tool to enhance the process of retrieving training materials that have been annotated using the terms4FAIRskills terminology.

Unlike annotation tools such as Rightfield⁵⁵ and Semaphora, this browser tool does not carry out annotation of online materials; rather, it allows the user to retrieve materials that have already been annotated, and open documentation allows the tool to be redeployed in different contexts. The materials need to be accessible online via a URL or DOI. The browser uses only the annotations that have been applied to the online materials in the collection. The browser user enters keywords in the search field and matches are suggested.

The tool was developed using open-source software and is released publicly so that others can extend or develop it.

The WP6 team has made the initial [requirements document](#)⁵⁶, [API documentation](#)⁵⁷; and [front documentation](#)⁵⁸ available, serving as a further example of sustainable outputs from WP6 activity with high potential for reuse. Reuse of tools such as the materials browser is likely to promote use of the T4FS terminology, which in turn contributes to a harmonisation of terms used across contexts to describe the skills for making and keeping research data FAIR.

⁵⁴ <https://github.com/terms4fairskills/FAIRterminology>

⁵⁵ <https://rightfield.org.uk/>

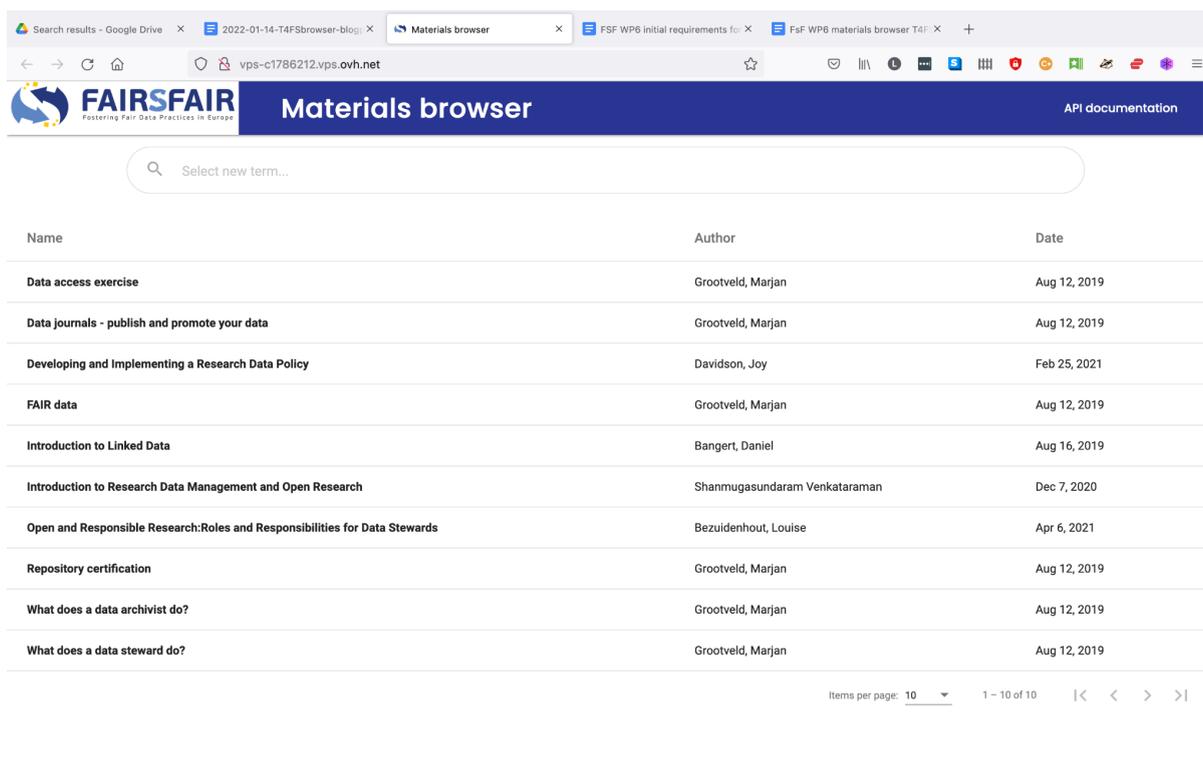
⁵⁶ https://docs.google.com/document/d/1RzjYLHsESYC91FViUNonOIBOdALe_sk6P2LhdJnzY2M/edit?usp=sharing

⁵⁷ Charly, M. (2021). materials-browser-front (Version 1.0) [Computer software]

