BioData.pt elixir PORTUGAL

TRAINING DATA STEWARDS

FOR LIFE SCIENCES

FAIR data stewardship landscape

André Vieira & Pedro Principe,
University of Minho, Portugal
21/02/2022

Universidade do Minho
Serviço de Documentação e Bibliotecas

AGENDA

FAIR data stewardship landscape

1. FAIR data skills frameworks
2. FAIR data education and training programs
3. FAIRsFAIR project relevant outputs
    1. FAIR competence framework for higher education
    2. FAIR Data Training Handbook
    3. Good practices in FAIR competence training
4. Final remarks

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

OUR BACKGROUND

Scientific Information Management, Repositories and Open Science Office at University of Minho Documentation and Libraries Services

- https://openscience.usdb.uminho.pt/
- OpenAIRE - www.openaire.eu/support
- FAIRsFAIR – www.fairsfair.eu
- FOSTER Open Science - www.fosteropenscience.eu/
- FIT4RRI - https://fit4rri.eu/
- EOSC Future (WP9 – training & skills)
- Essencial GDI (MOOC) https://www.nau.edu.pt/curso/o-essencial-da-gestao-de-dados-de-investigacao/

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

Briefing on FAIR Competences and Synergies

An overview and analysis of different existing:

- skills and competence frameworks,
- training initiatives and higher education courses and programes

in RDM, Open Science and FAIR data.

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

D7.2 Briefing on FAIR Competences and Synergies

Work Package	WP7 - FAIR Data Science and Professionalisation
Lead Author (Org)	Pedro Principe, André Vieira (UMINHO), Lennart Stoy (EUA)
Contributing Author(s) (Org)	Bregt Saenen (EUA), Claudia Engelhardt (UGOE), Joy Davidson, Angus Whyte, Patricia Herterich (DCC), Yuri Demchenko (UvA), Ellen Leenarts (DANS)
Due Date	31.08.2020, M18
Date	02.09.2020

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

https://doi.org/10.5281/zenodo.4009006

Briefing on FAIR Competences and Synergies

The report confirms that despite a relatively large availability of training resources…

- **more formal definition of disciplinary RDM/FAIR data skills and competences are lacking,**

- **as well as a wider integration of RDM skills in bachelor or master-level curricula.**

In recent years, different initiatives and projects have made efforts to structure and define the competences and skills required for RDM/FAIR.

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

FAIRSFAIR
Fostering Fair Data Practices in Europe

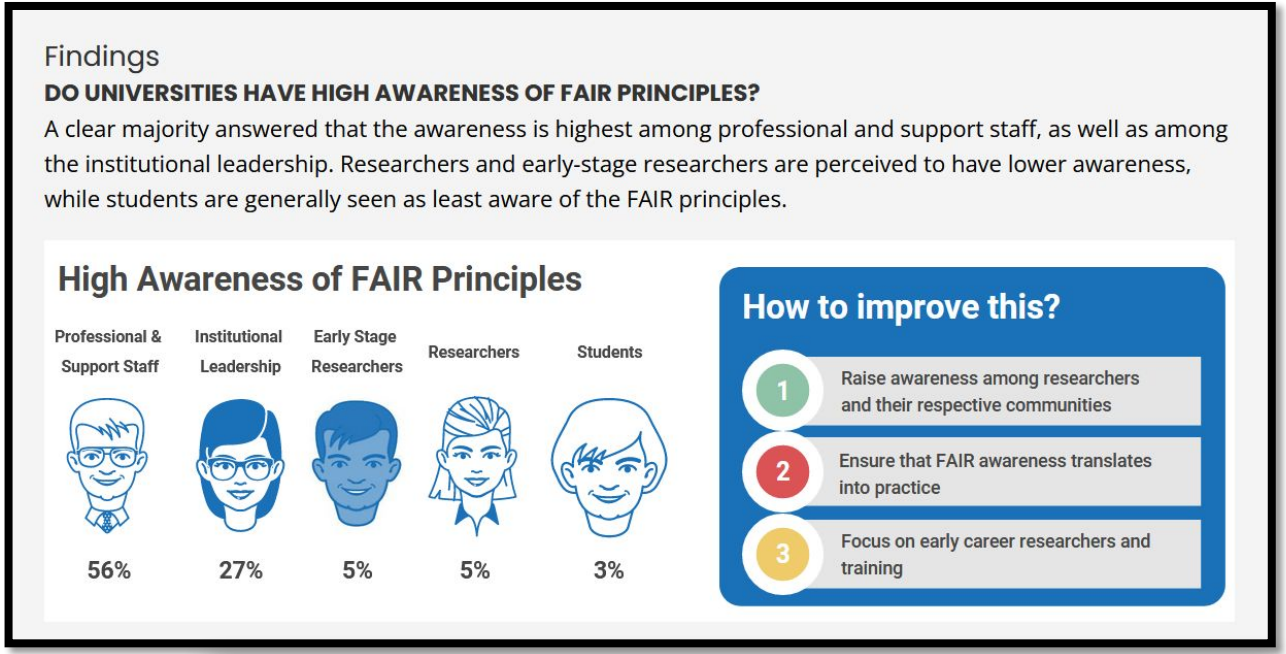
Project Title	Fostering FAIR Data Practices in Europe
Project Acronym	FAIRsFAIR
Grant Agreement No	831558
Instrument	H2020-INFRAEOSC-2018-4
Topic	INFRAEOSC-05-2018-2019 Support to the EOSC Governance
Start Date of Project	1st March 2019
Duration of Project	36 months
Project Website	www.fairsfair.eu

