

# RDM Training & Support catalogue landscape

Open Science FAIR, 21 September 2021, 9.30-11.00 CEST

For introductions we will use menti later, you can already go to the correct location for it: <a href="https://www.menti.com">www.menti.com</a> for voting code: **5066 1882** 

...sharing, collaborating, contributing, coordinating



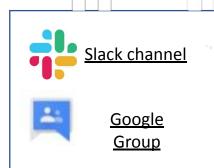
### **Open Science Training Coordinators**

# **Community of Practice**

www.openaire.eu/cop-training

# Owned and driven by members...







"The CoP can be regarded as a discipline transcending network of trainers and training organisers. We have built a community. We are exchanging information and best practices."

CoP member, 2020





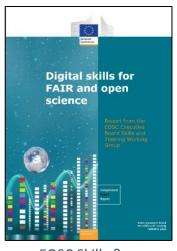
### **Open Science Training Coordinators**

# **Community of Practice**

www.openaire.eu/cop-training



Workshop report & recommendations



EOSC Skills & Training WG Report

Recommendation	Theme	Activity	Activity
Utilise the Framework of Actors in the EOSC Ecosystem in the development of initiatives, skills, training, reward and recognition frameworks and career paths necessary to support further development and mainstreaming of FAIR and open science.	Utilise Framework of Actors	EOSC Association - Task Force Data stewardship curricula and career paths	EOSC Association - Task Force Research careers, recognition and credit
Coordinate and align relevant skills curricula and training frameworks by generating a consensus on a core European higher education curriculum to deliver FAIR and open science skills at university level.	Coordination and alignment of curricula	EOSC Future - Task 9.1.2 Cross-domain EOSC Training alignment and EOSC wide training infrastructure (EOSC focus only)	FAIRsFAIR - WP6 FAIR competences
Encourage and support the competence centres approach as a framework for increasing coordinated provision of aligned training to support FAIR and open science.	Competence centres	FAIRsFAIR - WP6 Competence Centre and WP7 FAIR data science curricula and professionalisation	EOSC Association - Task Force Upskilling countries t engage in EOSC
Facilitate increased integration of FAIR and open science courses with university qualifications	Integration of FAIR & OS into curricula / qualifications	EOSC Synergy - WP6 Task 6.3 Embedding training resources in national education programmes	FAIRsFAIR - WP7 FAIR data science curricula and professionalisation

Mapping member activities
to EOSC report
recommendations

...sharing, collaborating, contributing, coordinating



# Workshop programme

9.30 - 9.40 Welcome by the CoP by Iryna Kuchma, EIFL

**9.40 - 10.00 Updates and introductions** (Ellen will manage the clock):

- RDA IG ETHRD minimal metadata learning resources focus group by Elizabeth Newbold, STFC
- Terms4FAIRskills by Laura Molloy, CODATA and Allyson Lister, University of Oxford
- Two minute madness on catalogues:
  - EOSC Pillar by Paula Oset Garcia, Ghent University
  - SSHOC Training Discovery Toolkit by Ellen Leenarts, DANS
  - Dariah Campus by Vicky Garnett, DARIAH
  - ELIXIR Training Portal TeSS by Celia van Gelder, DTL/ELIXIR-NL
  - EOSC Future by Lucia Vaira, LifeWatch ERIC

### 10.00 - 10.55 Challenges and discussions:

- Controlled vocabularies
- Curation process
- Sustainability

Collaborative notes: <a href="https://bit.ly/3hxWNX3">https://bit.ly/3hxWNX3</a>



# Mentimeter

- https://www.menti.com/qztcq9bwbk or
- <a href="https://www.menti.com">www.menti.com</a> and voting code:**5066 1882**



# RDA IG ETHRD (<u>Education and</u> <u>Training on handling of research data</u>)

# Focus group: Minimal Metadata Enabling Discovery & Findability of Learning Resources

To join the group or for further information see the RDA website: <a href="https://www.rd-alliance.org/groups/education-and-training-handling-research-data.html">https://www.rd-alliance.org/groups/education-and-training-handling-research-data.html</a>



# Challenge the focus group addresses

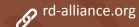


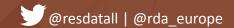
When looking to learn or upgrade research data skills, what's the best way to discover learning resources that help you do that?

- What are the most important descriptors that you can use to find appropriate learning resources – as a learner? -- as a trainer? as a service / infrastructure provider?
- Can you count on using similar descriptors for learning resources when searching places where they are collected (catalogues, registries, training portals)?