<https://github.com/FAIRSFAR/materials-browser-api>

⁵⁸ Charly, M. (2021). materials-browser-api (Version 1.0) [Computer software]

<https://github.com/FAIRSFAR/materials-browser-front>



Name	Author	Date
Data access exercise	Grootveld, Marjan	Aug 12, 2019
Data journals - publish and promote your data	Grootveld, Marjan	Aug 12, 2019
Developing and Implementing a Research Data Policy	Davidson, Joy	Feb 25, 2021
FAIR data	Grootveld, Marjan	Aug 12, 2019
Introduction to Linked Data	Bangert, Daniel	Aug 16, 2019
Introduction to Research Data Management and Open Research	Shanmugasundaram Venkataraman	Dec 7, 2020
Open and Responsible Research: Roles and Responsibilities for Data Stewards	Bezuidenhout, Louise	Apr 6, 2021
Repository certification	Grootveld, Marjan	Aug 12, 2019
What does a data archivist do?	Grootveld, Marjan	Aug 12, 2019
What does a data steward do?	Grootveld, Marjan	Aug 12, 2019

Figure 4.5: A screenshot of the FAIRSFAR Materials Browser: an open-source tool to retrieve objects that have been annotated by the T4FS terminology.

5. Evaluation of minimal viable criteria for a competence centre

We have considered what would be an appropriate approach to evaluate the effectiveness of a competence centre. D6.2 set out the structures for a FAIR competence centre outlining a set of characteristics for a minimal viable criteria for a competence centre:

1. Defines the organisation's mission to support research data management or stewardship, including adoption and implementation of FAIR and open principles
2. Provides one or more services to support research data management or stewardship for identified users or audiences, user communities and provides metadata about the service to an EOSC catalogue
3. Identifies the scope of the competences it aims to develop, in fulfilling its mission
4. Documents the training or other methods it uses to develop competences
5. Applies a CC-BY or CC-0 license to its learning resources or other outputs
6. Makes learning resources metadata findable through a local and/or third-party catalogue or discoverable through third-party catalogues
7. Makes learning resources accessible under clear conditions, and described using community agreed terms

8. Advises its users on the assessment of FAIR data stewardship activities and outputs
9. Plans for the sustainability and continuity of its operations
10. Leads or participates in FAIR policy development

We have reviewed this set of characteristics for their continued relevance over the course of the last two years, drawing on experiences and developments in the FAIR landscape. The core characteristics provide a useful basis against which to evaluate the effectiveness of a competence centre but should not be expected to be a definitive list nor intended to be quantifiable. It is a starting point that can be adapted with the emphasis on different characteristics as appropriate to an institution or organisational setting. There are many activities and initiatives as outlined in the landscaping work of [D6.1](#)⁵⁹ and [M6.9](#)⁶⁰ that display these characteristics but do not define themselves as a competence centre and whilst these criteria have been defined in the EOSC context they could be applied more broadly.

Whilst most of the characteristics would appear to be of continuing relevance, some points may require modification or expansion.

Criteria 1 and 9: The mission of the competence centre can be expressed at different levels to reflect the global mission of the organisation, field or sector that it is supporting. Having a clear set of objectives that ascribed to an assigned mission provides a checkpoint against which the relevance of activities can be assessed. Sustainability is an important criteria for a competence centre to consider. However, as highlighted in the landscaping work and in the milestone 6.9, sustainability and continuity is a challenge when activities are bound to short-term project funding. In the FAIRsFAIR project, there is an activity entirely focused on sustainability and it has been embedded throughout the work. A suitable definition of the concept of sustainability needs to be clarified, as well as how it translates into practice. Thought should also be given to what is not sustainable or discontinuation of operations. Sustainability should be incorporated into the competence centre mission.

Criteria 6 and 7 can be viewed together as there are interdependencies between the activities that are required for implementation. The extent to which these criteria can be fully realised is in some cases reliant on the maturity of the community e.g. the availability of catalogues and the development of community agreed terms. As described above in previous sections the work package members have been contributing to community initiatives to develop and agree on a terminology through the T4FS and the minimal metadata recommendations. Therefore it is probably worth expanding the criteria and include that a competence centre should also be active in the development and maintenance of community's terms and standards.

