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
Version	1.0
DOI	https://doi.org/10.5281/zenodo.4009006

Dissemination Level

x	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)

Abstract

This report provides an overview and analysis of different existing competence frameworks, resources and training initiatives in the field of RDM, Open Science and FAIR data, with the purpose to inform and support the future design and application of relevant skills frameworks, in particular to-be-developed within FAIRsFAIR project activities. It presents eight different skills frameworks in the field of RDM, data stewardship, and Open Science and provides a summary of the main elements and coverage of skills and competences of each. This report is complemented with a brief overview of several training initiatives from a selection of European projects and research infrastructures, and an analysis of a sample of higher education courses and programmes, identifying the main elements and coverage of skills of each case to analyse whether existing competence frameworks are utilised in this process.

Versioning and contribution history

Version	Date	Authors	Notes
0.1	12.03.2020	Pedro Principe, André Vieira (UMINHO) Lennart Stoy, Bregt Saenen (EUA), Claudia Engelhard (UGOE), Joy Davidson, Angus Whyte (DCC), Yuri Demchenko (UvA)	Initial draft structure of the briefing
0.2	18.06.2020	Pedro Principe, André Vieira (UMINHO)	Initial draft of the briefing
0.3	24.07.2020	André Vieira (UMINHO) Lennart Stoy, (EUA), Claudia Engelhard (UGOE)	Draft for WP7 review
0.4	07.08.2020	Pedro Principe, André Vieira (UMINHO) Lennart Stoy, Bregt Saenen (EUA),	Draft for internal review
0.5	21.08.2020	Patricia Herterich (DCC), Ellen Leenarts (DANS)	Internal review
1.0	02.09.2020	Pedro Principe (UMINHO) Lennart Stoy (EUA)	Revision following the internal review and final edits

Disclaimer

FAIRsFAIR 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.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Abbreviations and Acronyms

BMBF	Federal Ministry of Education and Research (Germany)	IFDS	Internet of FAIR Data and Services
BoF	Birds of a Feather format	IG	Interest Group
CESSDA	Consortium of European Social Science Data Archives	IN	Implementation Networks
CF-DS	Data Science Competence Framework (EDSF output)	ISC	International Science Council
DANS	Data Archiving and Networked Services	IT	Information technology
DARIAH	Digital Research Infrastructure for the Arts and Humanities	KPI	Key performance indicator
DeiC	Danish e-Infrastructure Cooperation	KSA	Knowledge, skills and abilities
DH	Digital Humanities	LCRDM	National Coordination Point Research Data Management (Netherlands)
DINI	Deutsche Initiative für Netzwerkinformation	LERU	League of European Research Universities
DM FORUM	Danish National Forum for Research Data Management	MC-DS	Data Science Model Curriculum (EDSF output)
DMP	Data management plan	MOOC	Massive Open Online Course
DOI	Digital object identifier	ODI	Open Data Institute
DS-BoK	Data Science Body of Knowledge (EDSF output)	OECD	Organisation for Economic Cooperation and Development
DSCC	Data Stewardship Competence Center	OS	Open Science
DSCC-IN	Data Stewardship Competence Centers Implementation Network	OS-CAM	Open Science Career Assessment Matrix
DSPP	Data Science Professional Profiles (EDSF output)	PaN RIs	Photon and Neutron Research Infrastructures
ECR	Early-career researchers	RDA	Research Data Alliance
ECTS	European Credit Transfer System	RDM	Research data management
EDSF	EDISON Data Science Framework	RPI / RPO	research performing institution / organisation
e-I&DM	Belmont Forum e-Infrastructure and Data Management initiative	SSHOC	Social Sciences & Humanities Open Cloud
EOSC	European Open Science Cloud	TG	Task Groups
FAIR	Findable, Accessible, Interoperable, Re-usable	UCL	University College London
GDPR	General Data Protection Regulation	WG	Working Group
ICT	Information and communication technology	WP	Work Package

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Executive Summary

This report provides an overview and analysis of different existing skills and competence frameworks, training initiatives and higher education courses and programmes in the field of RDM, Open Science and FAIR data, with the purpose to inform and support the future design and application of relevant skills frameworks, in particular to-be-developed within FAIRsFAIR project.

Research data management skills are relevant in the development of the European Open Science Cloud, as well as the attainment of European policy objectives on the advancing of Open Science skills within higher education curricula. How far data related skills can be embedded in higher education curricula, i.e. Bachelor and Master level, has proven a challenge due to limitations of available teaching staff, lack of space in programmes to add existing skills, and the way in which data-related skills.

As a basis for further discussion about integrating data management skills in higher education programmes, the report presents eight different skills frameworks in the field of RDM, data stewardship and Open Science and provides a summary of the main objectives, elements and coverage of skills and competences of each. We compared and analysed the different skills frameworks and find that, while a variety of resources already aim to define data science and data stewardship competences, domain-specific resources on data management skills seem scarce. Moreover, these frameworks are not specifically designed for use at Bachelor or Master level.

Additionally to the overview and comparative analysis of the selected frameworks, this briefing was supplemented with an overview of different training initiatives from a selection of European projects and research infrastructures. We found a 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 rarely listed as target groups.

An analysis of a sample of higher education courses and programmes revealed was conductedto identify the coverage of skills and competence of each case to analyse whether existing competence frameworks and training providers are utilised in this process. This analysis highlights that the practical use and impact of the skills frameworks is unclear and finds that, outside of a targeted projects to advance bachelor or master-level RDM education, RDM training appears targeted at doctoral or postdoctoral level in a more generic fashion.

The report confirms overall, that despite a relatively large availability of training resources, more formal definition of disciplinary RDM/FAIR data skills and competences in skills and competence frameworks are lacking, as well as a wider integration of RDM skills in bachelor or master-level curricula. On the other hand, the plethora of resources and activities identified in this briefing prove a high potential for collaboration and mutual learning across institutional, disciplinary or national boundaries.

FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Table of contents

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

List of Tables

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

List of figures

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

1. Introduction

Background and scope

FAIRsFAIR – Fostering Fair Data Practices in Europe – aims to support the emerging FAIR data culture (i.e., one where research data are findable, accessible, interoperable and reusable) in the context of the European Open Science Cloud (EOSC). For the sharing of research data, this calls for a wider adoption of research data management (RDM) practices and the FAIR principles[1] as increased production and use of FAIR data is seen as a cornerstone of EOSC.

This report is a step towards the "FAIR Competence Framework for Higher Education" (FAIRsFAIR Deliverable 7.3 due February 2021) which is intended to include FAIR data competences which can be acquired through higher education. It does so by providing a non-exhaustive overview and analysis of different existing frameworks, resources and training initiatives in the field of RDM and FAIR data.

It has the purpose to inform and support the future design and application of relevant frameworks, in particular the one to-be-developed by FAIRsFAIR, as well as the design of training offers. It also seeks to identify where and at which level other frameworks or initiatives are active. It complements and adds to earlier mapping exercises in FAIRsFAIR, which investigated the teaching of data science and RDM competences within higher education programmes, and establishes whether existing competence frameworks and training providers are utilised in this process.[2]

Structure

Chapter 1 continues with a primer on the role of training and skills in the EOSC process (Chapter 1.2). Previous insights about the state of FAIR and RDM within higher education and the role of skills frameworks are discussed in Chapter 1.3.

Chapter 2 contains an overview over a number of skills frameworks and training initiatives related to RDM and FAIR data. It presents eight different skills frameworks in the field of data stewardship and open science and provides a summary of the main objectives, elements and coverage of skills and competences of each. Chapter 2.2 compares and analyses the different skills framework and finds that, while a variety of resources already aim to define data science and data stewardship competences, domain-specific resources on data management skills seem scarce. Moreover, these frameworks are not specifically designed for use at Bachelor or Master level.

Chapter 3 contains an introduction into more than a dozen education and training initiatives from a selection of European research infrastructures and projects, again with a summary of their main

[1] Wilkinson, M., Dumontier, M., Aalbersberg, I., *et al.* (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018, pp. 1-9. https://doi.org/10.1038/sdata.2016.18.

[2] Stoy, Lennart, Saenen, Bregt, Davidson, Joy, Engelhardt, Claudia, & Gaillard, Vinciane. (2020). *D7.1 FAIR in European Higher Education* (Version v1.0_draft). Zenodo. https://doi.org/10.5281/zenodo.3629683

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

objectives, elements and coverage of skills and competences of each case. These initiatives are compared and analysed in Chapter 3.10 finding a 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 rarely listed as target groups.

Chapter 4 presents and analyses a sample of higher education courses and programmes, enriching the report with data on programme-level and institutional approaches to data-related skills. This analysis highlights that the practical use and impact of the skills frameworks is unclear and finds that, outside of a targeted projects to advance bachelor or master-level RDM education, RDM training appears targeted at doctoral or postdoctoral level in a more generic fashion.

Chapter 5 closes the report with a summary of the analysis provided in this report and an outlook for further work.

1.1. Training and skills in the context of EOSC

The recently published EOSC European Partnership Proposal outlines the ambitions of EOSC in the area of skills, training and competences. The document includes the proposed strategic objective that "Open Science practices and skills are rewarded and taught, becoming the 'new normal'".[3]

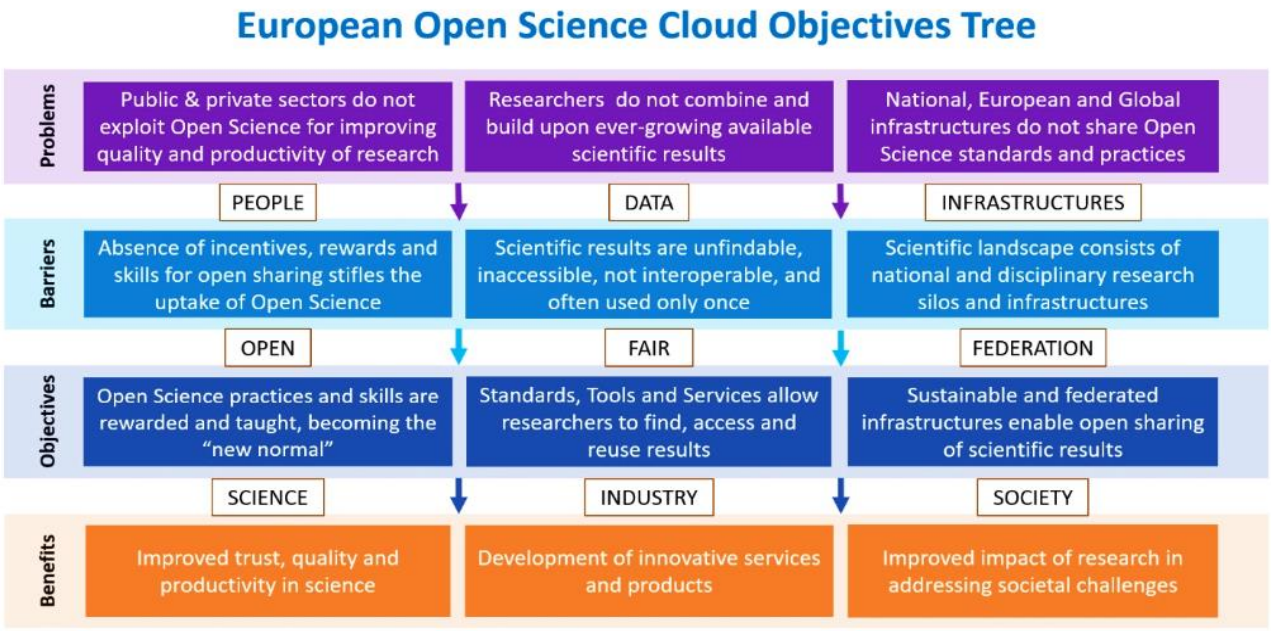

Figure 1. EOSC Objectives Tree (Source: *EOSC European Partnership Proposal*)

[3] *Draft proposal for a European Partnership under Horizon Europe European Open Science Cloud (EOSC) Partnership*, Version 28 May, 2020, p. 15.
https://ec.europa.eu/info/sites/info/files/research_and_innovation/funding/documents/ec_rtd_he-partnership-open-science-cloud-eosc.pdf

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

In the Partnership Proposal, skills development is integrated in the overall EOSC objectives tree (Fig. 1). The proposal – in its current form – states that EOSC will "promote the deployment of data stewards in academic institutions, transnational organisations and data-intensive industry by coordinating the development of training modules and certification mechanisms for EOSC-related digital skills."[4]

Therefore, one of the proposed operational objectives of EOSC by 2021-2023 is that "EOSC supports the training and deployment of professional data stewards together with European curriculum frameworks. (1.2)". A specific objective and Key Performance Indicator (KPI) associated with the objective would be that by 2025, "European curricula for data stewards are defined".