D7.2 Briefing on FAIR Competences and Synergies

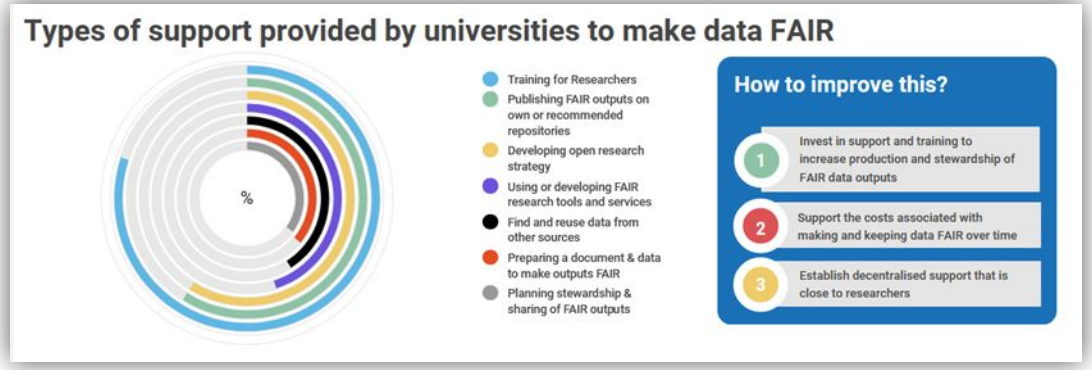
Work Package	WP7 - FAIR Data Science and Professionalisation
Lead Author (Org)	Pedro Principe, André Vieira (UMINHO), Lennart Stoy (EUA)
Contributing Author(s) (Org)	Bregt Saenen (EUA), Claudia Engelhardt (UGOE), Joy Davidson, Angus Whyte, Patricia Herterich (DCC), Yuri Demchenko (UvA), Ellen Leenarts (DANS)
Due Date	31.08.2020, M18
Date	02.09.2020

https://doi.org/10.5281/zenodo.4009006

FAIR data in European Higher Education



A groundbreaking study undertaken by FAIRsFAIR project - 90 survey responses received from universities in 26 countries
https://fairsfair.eu/fair-european-higher-education



TRAINING DATA STEWARDS
FOR LIFE SCIENCES

FAIR data skills (related) **frameworks**

Professional and PhD level

	Framework	Main Field	Addresses FAIR data?	Main elements
View	FAIR4S	FAIR data stewardship	Yes	Group of skills / competences for each research lifecycle stage; Identification of professional groups sharing stewardship responsibility.
View	FOSTER	Open Science	Yes	Open Science Taxonomy; Open Science training toolkit; Open Science Training Handbook oriented to practical teaching; Open Science Learning objectives.
View	Towards FAIR data steward as profession for the lifesciences	Data steward for Life sciences	Yes	Data Steward roles, function descriptions (responsibilities and tasks), competencies (knowledge, skills and abilities) and learning objectives.
View	BELMONT FORUM	Data skills in Global Environmental Change Research	No	Curricula goals, Curricula structure (core and optional modules, and for Principal Investigators).