Another criterion whose relevance came to light during the project was the establishment of formal and informal *collaborations*. A competence centre should adopt an approach that consists in liaising

⁵⁹ Herterich, Patricia, Davidson, Joy, Whyte, Angus, Molloy, Laura, Matthews, Brian, & Kayumbi Kabeya, Gabin. (2019). D6.1 Overview of needs for competence centres (1.0 DRAFT). FAIRsFAIR. <https://doi.org/10.5281/zenodo.3549791>

⁶⁰ Newbold, Elizabeth, & Kayumbi-Kabeya, Gabin. (2021). M6.9 Report on Final Review of Fair Competencies Landscape (V1.0). Zenodo. <https://doi.org/10.5281/zenodo.5708089>

with projects and initiatives sharing similar or presenting overlapping scopes. That should not be limited to a networking exercise but translate into actively developing synergies between existing and new partners. Whilst this has been implicit in the work of FAIRsFAIR and other EOSC projects, it was not explicitly stated as a criteria for a minimal viable competence centre structure.

Finally, the aforementioned review of the set of characteristics for a minimal viable competence centre has led us to the following refined set:

1. Defines the organisation's mission to support research data management or stewardship, including adoption and implementation of FAIR and open principles; *plans for sustainability and continuity of operations should be considered when defining the organisation's mission.*
2. Provides one or more services to support research data management or stewardship for identified users or audiences, user communities and provides metadata about the service to an EOSC catalogue
3. Identifies the scope of the competences it aims to develop, in fulfilling its mission
4. Documents the training or other methods it uses to develop competences
5. Applies a CC-BY or CC-0 license to its learning resources or other outputs
6. Makes learning resources metadata findable through a local and/or third-party catalogue or discoverable through third-party catalogues
7. Makes learning resources accessible under clear conditions, and described using community agreed terms; *participates in and contributes to the development and maintenance of community agreed terms*
8. Advises its users on the assessment of FAIR data stewardship activities and outputs
9. Leads or participates in FAIR policy development
10. *Establishes and develops collaborations liaising with projects and initiatives sharing similar or present overlapping scopes.*

6. Key Messages and Recommendations

This final section presents a series of key points derived from the work conducted by WP6 during the lifetime of the FAIRsFAIR project. It is the fruit of the work within WP6 as defined in its initial scope but also draws upon the experience of interactions with external stakeholders and opportunities for collaborations that developed as the project progressed. The following points include recommendations for future work including work with communities.

6.1 Collaboration

The establishment of cross-project collaboration with the INFRAEOSC-5b projects through a formal agreement and the EOSC-5 Task Force on Training and Skills was an effective mechanism to develop relationships and establish synergies. It provided a focus for activities and led to concrete collaborations as outlined in previous sections. Additional collaborations (without formal agreements) with the ESFRI cluster projects, the EOSC Future project, the Open Aire Community of

Practice of Training Providers, and the terms4FAIRskills initiative as well as the participation in RDA-ETHRD-IG were also very successful in aligning activities and providing routes for the promotion of harmonisation and identifying opportunities for synergies. Through these different collaborations there has been demonstrable progress on issues related to the findability, accessibility and interoperability of training resources and training catalogues. Establishing formal agreements of collaboration between stakeholders, creates an environment for cross-project working and incentivises actions that could lead to further initiatives.

As the INFRAEOSC-5b projects near completion and new ones are initiated it is recommended that the structures that support collaboration e.g. the task force approach or communities of practice are either maintained or established.

6.2 Community building

A particular focus has been on supporting data stewardship training and aimed at those communities that are not already well served this has translated into a focus on data stewards and other professionals mainly in Higher Education Institutes and other Research Performing Organisations who are new in roles; the audience for the training was broad in the spectrum of roles, skills and provenance. Building communities takes time and effort and needs infrastructure support in terms of both human effort and mechanisms for networking (both online and offline). Setting up a framework that defines the modus operandi for communities to embrace a FAIR domain-agnostic approach in FAIR is challenging. The tension between on the one hand a pragmatic domain-specific approach carrying the inherent advantage of being close to already established community networks, and on the other hand a general approach leading to a coherent global set of policies and practice applicable across the board, was palpable in the course of interactions we had with communities during this project. The difficulties for the FAIRdata forum to get substantial discussion was testament to such challenges. Feedback from the participants in the data stewards training would imply that communities of practice are important and are being developed taking multiple forms: informal networks built on in personal relationships and formal networks at national levels. The most common approach and one that seemed to be of most interest in developing were those on a regional or national level.