In the current governance structure of EOSC with a mandate until the end of 2020, a Skills & Training Working Group has been tasked with identifying skills needs and advising on a "Skills development framework (competences) for organizational culture change and service development"[5]. One of the main activities of the Working Group (WG) is further defining the breadth of professional profiles in the EOSC 'ecosystem' – and, by extension, the skills and competences required to build and use EOSC. Preliminary work contains a range from ICT and research infrastructure experts over data stewards to researchers with increased expertise in RDM and FAIR data (Fig. 2).

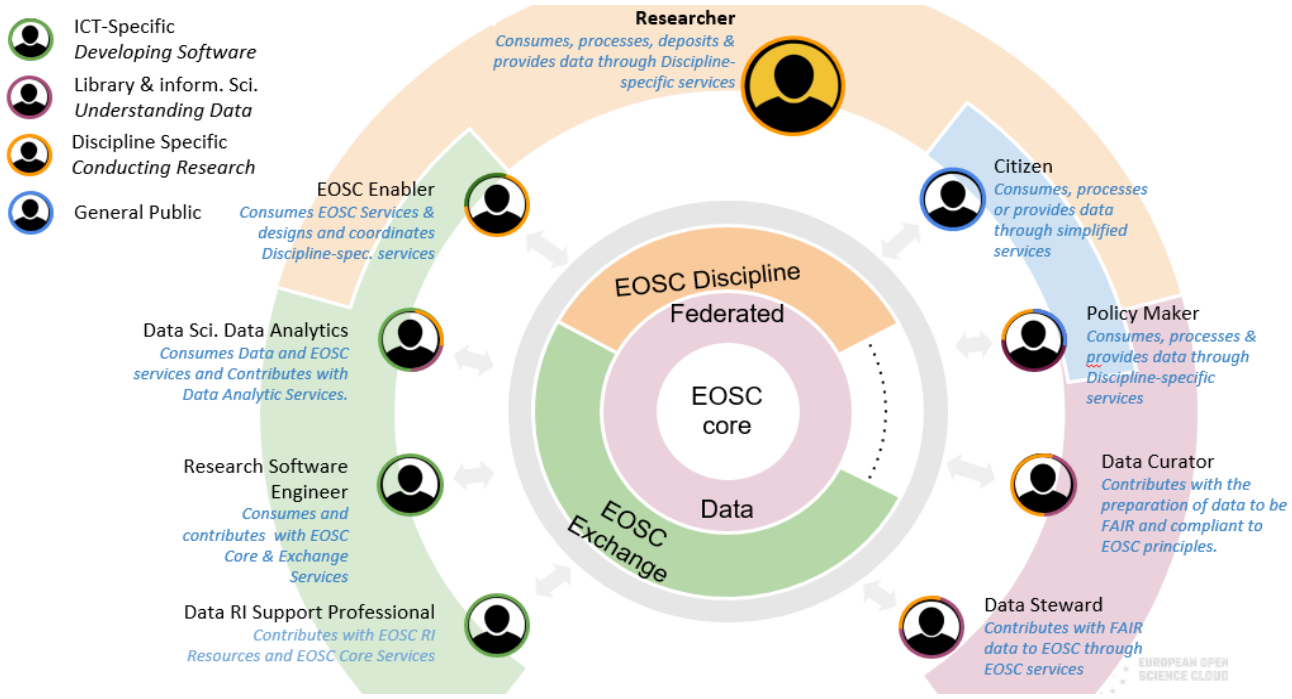

Figure 2. Preliminary of EOSC professional profiles by the EOSC Skills & Training Working Group (Source: www.eoscsecretariat.eu/sites/default/files/files/eosc_users.png)

[4] *Draft proposal for a European Partnership under Horizon Europe European Open Science Cloud (EOSC) Partnership*, Version 28 May, 2020, p. 12

[5] https://www.eoscsecretariat.eu/working-groups/skills-training-working-group

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

The FAIRsFAIR project is specifically aiming to create the conditions for a FAIR data research culture. Therefore, the project includes actions on skills and competences of researchers and professional or research support staff related to research data management and FAIR data. Research data management and FAIR data competences can be understood as a subset of the broader professional profiles and their different skills and competence groups that can be found in the recommendations and ambitions of EOSC at-large.

Work conducted by FAIRsFAIR in the context of the FAIRsFAIR Synchronization Force moreover shows that FAIRsFAIR is the only project (across EOSC-related projects participating in the mapping exercise) that directly contributes to the objective of establishing FAIR data stewardship curricula in the sense of formal education and curriculum frameworks.[6] Recently, this objective has received additional prominence through the updated European Skills Agenda, which calls for the creation of European open science curricula for researchers.[7] RDM and FAIR data skills can certainly be understood as core components of open science skills.[8]

1.2. Skills frameworks and university education for FAIR data

In recent years, different initiatives and projects have made efforts to structure and define the competences and skills required for RDM. This deliverable aims to provide a comprehensive overview of these existing skills and competence frameworks. It presents and analyses their main contents, components, and other defining qualities - and where available, their use cases. It complements and adds to earlier mapping exercises in FAIRsFAIR, which investigated the teaching of data science and RDM competences within higher education programmes, and whether existing competence frameworks and training providers are utilised in this process.[9]