### Goals:

- Identify a minimal set of metadata descriptors that can be recommended as:
  - Important for resource users from different learning sectors (both formal and informal) and from different perspectives
  - A core set that can adopted by learning resource creators and service providers
- Help to reduce duplication & identify gaps among existing & prospective learning resource service providers









# Developing the minimal metadata set

- Initial mapping across different schemas generates a master list of metadata Development of user stories: Learners, Trainers, Service Providers, Developers, Funders
- Metadata assessed for relevance to support the user stories
- Community engagement and feedback via RDA Plenaries
- Resulting in a set of 14 elements

Title - The human readable name of the resource.

Abstract / Description - A brief synopsis about or description of the learning resource.

Author(s) - Name of entity(ies) authoring the resource.

Primary Language - Language in which the resource was originally published or made available.

**Keyword(s)** - Keywords or tags used to describe the resource.

License - A license document that applies to this content, typically indicated by URL

Version Date - Version date for the most recently published or broadcast resource.

URL to Resource - URL that resolves to the learning resource or to a "landing page" for the resource that contains important contextual information including the direct resolvable link to the resource, if applicable. Should be a PID, if possible.

Resource URL Type - Controlled vocabulary designating the persistent identifier schemes used for the resource, e.g., DOI, ARK, Handle.

Target Group (Audience) - Principal users(s) for which the resource was designed.

Learning Resource Type - The predominant type or kind that characterizes the learning resource.

Learning Outcome - Descriptions of what knowledge, skills or abilities a learner should acquire on completion of the resource.

Access Cost - Access cost: Choice stating whether or not there is a fee for use of the resource (CV = Y/N/Maybe with recommendation that further explanation of "Maybe" goes in the Description field for "It depends" or "It changes" explanations).

Expertise (Skill) Level - Target skill level in the topic being taught; example values include: beginner, intermediate, advanced.







# terms4FAIRskills

Info: <a href="https://terms4fairskills.github.io/">https://terms4fairskills.github.io/</a>

Contact: <u>terms4fairskills@codata.org</u>

Dr Laura Molloy, CODATA

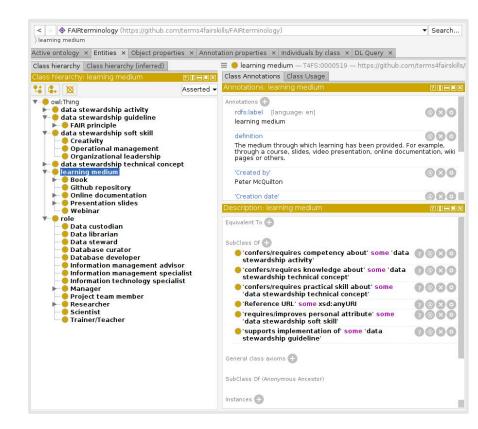
laura@codata.org

Dr Allyson Lister, FAIRsharing

allyson.lister@oerc.ox.ac.uk / contact@fairsharing.org

On behalf of the terms4FAIRskills community
Open Science Forum, 21 Sep 2021

CC BY-SA 4.0, <a href="https://creativecommons.org/licenses/by-sa/4.0/">https://creativecommons.org/licenses/by-sa/4.0/</a>











# terms4FAIRskills: scope and uses

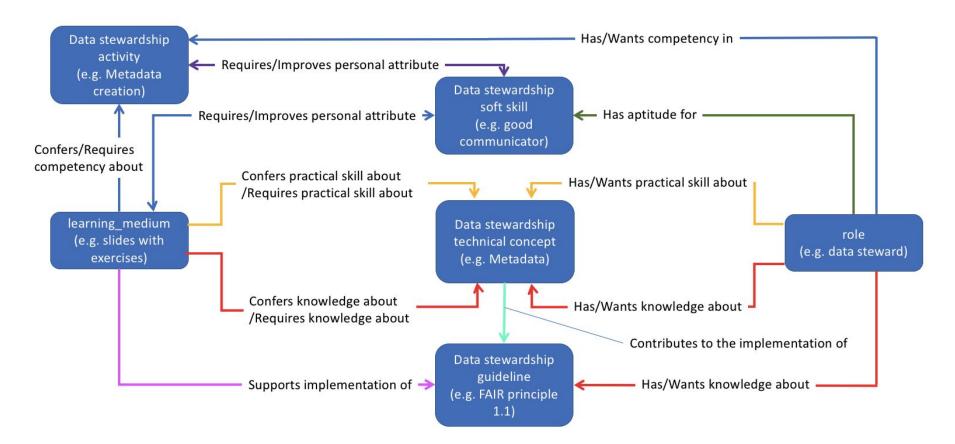
Training resources for Data Stewards are fragmented. T4FS is a terminology for the *competencies*, *skills* and *knowledge* necessary to make data FAIR and to keep it FAIR.