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

FAIR data skills (related) **frameworks**

Professional, PhD and Bachelor levels

Framework	Main Field	Addresses FAIR data?	Main elements
View DeiC	Data Stewardship	Yes	Data Steward roles; Models for data stewardship education; Analysis of current educations in data stewardship and an evaluation of the needs and expectations to the role of a Data Steward across the private and public sector.

Master and Professional learners

Framework	Main Field	Addresses FAIR data?	Main elements
View GERMAN INFORMATICS SOCIETY	Data science; data science basic for other disciplines/domains	No	Curriculum recommendations; entry requirements for curricula; comparison of other skills frameworks.

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

FAIR data skills (related) **frameworks**

Master and Bachelor

Framework	Main Field	Addresses FAIR data?	Main elements
EDISON building the data science profession	Data science	Yes	- Data Science Competence Framework (CF-DS) (Part 1) - Data Science Body of Knowledge (DS-BoK) (Part 2) - Data Science Model Curriculum (MC-DS) (Part 3) - Data Science Professional Profiles (DSPP) (Part 4) - Data Science Taxonomies and Scientific Disciplines Classification

View

Other: data-related skills in a company / organisation setting

Framework	Main Field	Addresses FAIR data?	Main elements
CDI open data institute	Data science in a business / organisation context	No	Includes examples for training courses

View

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

FAIR4S key skills - https://eosc-fair4s.github.io/keyskills.html

Skills group	Key skill
Plan and design	Plan stewardship and sharing of FAIR outputs
Capture and process	Reuse data from existing sources
Integrate and analyse	Use or develop open research tools/services
Appraise and preserve	Prepare and document for FAIR outputs
Publish and release	Publish FAIR outputs on recommended repositories
Expose and discover	Recognise, cite and acknowledge contributions
Govern and assess	Develop open research strategy and vision
	Apply policies to comply legal requirements, ethical & FAIR principles
Scope and resource	Secure funding for open science /support
Advise and enable	Lead good practice by example

Skills and competence frameworks

Main findings (Briefing on FAIR Competences and Synergies)

- Focus on data science and data stewardship
- Domain-specific resources on data management skills seem scarce (except ELIXIR & Belmont)
- Not specifically designed for use at Bachelor or Master level
- Gap of resources providing information for programme designers about discipline specific practices and supporting material in form of frameworks and model courses or curricula

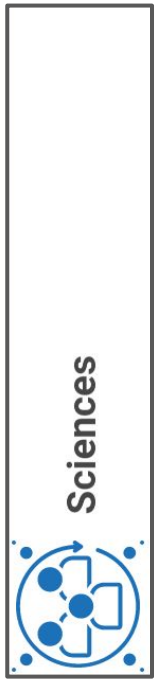
FAIR data education and training programmes

Generic Professional groups

Initiative	Professional Group Covered	Topics covered
RDA RESEARCH DATA ALLIANCE	All stakeholders involved in RDM activities.	All topics related with RDM, FAIR data and Open Science.
OpenAIRE	Researchers, data managers, repository managers, research support staff, funders, policy makers.	Addresses all topics related with RDM, FAIR data and Open Science.
GO FAIR	Individuals, institutions and organisations dealing with FAIR data principles.	FAIR data, Open Science.
FOSTER	Doctoral Students, Researchers, Research Project Managers, Knowledge Managers, Librarians, Funding Agencies.	Addresses all topics related with RDM, FAIR data and Open Science.
CODATA COMMITTEE ON DATA INTERNATIONAL SCIENCE COUNCIL	Early Career Researchers.	FAIR data, RDM, Data Science.
terms4FAIRskills	The completed terminology will be of use, for example, to trainers who teach FAIR data skills, researchers who wish to identify skill gaps in their teams and managers who need to recruit individuals to relevant roles.	FAIR data.

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

FAIR data education and training programmes



Initiative	Professional Group Covered	Topics covered
cessda	Social Sciences: Data producers, data users, data professionals, CESSDA Service Provider staff.	RDM, FAIR data.
SSHOC	Social Sciences and Humanities: Data producers, data users, data professionals..	RDM, FAIR data, Open Science, Data science.
DARIAH-EU	Digital Humanities: researchers.	RDM, FAIR data, Open Science.
PARTHENOS	Digital humanities and digital heritage: researchers, educators, managers, policy makers.	RDM, FAIR Principles and Open Science.
ENVRI FAIR	Environmental and Earth sciences: ENVRI Community.	RDM, FAIR data.
elixir	Life Sciences domain: researchers, Data Stewards, Reviewers and funders, Trainers.	RDM, FAIR data.
ExPaNDS	Photon and Neutron Research Infrastructures.	RDM, FAIR data.
panosc photon and neutron open science cloud	Neutron and Photon science.	RDM, FAIR data.