According to these results, universities see a need to increase the teaching of data science and data management across domains and qualification levels. Where applicable, universities reported ample use of training material and projects (Fig. 3). This contrasts earlier findings which concluded that "there is very good coverage of the availability of open training resources and of international initiatives to provide researchers and others with relevant skills. There is patchier evidence on the take-up of these within individual institutions".[10] In particular, resources from the Research Data

[6] Ingrid Dillo, Marjan Grootveld, Simon Hodson, & Sara Pittonet Gaiarin. (2020). *Second Report of the FAIRsFAIR Synchronisation Force (D5.5)* (Version 1.0). Zenodo. https://doi.org/10.5281/zenodo.3953979

[7] *European Skills Agenda for sustainable competitiveness, social fairness and resilience*, COM(2020) 274 final, https://op.europa.eu/s/n92h

[8] See e.g. McCaffrey, Ciara, Meyer, Thorsten, Riera Quintero, Clara, Swiatek, Cecile, Marcerou-Ramel, Nathalie, Gillén, Camilla, … Egerton, Frank. (2020, March 10). *Open Science Skills Visualisation*. Zenodo. http://doi.org/10.5281/zenodo.3702401

[9] Stoy, Lennart, Saenen, Bregt, Davidson, Joy, Engelhardt, Claudia, & Gaillard, Vinciane. (2020). *D7.1 FAIR in European Higher Education* (Version v1.0_draft). Zenodo. https://doi.org/10.5281/zenodo.3629683

[10] Ashley, Kevin (2016) *Developing Skills for Managing Research Data and Software In Open Research*. Wellcome Trust. https://dx.doi.org/10.6084/m9.figshare.4133916

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Alliance (RDA), OpenAIRE, or provided by the FOSTER project were utilised – and others from GO-FAIR, CODATA, and research infrastructures such as ELIXIR.

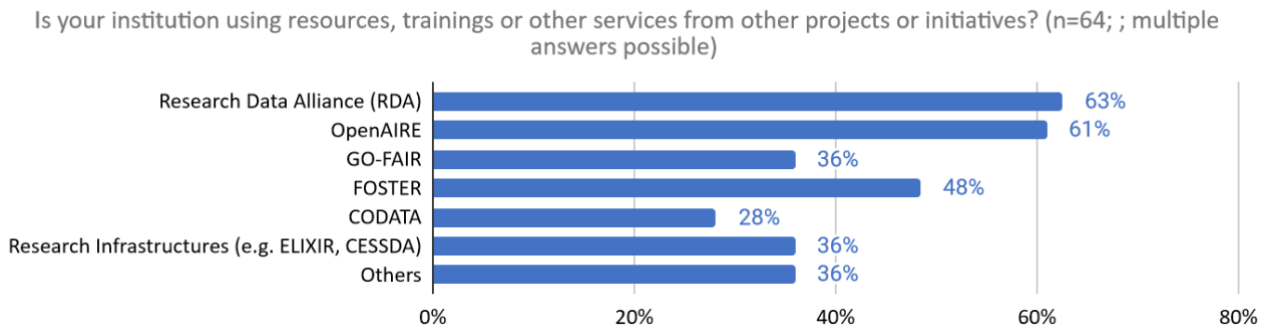

Figure 3. Use of existing resources, trainings or other services from other projects or initiatives (Source: D7.1 FAIR in European Higher Education)

In contrast, specific European (or other) frameworks were barely employed for the development of training activities or curricula (Fig. 4). Only 2% of responding universities indicated to have used the EDISON Data Science Framework (EDSF), and 8% reported to have used FAIR4S. The Open Science Career Assessment Matrix (OS-CAM) was used at 6% of universities (although not a competence framework in the strict sense). Where applicable, universities indicated to use "national frameworks or guidelines" (40%). However, a majority of 49% stated that "no frameworks or resources are used".

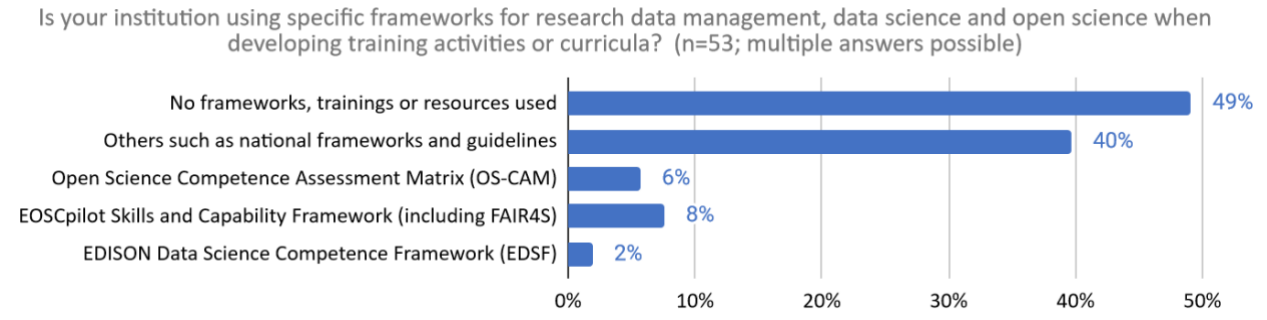

Figure 4. Use of frameworks for RDM, data science and open science for training or curricula development (Source: D7.1 FAIR in European Higher Education)

In short, despite an increasing availability of frameworks defining data science and/or RDM-related skills and competences, there is limited evidence how they affect the level of study programme

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

designs or individual study programmes. Continued exploration of the factors that may be hindering their utilisation in this context is needed.