**Discovery**: facilitate the annotation, search and evaluation of FAIR-enabling materials (e.g. training) and resources

**Design**: assist the creation and assessment of data stewardship curricula **Training**: help trainers who teach FAIR data skills, researchers and others who wish to identify skill gaps in their own skills base / teams

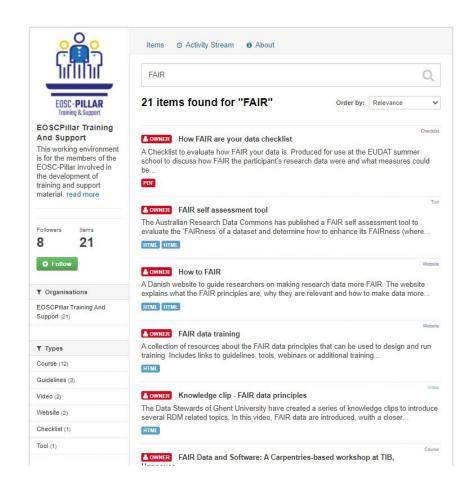
**Formalisation**: enable the definition of job descriptions and CVs with recognised, structured competencies

# terms4FAIRskills: model

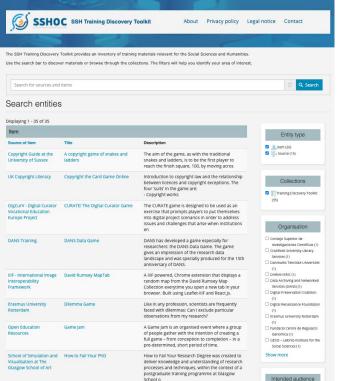


### **EOSC-Pillar**

- To curate RDM training and support materials and make them findable
- We are not the producers of the training material
- Based on our work and experiences as new data stewards team:
  - We look for resources & make them visible by publishing metadata + link to original source (several resources / URLs allowed per record)
  - Through feedback of community in EOSC-Pillar and in Flanders
- Both generic & discipline specific materials
- Any stage of research data lifecycle
- Training AND support materials



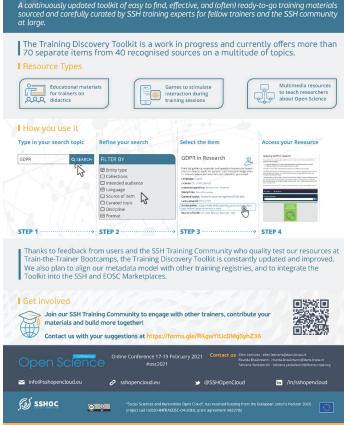
# **SSH Training Discovery Toolkit**



- Inventory reusable training materials
- Generic and discipline specific
- Faceted search
- Created in Drupal
- Data model based on SSH marketplace
- Curation by team
  - Input and feedback by the Training Community,
    Train-the-trainer bootcamps



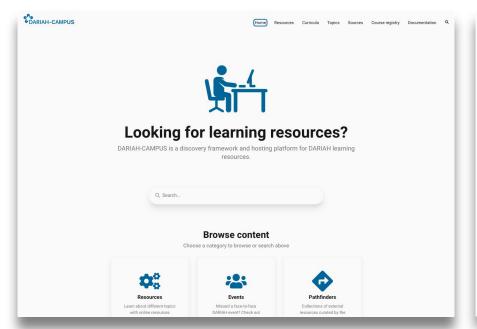
### A Training Discovery Toolkit for the Social Sciences and Humanities

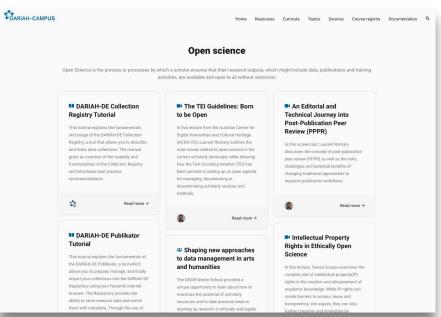


https://doi.org/10.5281/zenodo.4534192

https://training-toolkit.sshopencloud.eu/

# DARIAH-Campus (campus.dariah.eu)





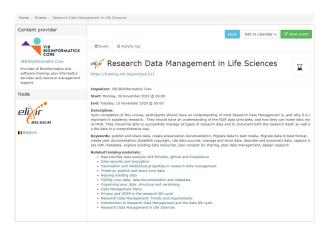
Find out more here: <a href="https://tinyurl.com/DARIAHCampusOpen">https://tinyurl.com/DARIAHCampusOpen</a>