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

Training Programmes

Main findings (Briefing on FAIR Competences and Synergies)

- Strong focus on RDM and FAIR data skills for professionals in data-related positions (data stewards etc.) or researchers at different career levels.
- Bachelor or master-level students were not listed as target groups.
- Training programmes ranged from more generic data stewardship or data science initiatives to research infrastructures servicing different domains and fields

Supporting the uptake of FAIR data teaching and training

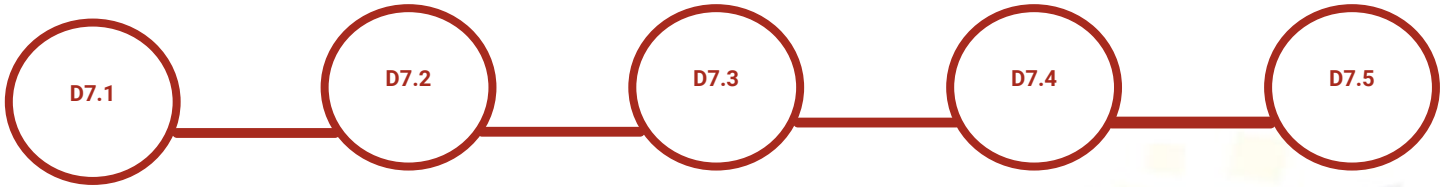
FAIR Data Science Curricula & profissionalisation

- **D7.1 FAIR Mapping analysis in European Higher Education** (February 2020)
- **D7.2 Briefing on FAIR Competences and Synergies** (August 2020)
- M7.4 Three stakeholder workshops (Feb. 2022)
- **D7.3 FAIR Competence Framework for Higher Education** (February 2021)
- **D7.4 FAIR Competences Adoption Handbook for Universities** (December 2021)
- **D7.5 Good Practices in FAIR Competence Training** (December 2021)
- M7.8 Three university workshops (Feb. 2022)

https://zenodo.org/communities/fairsfair

FAIR Data Science Curricula & profissionalisation

Through this work package higher education institutions will gain **practical tools** (D7.4 and 7.5), rooted in a comprehensive and up-to-date state-of-play survey (D7.1) and mapping of existing instruments (D7.2 and D7.3), **helping them to improve FAIR data science curricula** at Bachelor, Master and Doctoral level



Survey	Briefing	Competence framework	Adoption handbook	Good practices
FAIR in European higher education	FAIR competences and synergies	FAIR competence framework for higher education	FAIR competences adoption handbook for universities	Good practices in FAIR competence training
Link	Link	Link	Link	Link

Towards FAIR data competence framework

FAIR Competence Framework for Higher Education

Provides the competences definition structure that allows easy mapping to a **Body of Knowledge** and set of **Learning Outcomes** that can be used for defining academic curricula.

Provide a basis for building consensus on **defining Data Stewardship competences** and a **corresponding Body of Knowledge** and **Model Curriculum** that can be adopted by wider groups of adopters from the research community, academia and industry.

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

D7.3 FAIR Competence Framework for Higher Education (Data Stewardship Professional Competence Framework)

Work Package	WP7, FAIR Competences for Higher Education
Lead Author (Org)	Yuri Demchenko (UvA)
Contributing Author(s) (Org)	Lennart Stoy (EUA), Claudia Engelhardt (UGOE), Vinciane Gaillard (EUA)
Due Date	28.02.2021
Date	24.02.2021

DOI: https://doi.org/10.5281/zenodo.4562089

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

Body of Knowledge

Linking the competence framework with academic curricula.

A **Body of Knowledge** defines a set of **Knowledge Areas** and **Knowledge Units need to be included in a curriculum to achieve the intended Learning Outcomes** and defines a set of academic disciplines that can be taught in a curriculum.

Can be used as a basis for defining Data Science related curricula, courses, instructional methods, educational/course materials, and necessary practices for university post- and undergraduate programs and professional training courses.