A summary report of several RDM training projects in Germany, for instance, highlighted structural barriers related to study programme design which slow the integration of RDM competences. This includes the necessity to identify existing courses or modules that allow integrating RDM competences or the need for flexibility of teaching staff to adapt course contents. Nonetheless, the report recommends the integration of RDM into existing modules and programmes at Bachelor and Master level – initially for instance by including RDM in research methods courses. Moreover, it emphasises that – beyond a basic understanding of RDM that may be similar across different disciplines – RDM education should be embedded in disciplinary or domain-specific practices.[11]

A similar argument is articulated in a recent OECD report authored by an international group of experts for data-intensive science, which finds that "specialised digital skills needs differ across scientific domains and evolve over time in response to the state of the discipline and to new technologies and policy requirements." Therefore, the report further argues that "there is a baseline of more generic foundational digital skills for research that evolves more slowly and should ideally be embedded in undergraduate science education but currently is not".[12]

Universities UK has identified related challenges in a study on the status of undergraduate level data analytics education and training at British universities. Besides identifying challenges for new students at foundational level, such as "variable mathematical knowledge and training" or "variable computer science and coding skills" in the transition to university studies, the report warns that teaching quantitative methods isolated from the rest of the curriculum can cause the perception that "these methods are not a core component of a range of disciplines".[13]

The report also made reference to problems of data analytics not being equally relevant across different domains, observing that that "value of quantitative data is not universally accepted within a number of key academic fields"[14], which might deter academic staff from further engaging. Although discussing data analytics rather than research data management, these findings echo sentiments of developing a FAIR research culture across different disciplines.

---

[11] Einwächter, Sophie G., Esther Krähwinkel, and Frederik Ostsieker. (2020). „Lessons Learned: Thesen zur FDM-Kompetenzausbildung: Erkenntnisse aus dem Vernetzungstreffen der vom BMBF geförderten Projekte EeFDM Jena, FDMentor, FOKUS, PODMAN Und UniLLAB am 30. und 31. Januar 2019 in Marburg". *Bausteine Forschungsdatenmanagement*, Nr. 1 (April), 8-15. https://doi.org/10.17192/bfdm.2020.1.8101

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

[13] Universities UK (2015). *Making the most of data: Data skills training in English universities*, p. 22. www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2015/making-the-most-of-data-training-skills-english-universities.pdf

[14] Ibid., p. 21.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Table 1. Possible solutions to alleviate staff shortages

Possible solution	Pros	Cons
Upskilling current staff	Methods teaching embedded as part of the wider curriculum. Relatively low amount of resource and qualified staff required.	In some instances, staff members do not want to learn data analysis skills. Requires pre-existing institutional expertise. Potentially long lead-in time.
Service-level teaching from the statistics or equivalent department	Relies on existing institutional resource. Relatively short lead-in time.	More difficult to embed methods teaching within wider curriculum. In some institutions, existing staff may have limited capacity to do this. Staff members may not want their students to be taught methods they themselves do not feel confident using. High cost might mean that home department takes teaching in-house to save money.
Developing quantitative methods lectures centrally	Relatively low number of qualified staff required.	More difficult to embed methods teaching within wider curriculum. Potentially long lead-in time as it requires the careful design of both plenary lectures and smaller seminars with access to computers.

(Source: Making the most of data: Data skills training in English universities)

A further challenge identified by Universities UK was a shortage of skilled lecturers. Therefore, solutions must take into account the existing constraints of academic staff. Potential approaches could include, first, upskilling current staff, second, service-level teaching from the statistics or equivalent department, or, third, developing quantitative methods lectures centrally. All solutions come with associated advantages and disadvantages (Table 1). For RDM training, similar potential solutions as well as challenges can be assumed.

Beyond challenges related to content and available skills of teaching staff, degree structures also pose limitations. In a typical three-year undergraduate programme, "students across subject areas must be taught a range of technical skills as well as the domain knowledge required to develop appropriate research designs and critically evaluate data".[15] Consequently, "a major challenge facing

[15] Ibid., p. 23

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

academics is how to balance training in various skills and topics within a three-year curriculum".[16] All in all, to include additional data analytics - and by extension RDM skills - in curricula, different challenges pertaining to the knowledge of students, the availability of experienced academic staff, and curriculum structure need to be considered.

The Universities UK report also touches upon the topic of standardisation and structured overviews over skills by suggesting that "standards would be helpful in order to define core skills. Standards could offer insights into future career progression [...]. This will enable a distinction to be made between the core skills that should be taught at undergraduate level, and the targeted training and support that could be offered afterwards."[17] This recommendation relates back to the role of competence and skills frameworks discussed at the beginning of the section.

The aforementioned FAIRsFAIR report *D7.1 FAIR in European Higher Education* investigated whether different types of model courses and curricula or competence framework would be perceived as useful supporting resources for higher education institutions (Fig. 5). Here, respondents indicated that domain specific model courses and curricula, as well as competence frameworks that relate to domain-specific needs would be among the most helpful resources to draw from.

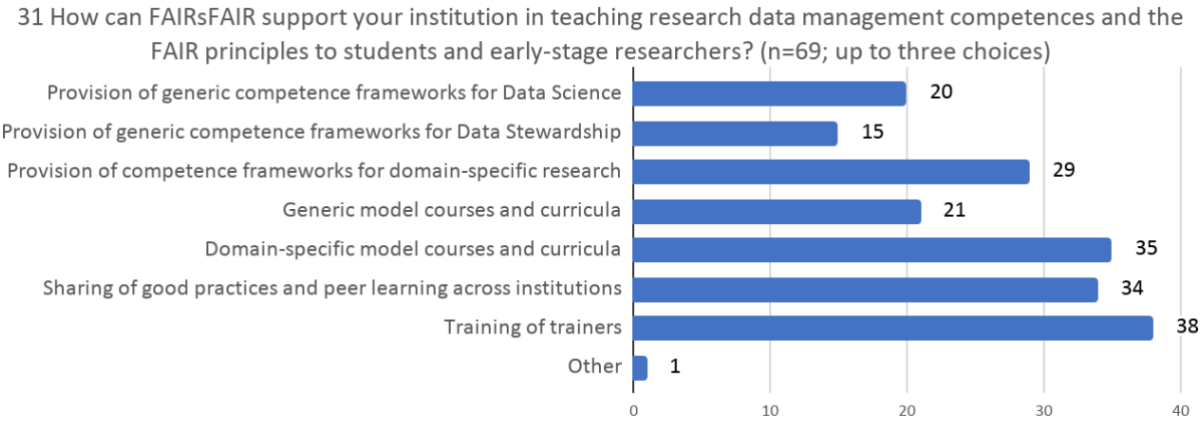

Figure 5. FAIRsFAIR support for teaching FAIR RDM competences
(Source: D7.1 FAIR in European Higher Education)

An additional use of such frameworks, according to the OECD report, is that they "can help to identify how many people with particular skills at a given level of expertise are required at different scales, from national to local. Some types of frameworks will also indicate how those skills can be acquired

[16] Ibid., p. 23

[17] Ibid., p. 23

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

- for example, through formal education, self-paced learning or mentoring. This allows appropriate policy interventions to be more effectively targeted".[18] Therefore, frameworks identifying skills and competences for different domains as well as at different levels, as well as recommendations on how to acquire them, can be useful tools for policy interventions as well as organisations which offer training and education or are developing institutional capacity.