In the provision of spaces for discussions and exchange between users and experts, locality and domain-specificity also should be considered in building communities.

Communication should be central to a competence centre action. Projects should envisage the development of a specialist role whose remit would encompass community engagement and facilitate transversal linkage between stakeholders, internal and external.

6.3 FAIR training materials

There has been significant progress in the areas of applying FAIR principles to training materials, but there are many strands to this work which still need to be progressed. Much of the work has developed through volunteer effort or support from projects through effort in kind. To maintain the

progress the outputs of the community efforts need to be adopted more widely beyond those directly involved.

The work on a recommended set of minimal metadata is a significant first step, the next step is to see how this is adopted in real world situations. The next stage in the development of the work for the RDA-ETHRD-IG is to consider the extended metadata and documentation which will be vital in filling in the gaps and possible limitations of the minimal set. As mentioned in section 6.4, recommendations and *guidance on controlled vocabularies is a key area of activity that should be continued*. To progress interoperability of training catalogues we *recommend developing DC TAB Metadata Application profile for use in the EOSC context*.

The work by the community on the development of the terms4FAIRskills terminology and related applications demonstrates the need for such a terminology. There is an ongoing requirement to maintain and develop the terminology which has so far been supported in the main from volunteer effort, in kind support and small scale project funding. In order to progress the work there is a need for adoption of the terminology and take up in the wider community. Further to the development and testing of the terms4FAIRskills terminology, it appears that there is demand for open source tools which can interrogate or redeploy the terminology in new ways.

In the scope of the Data Stewards Instructor training we have demonstrated that modules can be re-used successfully by instructors other than those who created them and the course curriculum translates across audiences. Working collaboratively with EOSC Synergy has provided an example to follow in future. It has allowed for the reuse of existing course templates, structures and materials, which resulted in a more efficient development process. Achieving that requires effort and commitment by both those creating the courses and those delivering as an element of customisation for the audience is often required, for example providing local context and it requires that materials are made openly available and discoverable.

Training courses should be planned in advance and all potential re-uses of the materials considered from the outset, particularly where the same materials are to be used for both instructor-led and self-directed learning. This should include the creation of a proposal and a course development plan outlining planned topics, learning outcomes, assessments and activities. By visualising the course structure and uses in this way with sufficient forward-planning, there are significant efficiency gains to be made in respect of time, resources and the avoidance of duplication of effort. For recorded materials, professional editing and the inclusion of opening and closing credits can add a professional and user friendly finish to videos.

We recommend that the initial phase of planning courses should systematically include the identification of potential re-uses of the materials considered. That can be achieved through a development plan that outlines the course structure and uses.

6.4 Delivery of training and skills

The CODATA-RDA Schools of Data Science and the FAIRsFAIR data stewards instructor training events provided useful models for skills and training development that could be adopted by others. Due to

necessity, the format of both the data stewardship and CODATA-RDA schools of data science were changed and redeveloped as virtual online events providing an opportunity to explore new models of delivery. Similar developments were experienced in the SSHOC train-the-trainer bootcamps⁶¹ which FAIRSFair contributed to.

Virtual online events have proved to be a useful model to engage with a geographically dispersed audience lowering the barriers to participation, with reduced costs (travel), scheduling flexibility and time commitments. However, elements of community building and informal knowledge sharing are not so easily replicated in virtual events, and therefore whilst it is attractive to maintain the virtual models further work is needed to look at hybrid models consisting of a mixture of in-person and online training activities. Due to the circumstances with restriction on travel and in person meetings it was not possible to develop these models during the project.

We recommend further exploration and experiments on hybrid models of training delivery in future projects to extend the knowledge on suitable formats.

Feedback from the participants on the data stewards training, via the [FAIRSFair National Roadshows](#)⁶² and the landscape activities, show there is an ongoing demand for training and skills development for data stewards, the training provided within the scope of the FAIRSFair project has not been domain specific, whilst we envisage that there will always be demand for this as new people enter the profession, there is also an increasing demand for domain focused activities and more advanced and specialist training which will increase as the cohort of new data stewards progress in their careers. Alongside this there is an additional need for training in national languages and with local context (regional or sector specific).

In addition to the necessary domain-agnostic data steward training, we also recommend the provision of a contextualised training. This latter will bring trainees closer to their professional fields, and cultural environment. Such a dual approach will develop synergies within specific communities but also beyond, across domains.