FAIR Data Training Handbook

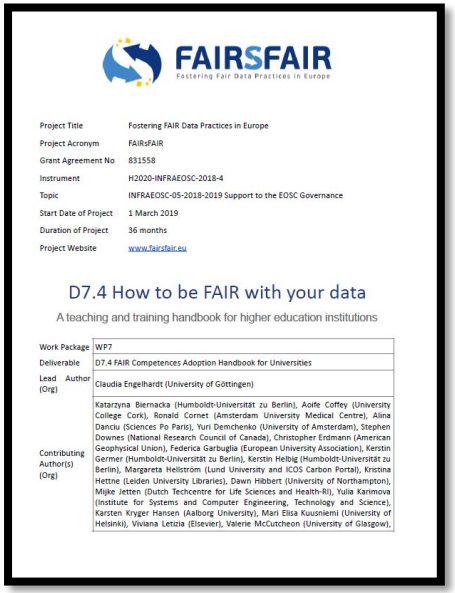
How to be FAIR with your data.

A teaching and training handbook for higher education institutions

Purpose:

Provide guidance and practical support with integrating the FAIR principles and related content into curricula and teaching.

DOI: https://doi.org/10.5281/zenodo.5837500



FAIRSFAIR
Fostering Fair Data Practices in Europe

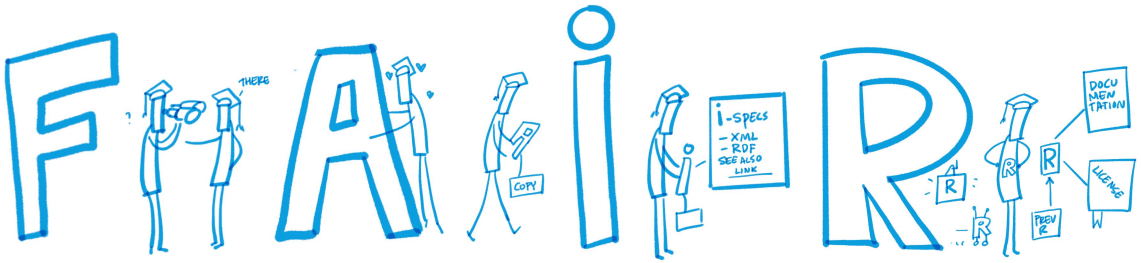
Project Title	Fostering FAIR Data Practices in Europe
Project Acronym	FAIRsFAIR
Grant Agreement No	831558
Instrument	H2020-INFRAEOSC-2018-4
Topic	INFRAEOSC-05-2018-2019 Support to the EOSC Governance
Start Date of Project	1 March 2019
Duration of Project	36 months
Project Website	www.fairsfair.eu

D7.4 How to be FAIR with your data

A teaching and training handbook for higher education institutions

Work Package	WP7
Deliverable	D7.4 FAIR Competences Adoption Handbook for Universities
Lead Author (Org)	Claudia Engelhardt (University of Göttingen)
Contributing Author(s) (Org)	Katarzyna Biernacka (Humboldt-Universität zu Berlin), Aoife Coffey (University College Cork), Ronald Cornet (Amsterdam University Medical Centre), Alina Dancu (Sciences Po Paris), Yuri Demchenko (University of Amsterdam), Stephen Downes (National Research Council of Canada), Christopher Erdmann (American Geophysical Union), Federica Garbuglia (European University Association), Kerstin Germer (Humboldt-Universität zu Berlin), Kerstin Helbig (Humboldt-Universität zu Berlin), Margareta Hellström (Lund University and ICOS Carbon Portal), Kristina Hettne (Leiden University Libraries), Dawn Hibbert (University of Northampton), Mijke Jetten (Dutch Techcentre for Life Sciences and Health-RI), Yulia Karimova (Institute for Systems and Computer Engineering, Technology and Science), Karsten Kryger Hansen (Aalborg University), Mari Elisa Kuusniemi (University of Helsinki), Viviana Letizia (Elsevier), Valerie McCutcheon (University of Glasgow),