1.3. Relationship to other FAIRsFAIR project and EOSC outputs

Over the entire project, FAIRsFAIR is building a body of insights and data about the practical implementation of the FAIR principles in research and education. Related activities are taking place in different EOSC-related projects and Working Groups. This body of FAIRsFAIR outputs and other EOSC related documents will help identify existing resources, training resources, gaps and synergies concerning RDM and FAIR data education within FAIRsFAIR and other EOSC projects.

Table 2. Overview of FAIRsFAIR and other relevant EOSC activities

Resources	Origin	Relevance
D7.1 FAIR in European Higher Education Data for D7.1 FAIR in European Higher Education	FAIRsFAIR: WP7	Survey results and underlying data data on data science and RDM education activities and needs at European universities.
D3.1 FAIR Policy Landscape Analysis	FAIRsFAIR: WP3	Analysis of data policy landscape at various levels (national, funder, publisher, institutional) in 2019 to identify policy elements that support or hinder FAIR data practice, including skills.
D6.1 Overview of needs for competence centres	FAIRsFAIR: WP6	Report on training needs as part of operating FAIR competence centres
FAIRsFAIR Policy and Practice Survey 2019 data for D3.1_D3.2_D6.1	FAIRsFAIR: WP3, WP6	Data collection on FAIR practice across research organisations
Seven Recommendations for Implementation of FAIR Practice FAIR in practice reference list (Google sheet version)	EOSC FAIR WG: FAIR Practice Task Force	Recommendations on developing FAIR practices in research, supported by a collection of resources on existing practices supporting FAIR data
D7.1: Skills landscape analysis and competence model	EOSCpilot: WP7	Outputs of the work of the EOSCpilot Skills and Capacity work package (WP7). Includes an overview

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

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

D7.2: Interim report and catalogue of EOSC skills training and educational materials D7.3: Skills and Capability Framework		of training initiatives and the EOSCpilot Skills Framework, which is used to categorise, assess and catalogue training resources.
FAIR-Aware tool https://fairaware.dans.knaw.nl/	FAIRsFAIR: WP4	Online tool helping researchers and other professionals to assess their level of awareness on making your datasets findable, accessible, interoperable and reusable (FAIR) before uploading them in a data repository.
Workshop report "Training in the EOSC"	DANS, EGI, EUDAT and OpenAIRE, funded by EOSCsecretariat.eu	Output of a three-day workshop on Training in the EOSC (February 2020), providing recommendations regarding Rules of Participation for training as well as recommendations regarding practical guidance for training service providers.

Table 2 shows a non-exhaustive list of relevant FAIRsFAIR and other EOSC activities that address skills and training. Taken together, they offer insights into FAIR data policies, RDM practices at different levels, activities of RDM competence centres, basic FAIR awareness levels and training initiatives both disciplinary and generic.

A major objective of FAIRsFAIR WP7 is to develop a "FAIR Competence Framework for Higher Education" (D7.3 due February 2021) which will include FAIR data competences which can be acquired through higher education and FAIR data competences for graduates continuing to work as professionals in FAIR data management. This work needs to be informed by practice and avoid overlapping with other projects and initiatives. Knowing the range of activities within FAIRsFAIR and in the EOSC ecosystem will allow WP7 to realise synergies and draw from the outputs and expertise of the partners in FAIRsFAIR.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

2. Overview and analysis of existing competence frameworks in Europe and abroad

For this briefing, we collected information about a sample of skills and competence frameworks, as well as about a number of education and training programmes at different levels. Skills and competence frameworks are presented in Chapter 2.1 and education and training programmes in Chapter 2.2.

As the objective was to deliver a non-exhaustive overview of relevant existing competence frameworks in Europe and abroad with a specific focus on the inclusion of Open Science and FAIR data competences, the focus in Chapter 2.1 lies on frameworks with a minimum content on skills and competences related to RDM and/or Open Science. The selection of frameworks is reported in Table 3 below. More basic frameworks such as DigComp[19] were deemed not detailed enough for insights into FAIR RDM practices and therefore not considered for this exercise - although they may form a basis for the more granular frameworks included in this briefing. The frameworks represent a combination of new documents with resources that have been discussed within previous projects.[20]

Addressing digital skills within higher education curricula necessarily needs to take into consideration that many graduates at bachelor, master or doctoral level will not be employed in academia but in the private or public sector. Therefore, the needs of the private or public sector should be considered when developing data-related training and education, which is reflected in the selection of frameworks.

Table 3. Overview of frameworks included in Chapter 2.1

Organization / Initiative/ Project	**Short description**	**Domain**	**Link**
EOSCpilot FAIR4S Framework	FAIR4S is the skills and capability framework that was developed in WP7 in the EOSCpilot project. FAIR4S aims to help research communities and	Generic	Webpage Deliverable

[19] See Vuorikari, R., Punie, Y., Carretero Gomez S., Van den Brande, G. (2016). *DigComp 2.0: The Digital Competence Framework for Citizens. Update Phase 1: The Conceptual Reference Model*. Luxembourg Publication Office of the European Union. EUR 27948 EN. doi:10.2791/11517 and Carretero, S.; Vuorikari, R. and Punie, Y. (2017). *DigComp 2.1: The Digital Competence Framework for Citizens with eight profi-ciency levels and examples of use*, EUR 28558 EN, doi:10.2760/38842