In respect to the provision of training, further work on alternative funding models are required in order to sustain and develop a skilled workforce. Although this topic has been explored in the report D6.7 and is being taken forward in the context of the CODATA-RDA data science summer schools, *models of funding for and sustainability of training and skills programmes is an area that would benefit from further work within the EOSC landscape.*

⁶¹ Darja Fišer, Erzsébet Tóth-Czifra, Judith Wehmeyer, Veronika Keck, Ellen Leenarts, Ricarda Braukmann, Tatsiana Yankelevich, & Cristina Magder. (2021). D6.12 Report on the SSHOC train-the-trainer bootcamps. <https://doi.org/10.5281/zenodo.5734301>

⁶² <https://www.fairsfair.eu/events/fairsfair-national-roadshow-series>

7. Conclusion

This report presented the different stages that characterised WP6's work in progressing towards the main objective of support for FAIR uptake through the provision of training on FAIR and related activities. By placing community engagement at the centre of its approach in the development of concepts and their implementation, this report showed how the FAIR Competence Centre shaped its actions along its assigned trinary role. Firstly, *Advisory*, by providing the online platform FAIRdata Forum to support communities and Data Steward Training events; secondly, the *Dissemination of Training & Skills*, including the provision of an example of knowledge base development with a FAIR Training Library; thirdly, *Harmonisation* in building consensus among internal and external partners whilst addressing issues of findability of learning materials, their accessibility through characterisation and curation, and interoperability, looking at EOSC project catalogues and terminology. WP6's present and previous reports showed how throughout the project, competence centre concepts have evolved and their implementation shaped by community engagement in a dynamic FAIR landscape. Following up on criteria to characterise a competence centre as previously presented in D6.2, and based on WP6's experience, the present report offered reflections on how those criteria could be amended to best suit their context.

Finally, considering the multiple facets pertaining to WP6's work, among these its objective, the approach in addressing it, the challenges encountered, the engagement with communities and FAIR stakeholders, this report extracted a synthesis of key messages and propounded recommendations on *Collaboration, Community Building, Training Material, and Training & Skills delivery*.

8. Bibliography

- Barker, Phil, and Whyte, Angus. (2020). Harmonizing Metadata for Exchange of FAIR Training Materials (1.1). Zenodo. <https://doi.org/10.5281/zenodo.4434615>
- Bezuidenhout, Louise, & Shanahan, Hugh. (2022). D6.5 Report on three annual schools in core data skills for researchers. Zenodo. <https://doi.org/10.5281/zenodo.6074589>
- European Commission, Directorate-General for Research and Innovation (2021). Digital skills for FAIR and Open Science: report from the EOSC Executive Board Skills and Training Working Group, (E,Lazzeri,editor, M,Barker,editor, N,Manola, editor, V,Gaillard,editor, I,Kuchma,editor, L,Stoy,editor) Publications Office. <https://data.europa.eu/doi/10.2777/59065>
- Garcia L, Batut B, Burke ML, Kuzak M, Psomopoulos F, Arcila R, et al. (2020) Ten simple rules for making training materials FAIR. PLoS Comput Biol 16(5): e1007854. <https://doi.org/10.1371/journal.pcbi.1007854>
- Heery, Rachel, and Manjula Patel. (2000). Application profiles: mixing and matching metadata schemas. Ariadne 25. <http://www.ariadne.ac.uk/issue/25/app-profiles/>
- Herterich, Patricia, Davidson, Joy, Whyte, Angus, Molloy, Laura, Matthews, Brian, & Kayumbi Kabeya, Gabin. (2019). D6.1 Overview of needs for competence centres (1.0 DRAFT). FAIRSFAR. <https://doi.org/10.5281/zenodo.3549791>
- Newbold, Elizabeth, & Kayumbi-Kabeya, Gabin. (2021). M6.9 Report on Final Review of Fair Competencies Landscape (V1.0). Zenodo. <https://doi.org/10.5281/zenodo.5708089>
- Newbold, Elizabeth, Kayumbi Kabeya, Gabin, Matthews, Brian, Davidson, Joy, Herterich, Patricia, Whyte, Angus, & Molloy, Laura. (2020). D6.2 Initial Core Competence Centre Structures (1.0 DRAFT). Zenodo. <https://doi.org/10.5281/zenodo.3732889>
- Newbold, Elizabeth, Kayumbi, Gabin, Whyte, Angus, & Lazzeri, Emma. (2021). Summary Report: Workshop on Harmonising Training Resource Metadata for EOSC Communities. Zenodo. <https://doi.org/10.5281/zenodo.4769468>
- Kayumbi-Kabeya, Gabin, Newbold, Elizabeth, Whyte, Angus, Cepinskas, Linas, & Molloy, Laura. (2021). D6.3 Established Competence Centre for Variety of Communities (1.0 Draft). Zenodo. <https://doi.org/10.5281/zenodo.4560474>
- Kühn, Eileen, & Streit, Achim. (2017). D7.2: Interim report and catalogue of EOSC skills training and educational materials (1.0). Zenodo. <https://doi.org/10.5281/zenodo.3395959>
- Charly, M. (2021). materials-browser-front (Version 1.0) [Computer software].