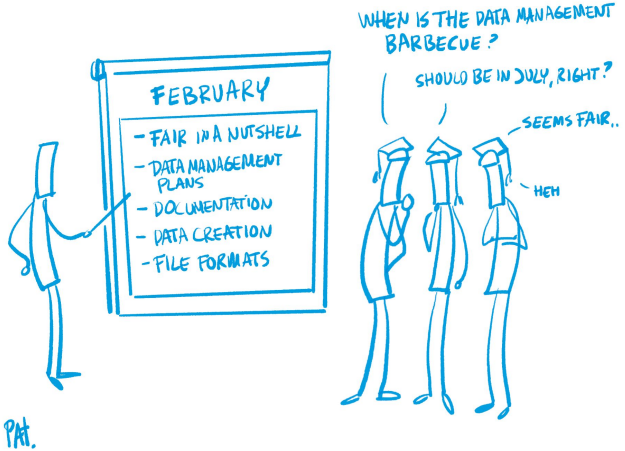
TRAINING DATA STEWARDS
FOR LIFE SCIENCES



FAIR Data Training Handbook

Contents – overview

1. Motivation
2. About this book
3. **FAIR Skills and Competences**
4. **Teaching and training designs for FAIR**
5. **FAIR lesson plans**
6. **Implementing FAIR**

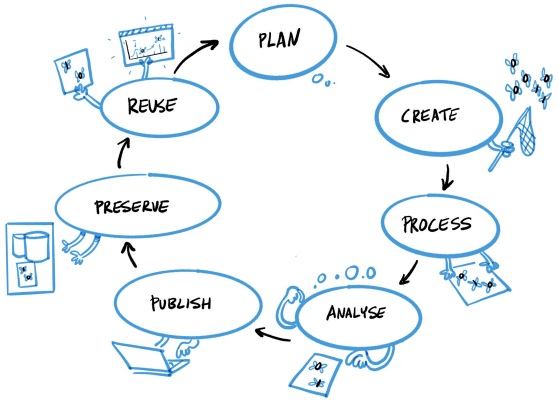


TRAINING DATA STEWARDS
FOR LIFE SCIENCES

FAIR Data Training Handbook

Lesson Plans

- FAIR in a nutshell
- DMPs
- Documentation
- Data Creation
- File Formats
- Data Standardisation and Ontologies
- PIDs
- Licences, Copyright, IPR
- Data reuse
- Repositories
- Sensitive data & ethical aspects
- Data Access
- FAIR Software/citable code
- RDM – overview & best practices
- Data Management and Governance in Industry and Research



TRAINING DATA STEWARDS
FOR LIFE SCIENCES

Good practices in FAIR Competence Education

- Collection of **7 case studies** highlighting institutional practices and training activities focused on teaching RDM and FAIR data skills at the institutional level.
- DOI: https://doi.org/10.5281/zenodo.5785252

Overview of good practices

Organiser	Name of initiative(s)	Type of initiative(s)	Target group(s)
U Bremen Research Alliance *Germany*	Data Train programme	Training programme	Doctoral candidates
NOVA University Lisbon *Portugal*	Research Data Management course	Teaching course	Doctoral candidates
TU Wien *Austria*	Center for Research Data Management and Data Stewardship course	Institutional unit dedicated to RDM and teaching course	Whole university community and Master students (course)
Tampere University *Finland*	Managing Research Information course and RDM: survey and interview data course	Teaching courses	Master students
swissuniversities *Switzerland*	Open Science programme	National funding scheme	Swiss higher education institutions
University of Minho *Portugal*	Essentials on Research Data Management MOOC	MOOC	Doctoral candidates, other researchers
University of Cape Town *South Africa*	Variety of training courses on RDM and FAIR data-related topics	Training courses	Whole university community

Additional tips and final remarks

FAIRSFAIR
Fostering Fair Data Practices in Europe

FAIR Aware

FAIR-Aware is:

- For researchers and data stewards
- An online tool to raise awareness and educate on data FAIRness
- 10 simple questions with practical tips to improve data FAIRness before deposit (English)

https://fairaware.dans.knaw.nl (English)

https://doranum.fr/enjeux-benefices/outil-fair-aware (French)

Provides an overview of awareness and willingness and suggestions for improvement

FAIR Aware

Thank you for your participation!

FAIR

Awareness:

High (8/10)

FAIR

Willingness to comply:

Low (27/50)

Guidance:

Based on your answers, you can find the guidance below to improve your awareness on some FAIR issues.

1. Are you aware that a dataset should be assigned a globally unique persistent and resolvable identifier when deposited with a data repository?