[20] EOSCpilot WP7 included a landscape analysis of education and training initiatives and relevant skills and training frameworks. See Whyte, A., Ashley K. (2017). *D7.1: Skills landscape analysis and competence model*. https://eoscpilot.eu/sites/default/files/eoscpilot-d7.1.pdf.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

	research institutions implement Research Data Management and FAIR stewardship in the open science and data science context.		
FOSTER Open Science	The FOSTER portal is an e-learning platform that brings together training resources addressed to those who need to know more about Open Science, or need to develop strategies and skills for implementing Open Science practices in their daily workflows. FOSTER also defined the Open Science Learning Objectives for the main stakeholders in the Research Ecosystem.	Generic	Webpage
Towards a community-endorsed data steward profession description for life science research	The aim of the project is to professionalise the data steward function in the life sciences domain in the Netherlands. The project works towards a common job description (including responsibilities and tasks) and an agreement on the required competencies (knowledge, skills and abilities (KSAs)) for the data steward function.	Life Sciences	Final report
DeiC - Models for a Data Steward Education	DeiC provides an analysis of current education in data stewardship and an evaluation of the needs and expectations to the role of a Data Steward across the private and public sector, aiming to inform future data stewardship education in Denmark.	Generic	Report
EDISON Data Science Framework (EDSF)	EDSF provides a basis for the definition of the Data Science profession and enables the definition of the other components related to Data Science education, training, organisational roles definition and skills management, as well as professional certification.	Data Science	Webpage Report
Data Science: Learning and training content (GFI)	The paper produced by the working group on Data Science/Data Literacy of the German Society for Informatics (GFI), in collaboration with the Federal Ministry of Education and Research (BMBF), details recommendations for learning content in Data Science at master level and for continued professional development.	Data Science	Working paper

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

	The report includes a comparison of five different frameworks for data science skills.		
Open Data Institute (ODI) Data Skills Framework	The framework is a high-level collection of 26 different technical and non-technical skills, aimed at organisational leaders, learning and development specialists or "teams aiming to develop skills and collaborate better". It intends to serve as an assessment tool to identify existing skills within an organisation and to, subsequently, address or fill gaps by developing training programmes.	Generic	Report
Belmont Forum - Data skills curricula framework	The Belmont Forum conducted a study on skills gap analysis and curricula for data intensive digital skills in Global Environmental Change Research, which resulted in a proposal for a curricula framework for the development of digital skills for data intensive global change research. This curriculum was endorsed by the Belmont Forum in November 2017.	Environmental Sciences	Report

A similar approach was chosen for the overview of education and training programmes in Chapter 2.2. They represent different actors and initiatives implementing or shaping education, training and skills development in RDM. It includes bottom-up initiatives such as RDA, e-infrastructures such as OpenAIRE, international initiatives like GO-FAIR and CODATA, projects such as and terms4FAIRskills, and a set of Europe-based research infrastructures developing more domain-specific RDM.

2.1. Skills frameworks

The following chapters briefly presents each framework and provides a summary of the main objectives, elements and coverage of skills and competences. A more structured analysis and comparison of the different frameworks is provided in Chapter 2.2.

2.1.1. FAIR4S Framework

FAIR4S (https://eosc-fair4s.github.io/) is the result of the EOSCpilot framework of FAIR data stewardship skills for science and scholarship.

The EOSCpilot FAIR4S Framework aims to help researchers and organisations to identify the capabilities and skills required to ensure research objects are 'FAIR' (Findable, Accessible,

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Interoperable and Reusable), and that they stay FAIR. In other words the skills to provide stewardship of research outputs. Stewardship skills development is likely to involve a range of 'data experts' - data stewards and other professional groups involved in producing research objects (data managers, data service engineers, and data scientists/analysts)[21].

FAIR4S groups the professional roles sharing stewardship responsibility, as follows:

- Data manager, data curator, data librarian,
- Data service engineer, data science engineer, research service developer, research software engineer.

FAIR4S sees the data steward role as combining a data management/curation function with at least one other role, i.e. research, data science, or engineering.

Skills framework with a focus on data stewardship[22]:

- About skills, competencies, capabilities, for people and organisations,
- A lifecycle-based approach to describing skills groups.

FAIR4S competences and capabilities

The identified skills are organised around the research life cycle and help to answer the following questions:

- What skills are needed to build, operate, support, use a particular EOSC service?
- What capabilities should an organisation build, through recruitment, training, staff development?
- What should job descriptions contain relating to data stewardship?

The following table presents the key skills for each research lifecycle stage (Skills group). Each key skill was selected as a priority for at least one role to develop its stewardship responsibilities. The full view of the competences and capabilities framework is available at https://eosc-fair4s.github.io/framework.

Table 4. FAIR4S competences and key skills

Skills group / Competences	Key skill
Plan and design	Plan stewardship and sharing of FAIR outputs
Capture and process	Reuse data from existing sources

[21] Key points for consultation at www.eoscpilot.eu/sites/default/files/fair4s_eoscpilot_skills_framework.pdf

[22] ibid.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

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

(Source: https://eosc-fair4s.github.io/keyskills.html)

FAIR4S identifies 9 skills groups and 59 topics relevant to professional groups involved in open science data stewardship, according to their role in service application as a user or operator. The complete skills and capability framework in which are also presented the recommended expertise by professional group and service role, is available at the EOSCpilot Skills and capability framework (D7.3)[23].

The following figure presents an overview of the FAIR4S skills groups and their relationship.

[23] www.eoscpilot.eu/content/d73-skills-and-capability-framework

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Figure 6. FAIR4S skills groups
(Source: EOSC FAIR4S)

2.1.2. FOSTER Open Science

FOSTER Open Science initiative is a result of two European Union funded projects. The FOSTER portal (www.fosteropenscience.eu) is an e-learning platform that brings together the best training resources addressed to those who need to know more about Open Science, or need to develop strategies and skills for implementing Open Science practices in their daily workflows. Many different users - from early-career researchers, to data managers, librarians, research administrators, and graduate schools - can benefit from the portal. In order to meet their needs, the existing materials are extended from basic to more advanced-level resources[24].