## The ELIXIR Training Portal - TeSS

- TeSS is a registry to aggregate and disseminate training events and materials
  - September 2021: 1519 training materials,
     128 upcoming events, > 12000 past events
  - also linked to the ELIXIR RDMKit (<a href="https://rdmkit.elixir-europe.org/">https://rdmkit.elixir-europe.org/</a>)
- majority scraped automatically
  - via dedicated scrapers and via Bioschemas
- > 70 training content providers (not only ELIXIR Nodes!)
  - ELIXIR Nodes & their institutes, Carpentries, Coursera, Galaxy, PRACE, EMBL-ABR (Australia)
  - Interest in TeSS from EOSCHub, EOSCPilot, EOSC-Life, EJP-RD, Australian Research Data Commons (DreSA Portal will be launched soon!)
  - TeSS / ELIXIR Training aims to become an EOSC Training Provider
- Tess currently uses EDAM ontology and is in the process of implementing terms4fairskills
  - ELIXIR Training is a terms4fairskills use case







https://tess.elixir-europe.org/

# **EOSC Future Training Catalogue**

#### **Main goals** of WP9 – Training and skills:

- support FAIR sharing and reuse of training resources in EOSC
- build synergies (we don't need to start from scratch and we cannot do this alone) with other projects, clusters, research infrastructures nodes, associations, in order to benefit from other similar initiatives to avoid duplication of effort and focusing on interoperability, service offering, and data provision.



#### **Target audience** - our users can be divided into:

- EOSC Consumers (individual researchers; research communities; citizen scientists; commercial entities like private companies)
- EOSC Providers (infrastructure training managers; trainers)
- EOSC Facilitators (research administrators; research funders and policy makers; librarians)









# **EOSC Future Training Catalogue**

### Main functionalities of the EOSC Training Catalogue:

- basic and advanced browse and search interface (GUI and APIs);
- display of landing pages for included training resources;
- definition and management of a training resource metadata set (flexibility of the metadata set starting from results of the RDA Minimal Metadata for Learning Resources WG);
- metadata harvesting from other catalogues and external portals;
- manual content creation (for providers that do not have a catalogue);
- review process (community-validated quality criteria for training resources will be defined);
- versioning;
- feedback mechanisms;
- link with the EOSC Training Platform.



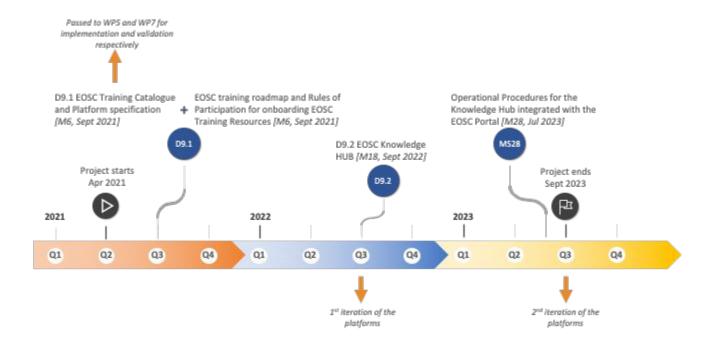






# **EOSC Future Training Catalogue**

#### **Timeline**











### Catalogue challenge: controlled vocabularies

### Challenges identified by the Catalogue Providers

- Hard to find/select suitable CVs e.g. resource type, keywords, topics/tags
  - domain specific versus domain-agnostic (Experienced by ELIXIR TeSS)
  - choose proper CVs: Ensure that Catalogue users know where (catalogues) and how (CVs) to search
  - when are new/emerging CVs mature enough to implement
- Lack of community agreement
  - experienced by EOSC-Pillar, SSHOC TDT
  - question: Who gets to decide for a for a specific CV for a specific catalogue?
- For an established catalogue: Adding or changing CV is a challenge
  - Curation and/or mapping needed (Experienced by SSHOC TDT, ELIXIR TeSS)
- Ensure interoperability between your catalogue and the rest of the catalogue ecosystem

### Challenge topic: controlled vocabularies

#### EOSC-Pillar:

- FAIR4S Framework: skills group (~ lifecycle stage, less intuitive), target group (-> EOSC actors?)
- Domain / discipline: re3data.org (-> <u>FAIRsharing Subjects Ontology</u>? Extends and enhances the re3data.org hierarchy)
- Other (e.g. ISO for languages and countries)
- Aim: terms4FAIRskills. Participated in "annotate-a-thon".