Selected datasets should be assigned a globally unique, persistent and resolvable identifier (PID) so they can be located unambiguously by humans or machines on the web. Persistent identifiers are maintained and governed so that they remain stable and direct the users to the same relevant object consistently over time. Examples of PIDs include Digital Object Identifier (DOI), the Handle System, identifiers.org, w3id.org and Archival Resource Key (ARK).

Identifiers are normally assigned by data repositories (or other service providers) when data and/or metadata are made available through their

https://fairaware.dans.knaw.nl/

Recognition of the need for professional data stewards

Priority Recommendation

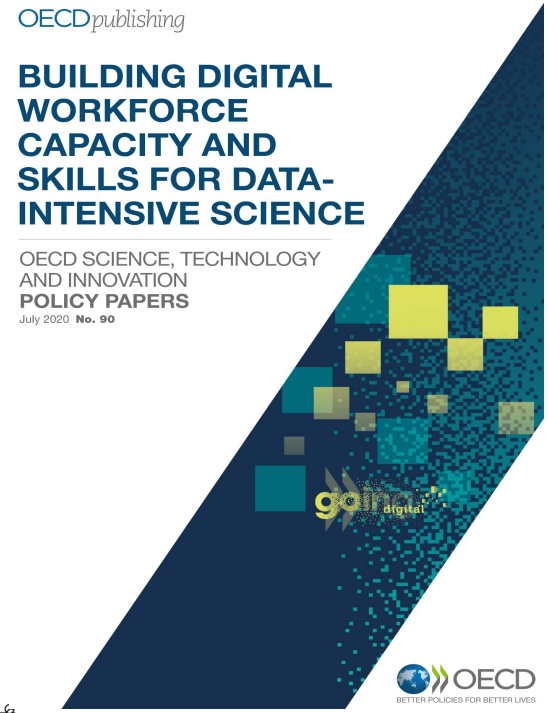
Rec. 10: "Professionalise data science and data stewardship roles and train researchers":

Steps need to be taken to develop **two cohorts** of professionals to support FAIR data: **data scientists** embedded in research projects, and **data stewards** who will ensure the management and curation of FAIR data. All researchers also need a foundational level of data skills.

European Commission. Directorate General for Research and Innovation. (2018), "TurningFAIR into reality: final report and action plan from the European Commission expert group on FAIR data", Publications Office, https://doi.org/10.2777/1524

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

Recognition of the need for professional data stewards



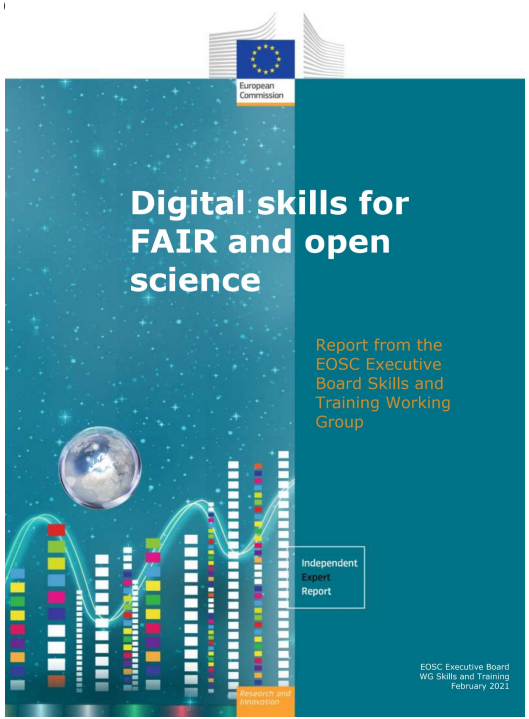
Key recommendations for universities and libraries

Support the development of **professional** communities in emerging roles such as **data stewards and RSEs**, and for trainers and leaders of digital skills initiatives.