<https://github.com/FAIRsFAIR/materials-browser-front>

Charly, M. (2021). materials-browser-api (Version 1.0) [Computer software].

<https://github.com/FAIRsFAIR/materials-browser-api>

Molloy, Laura, McQuilton, Peter, & Le Franc, Yann. (2021). EOSC Co-creation funded project 074: Delivery of a proof of concept for terms4FAIRskills: Technical report (1.0.0). Zenodo.

<https://doi.org/10.5281/zenodo.4772741>

Whyte, Angus, Leenarts, Ellen, De Vries, Jerry, Huigen, Frans, Kuehn, Eileen, Sipos, Gergely, Kalaitzi, Vasso, Dijk, Elly, Jones, Sarah, & Ashley, Kevin. (2019). D7.5: Strategy for Sustainable Development of Skills and Capabilities (1.1). Zenodo.

<https://doi.org/10.5281/zenodo.5095052>

Appendix 1 Scale of the trainers network

The following details the scale of the trainers network resulting from the training activities for both data science and data stewardship over the course of the project (and constitutes the content of the milestone report M6.7).

A.1. Core Skills in Data Science Summer Schools

The core skills in Research Data Science Summer Schools activities are outlined in the report D6.5 Report on Three Annual Schools in Core Data Skills for Researchers⁶³. The summer schools are aimed at early career researchers and provide practical introduction to the topics and have focused on participants from Low and Middle Income Countries. The following table details the scale of the training activity over the three years. The Schools have been run since 2016, FAIRSFAR provided support for the schools during 2019, 2020 and 2021. With the support of FAIRSFAR and the move to online delivery the geographic scope for participation in the 2021 Schools was able to be expanded to include participation from European countries, previous events had been targeted specifically to low and middle income countries.

	Number of attendees	Number of Countries Represented
2019 (in person)	38	18
2020 (online)	123	42
2021 (online)	76 (8 EU, 5 UK)	32 (6 EU countries and UK)
Total	243	56 (unique countries)

Table A.1 CODATA-RDA School of Research Data Science Summer School participation

A.2 Data Stewards Instructor Training

The activities of the data stewards instructor training are outlined in detail in D6.6 Data Steward Instructor Training⁶⁴. The following table details the scale of the training activity over the three years. A total of nine instructor training events were held, due to the Covid-19 pandemic and restrictions on travel only one event was held in person (Trieste, 2019) all other events were held online. The move to online enabled a greater number of individuals to participate in the training and increased the reach of the training network. Online events enabled FAIRSFAR to engage with institutions and individuals outside of the core European stakeholder group demonstrating the international interest in data stewardship instructor training. Participation in the data stewards event was through registration, events held in collaboration with a regional partner were restricted

⁶³ <https://doi.org/10.5281/zenodo.6074589>

⁶⁴ <https://doi.org/10.5281/zenodo.6074458>

to those regions, the “open” events were not restricted geographically and the two events were staggered with one held during the morning and the other during the afternoon to accommodate different time zones. The open events attracted participants from across Europe but also from South American countries. Whilst the majority of attendees were from higher education institutions other research performing organisations were represented in both the open events and the regionally focused events. For example, at the NORF event, there were 10 related institutions including the Digital Repository of Ireland, the Marine Institute and the Irish Centre for High-End Computing and at the two open events there were 20 related institutions including the Portuguese Institute of Oncology of Porto, the Italian Institute of Technology and the Slovak Centre of Scientific and Technical Information.

As highlighted in Section 3 an additional data stewards session is due to be delivered by the DCC at Teesside University at the end of February. This is an additional event and the final details were not available for inclusion in this report.