### SSHOC Training Discovery Toolkit:

- Our catalogue was based on an existing collection, challenge of adding or changing CV once you established the catalogue
- Curation of lists you introduced yourself can be a burden

#### EOSC Future:

 The catalogue will re-use of RDA WG solutions trying to avoid customised codelists as much as possible and to use common and standardised CVs

### Challenge topic: controlled vocabularies

- DARIAH-Campus:
  - EOSC- based categories
  - GitHub-based data files with pre-determined tags
  - Drop-down lists on CMS to control tags applied
- ELIXIR Training Portal TeSS:
  - uses BioSchemas
  - domain/disciplline specific & domain agnostic (EDAM & terms4fairskills)
    - challenge: terms4fairskills not complete/finalized yet
  - Tess aiming to incorporate both since we want to annotate our trainings with both CVs
  - ELIXIR Training is use case for terms4fairskils
    - participated in several terms4fairskills annotatethons
    - held two ELIXIR dedicated mini annotatethons to test suitability of the CVs

### Catalogue challenge: data model

- Lack of community agreement, or requirements for EOSC (EOSC-Pillar, SSHOC TDT)
- EOSC-Pillar:
  - CKAN. Metadata profile: common metadata + item specific metadata (field name, field type, mandatory, controlled vocabularies, validation rules...)
  - o Interoperable with other systems, e.g. DCAT and OAI-PMH
  - Metadata based on FAIRsFAIR D6.2 Initial Core Competence Centre Structures.
    - Other: report from the EOSC Skills and Training WG, RDA Metadata Standards Catalog WG

### SSHOC Training Discovery Toolkit

- Challenge of changing the data model once you established the catalogue
- Guidelines and the work from the RDA-ETHRD were still under development (and is evolving) → How do you remain flexible to adjust to new developments?
- Not interoperable yet.

#### EOSC Future

 The metadata set of the training resources will be based on the RDA WG and will be flexible to reflect the future developments

### Catalogue challenge: Curation process

- No systematic approach (EOSC-Pillar, SSHOC TDT, TeSS), but desired
- Small curation team, limited time (EOSC-Pillar, SSHOC TDT)
- Expertise & bias
  - E.g. (EOSC Pillar): Over-representation of generic or domain-agnostic resources; difficulty to judge quality of discipline specific content.
- Need for curation / quality evaluation criteria: collaborative work?
  - Establish contribution workflows, increase catalogue contributors or size of curation team
- No funding for curation or updating after the end of the project
  - Need for automatic expire date?

### Catalogue challenge: Curation process

### ELIXIR Training Portal TeSS

- TeSS is a registry: Number of materials is huge and from many different providers. Only control about our "own" (ELIXIR) training materials (from 23 countries).
- No clear process implemented yet but long on wishlist
- For 2022 funding to work out the process. First step: Curate on a higher level (e.g. Collections (e.g. on a topic, or for a specific Community).

### DARIAH-Campus

- Contributions made in first instance via an 'issue' form on GitHub
- Topics organised by use of 'tags'
- 'Sources' organised into 'channels' for multiple resources from single contributors
- Curricula organised content into courses
- GitHub provides structure and framework

#### EOSC Future:

- community-validated quality criteria for training resources
- review process among the functionalities of the catalogue → redactional workflow until the end of the project

# Catalogue challenge: sustainability

### EOSC-Pillar:

- mass export and import possibilities
- curation after project end

### SSHOC Training Discovery Toolkit

- o no funding for curation or maintenance after the project → responsibility problem: the catalogue is a (EOSC!) project output, who is responsible after?
- o difficult to transfer as new initiatives are also often project-based and have their own targets
- lack of vision on how catalogues should integrate across projects

#### EOSC Future

- it is built upon the work of previous EOSC projects and EOSC Association (EOSC Sustainability WG)
   → long-term sustainability
- community-validated quality criteria for training resources (re-use of RDA solutions in the EOSC context)
- harvesting from authoritative and validated sources → it allows high autonomy
- just started → still have 3 years to figure it out and to provide recommendations for sustainability

### Catalogue challenge: Sustainability

- ELIXIR Training Portal TeSS
  - Tess originated as ELIXIR-UK Node Service
  - From 2022 on it will become an ELIXIR Infrastructure Service, (limited) funding by ELIXIR Core Funding
  - Moving to co-development by several Nodes
  - Dividing "technical" task and "training" tasks

### DARIAH-Campus

- 'ERIC' status ensures longevity
  - CHALLENGE What to do with project resources that become obsolete?
- O DARIAH-Campus Reuse Charter ensures good practice in data management and reuse
  - https://campus.dariah.eu/docs/reuse-charter