OECD (2020), "Building digital workforce capacity and skills for data-intensive science", OECD Science, Technology and Industry Policy Papers, No. 90, OECD Publishing, Paris, https://doi.org/10.1787/e08aa3bb-en

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

Recognition of the need for professional data stewards



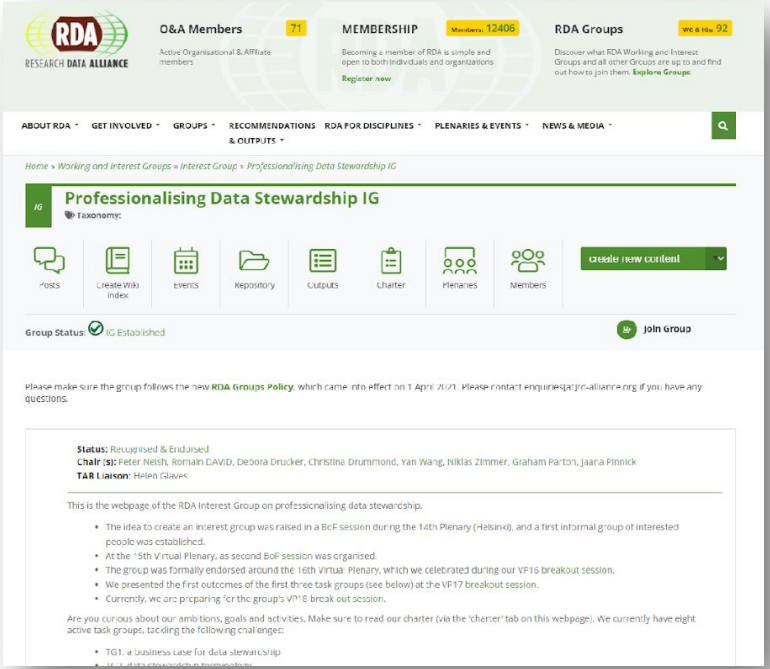
Research Software Engineers and Data stewards/Data librarians are identified as actors in the EOSC ecosystem

"What is clearly missing is a set of guidelines or similar support measures to help policy makers develop and formalise clear **career pathways** that are custom designed to target new research staff profiles aligned with open science." (p. 56)

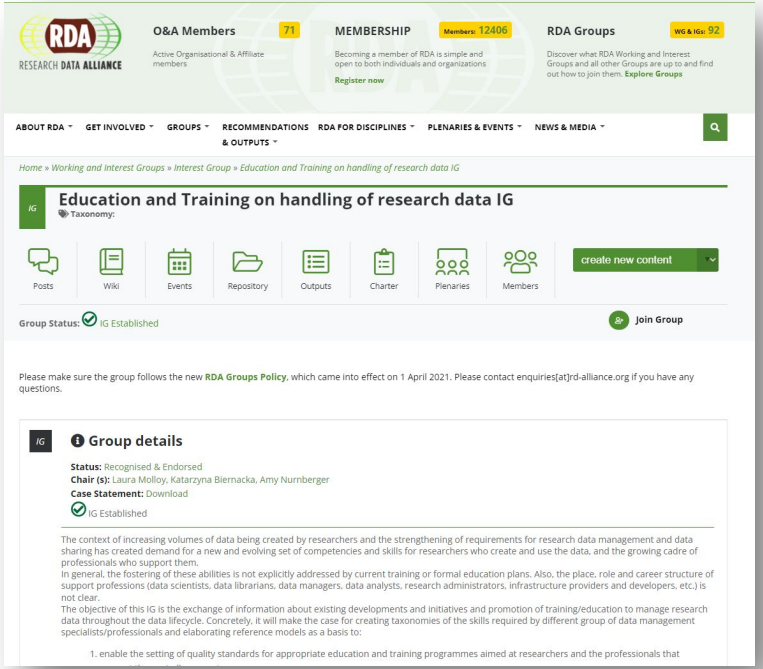
European Commission. Directorate General for Research and Innovation. & EOSC Executive Board (2021), "Digital skills for FAIR and Open Science: report from the EOSC Executive Board Skills and Training Working Group", Publications Office, https://doi.org/10.2777/59065

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

Follow RDA groups



www.rd-alliance.org/groups/professionalising-data-stewardship-ig



www.rd-alliance.org/groups/education-and-training-handling-research-data.html

TRAINING DATA STEWARDS
FOR LIFE SCIENCES

No single blueprint for data stewardship provision or FAIR data stewardship

…

Generic data steward (central support service) VS
Embedded data steward (directly involved)

…

Working with others, gathering best practices, implementation stories on how institutions are enabling FAIR data.

BioData.pt elixir PORTUGAL

TRAINING DATA STEWARDS

FOR LIFE SCIENCES

THANKS!

André Vieira andre.vieira@usdb.uminho.pt
Pedro Principe pedro.principe@usdb.uminho.pt

Universidade do Minho
Serviço de Documentação e Bibliotecas