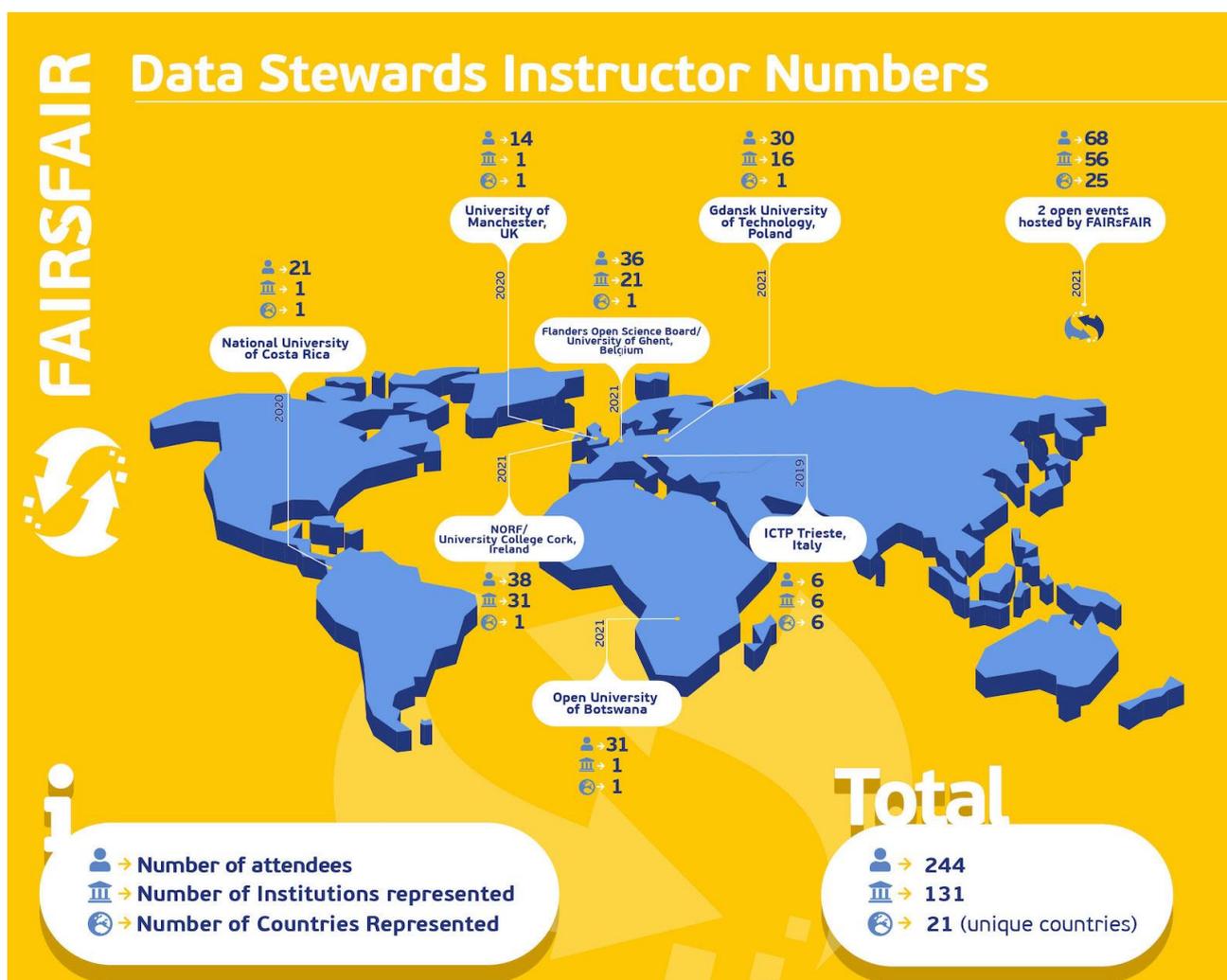


Figure A.1: Data Stewards Instructor Training: participation numbers.

A.3. Collaboration

Collaboration has been central to the approach of the work undertaken in WP6 including the delivery of training. The CODATA-RDA data science summer schools already demonstrated the value of collaboration and these principles were applied when developing the data stewardship activities. In addition to the skills of the instructors and institutions represented in the FAIRsFAIR project, the data stewards instructor training events drew on expertise from INFRAEOSC-5b projects (EOSC Synergy and EOSC Pillar) as well as collaboration with Universities and national bodies. The table below shows the organisations (external to the FAIRsFAIR project) who contributed to the training activities alongside their role in the delivery of the training.

Role	Collaborating organisation
Event co-host	<ul style="list-style-type: none"> ● International Centre for Theoretical Physics, Italy ● University of Manchester, UK ● National University of Costa Rica, Costa Rica ● University Library of Botswana, Botswana ● University of Ghent, Belgium ● University College Cork, Ireland ● Gdansk University of Technology, Poland
Supporting organisations	<ul style="list-style-type: none"> ● Flemish Open Science Board, Belgium ● National Open Research Forum, Ireland
Course content and delivery	<ul style="list-style-type: none"> ● EOSC Pillar ● EOSC Synergy

Table A.2 Collaborations with external organisations.

A.4 Promoting the training events

The instructor training events were promoted to local audiences using the host institutions' networks as well as dedicated space on the FAIRsFAIR website and communication channels. In addition the open calls were promoted to attendees of the FAIRsFAIR National Roadshow series.

A.5 Blog posts

Blog posts produced by participants and collaborators:

- [Eurodoc participated in the 1st Data Steward School](#)⁶⁵ by Sothearath Seang, Treasurer of Eurodoc
- [Data Stewardship Training - Manchester](#)⁶⁶ by Eleanor Warren, Research Services Librarian, University of Manchester
- [Data Stewardship Training - Botswana](#)⁶⁷ by Naniki Maphakwane, Botswana Open University Library and Information Services
- [EOSC Synergy helps boost confidence of trainers](#)⁶⁸ by Linas Cepinskas, Research Data Management Specialist, DANS
- [Data Stewardship Training - Belgium](#)⁶⁹ by Paula Oset Ghent University
- [NORF collaborates with FAIRsFAIR and EOSC Synergy on Data Stewardship Training](#)⁷⁰ by Dr Daniel Bangert (National Open Research Coordinator) and Dr Aiife Coffey, UCC (NORF FAIR Working Group Co -Chair)
- [A Focus on...the emerging role of 'Data Steward in Ireland'](#)⁷¹ by Dr Deborah Thorpe at the Digital Repository Ireland,
- [FAIRsFAIR-CODATA-RDA Sata Steward Training Series - Poland](#)⁷² by Katarzyna Dudek and Piotr Krajewski

A.6 Feedback

Feedback from the participants was taken into account in the planning of the sessions and changes to timings and structure were incorporated. Changes were also made on the suggestion of the local hosts. Much positive feedback was received across all the schools and Data Steward events. A selection of these are highlighted below.

Attendees, Gdańsk University

"It is a great opportunity to get knowledge and skills about teaching others how to be a data steward."

"There are too few of such training!"

⁶⁵ <http://www.eurodoc.net/news/2019/eurodoc-participated-in-the-1st-data-steward-school>

⁶⁶ <https://fairsfair.eu/articles-publications/data-stewardship-training-manchester>

⁶⁷ <https://fairsfair.eu/articles-publications/data-stewardship-training-botswana>

⁶⁸ <https://www.eosc-synergy.eu/eosc-synergy-helps-boost-confidence-of-trainers/>

⁶⁹ <https://fairsfair.eu/articles-publications/data-stewardship-training-belgium>

⁷⁰ <https://norf.ie/index.php/2021/09/10/norf-data-stewardship-training/>

⁷¹ <https://dri.ie/blog-data-stewardship>

⁷² <https://www.fairsfair.eu/articles-publications/fairsfair-codata-rda-data-steward-training-series-poland>

“The idea of sharing knowledge and experience regarding data stewards training is very important for our University.”

Attendees, University of Costa Rica

“It is extremely important to clarify concepts, and see the practical usefulness of the concepts in our reality as a university.”

“With this training It was evident that some of the tasks that we have been doing with the repositories are related to the tasks of a Data Stewardship. We can work on some more things from this area.”

Attendees, NORF/University College Cork

“There is a critical lack of formal data stewardship services. I have seen it first-hand in the project I have been involved with, but the case studies highlighted to me that this was an even more general problem than I assumed. Even if full-time stewards are unlikely to happen in the near future, training and part-time stewardship would go a long way to help our research community in publishing/maintaining their data products, which are fast becoming an alternative to papers as an expected research output.”

“We need lots more events like this if support for Irish research is to keep up with international developments in the area of Data Stewardship.”