FOSTER wanted to specify the Open Science concept with the use of a taxonomy (Open Science Taxonomy: www.fosteropenscience.eu/resources), aiming to map the field and offer an in-depth representation of the concepts around Open Science, and also to provide a structured and consistent terminology that would reduce the descriptive conflicts in the field of Open Science and promote its development through the consistency of the open practices that relate to the Open Science terminology. Therefore, the project's goal was not only to use the taxonomy to classify the subject field, but also to take advantage of its organising ability, using the taxonomy to systematise the

[24] www.fosteropenscience.eu/about

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

educational resources hosted in the portal, which would also assist both the portal managers and the general public with the material and events review process.[25]

Among all these resources, we highlight the Open Science training toolkit (www.fosteropenscience.eu/toolkit), which offers a set of Open Science Training Courses. These courses answer some of the most common questions about putting Open Science into practice. Each course takes about 1-2 hours to work through, including practical tips on getting started with Open Science as well as providing information on discipline-specific tools and resources. All of the courses provided by FOSTER are listed and searchable in the portal at www.fosteropenscience.eu/courses.

FOSTER Open Science Learning Objectives

FOSTER defined the Open Science Learning Objectives for the main stakeholders in the Research Ecosystem (https://doi.org/10.5281/zenodo.608586). Learning Objectives are structured by Open Science Topics according to a functional Open Science Taxonomy, that accompany the main responsibilities of each stakeholder along the Research Lifecycle.

The ultimate objective is to support the integration of Open Science best practices into the daily routine of performing and supporting research, to underpin implementation of Horizon 2020 Mandate on Access to Scientific Information, and augment the "societal impact" and uptake of research, for the benefit of all stakeholders in the knowledge creation process.

Specific Learning Objectives are structured in increasing level of competence, frequently ending with successful integration of Open Science best practices in the daily research routine, facilitating self-assessment of the personal workflow.

The Learning Objectives can provide a backbone for a structured learning plan for Doctoral Schools with the ambition to train future researchers in optimizing their societal impact, alongside research excellence training, as well as preparing graduates for new and emerging research impact measures and criteria.

To maintain the coherence between an Open Science taxonomy that accompanies the target audience workflow, and to be able to match it with relevant training content (via the FOSTER Portal), the following logic is applied in structuring the Learning Objectives:

Topic → General Objective → Specific Objective → Learning Activity

For each main topic and subtopic of the Portal taxonomy (www.fosteropenscience.eu/resources), we define one general objective that can be structured in specific learning objectives. These specific

[25] Pontika, Nancy; Knoth, Petr; Cancellieri, Matteo and Pearce, Samuel (2015). Fostering Open Science to Research using a Taxonomy and an eLearning Portal. In: iKnow: 15th International Conference on Knowledge Technologies and Data Driven Business, 21-22 Oct 2015, Graz, Austria. https://doi.org/10.1145/2809563.2809571

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

learning objectives will be the basis for the course creation through a variety of possible approaches (face-to-face, blended, or e-Learning) and allow the course creator to choose which specific objectives are relevant to which target audience.

In conclusion, the list of learning objectives are defined and grouped as follows:

- Topics (following the research lifecycle);
- Core learning objectives;
- Learning objectives (as basis for a learning plan);
- Stakeholders (Doctoral Students, Researchers, Research Project Managers, Knowledge Managers and Librarians, Funding Agencies).

The full list of learning objectives are accessible at:
https://docs.google.com/spreadsheets/d/1UwsYf8fEFZzK8IPfK-7rFE3BO_VbjvOjQm3CigqBqyk/edit?usp=sharing

2.1.3. Towards FAIR data steward as profession for the lifesciences

The project "Towards a community-endorsed data steward profession description for life science research" (ELIXIR) aimed to professionalise the data steward function within the life-sciences domain in the Netherlands, with a special focus on implementation of the FAIR principles. Their results are presented in the final report "Towards FAIR Data Steward as profession for the Lifesciences"[26], and meant to be applicable outside the Netherlands and even outside the specific domain.

Based on their findings, three different data steward stakeholder-fields were defined in the data stewardship landscape depending on their interaction with specific stakeholders. All three stakeholder-fields together ensure that research carried out at institutes and in projects produces FAIR data along the data life cycle.

Each of them have their own focus and thus different data steward role: policy, research and infrastructure, as showcased in the following table.

Table 5. Data steward roles and their focus

Data steward role	**Main stakeholder field**
Data steward - policy	Policy makers, funders, management, universities board, and the

[26] S. Scholtens, M. Jetten, J. Böhmer, C. Staiger, I. Slouwerhof, M. van der Geest, C.W.G van Gelder (2019, October 5). Towards FAIR data steward as profession for the lifesciences. Report of a ZonMw funded collaborative approach built on existing expertise. Zenodo. http://doi.org/10.5281/zenodo.3471708

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

	institutes' deans. This group of stakeholders has a say in, and are ultimately responsible for, how data should be handled.
Data steward - research	Researchers and data scientists who produce data and work with the data on a daily basis with a focus on research. Researchers need to be empowered to handle data in a way, which is policy compliant without losing the power to execute their research.
Data steward - infrastructure	Data and IT infrastructure providers, e.g. IT staff, technicians and application managers. This stakeholder group provides tools to enable implementation of certain data policies and hence make it easier for researchers to manage their data in a policy-compliant way.

(Source: Adapted from: Towards FAIR data steward as profession for the lifesciences)

For all three data steward roles, eight competence areas were defined, as shown in the following table.

Table 6. Data steward competence areas

Competence area	**This concerns**
Policy/strategy	Development, implementation and monitoring of research data management policy and strategy for the research institute
Compliance	Compliance to the Netherlands Code of Conduct for Academic Practice, the Netherlands Code of Conduct for Research Integrity, the General Data Protection Regulation (GDPR), and other relevant legal and ethical standards
Alignment with FAIR data principles	Alignment to the FAIR data principles and the principles of Open Science
Services	Availability of adequate support on research data management, in staff or services
Infrastructure	Availability of adequate data infrastructure for research data management
Knowledge management	Adequate level of knowledge and skills on research data management within the institute, department or project

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Network	Obtaining and maintaining a network of aligned expertise areas and relevant departments and organisations inside and outside the institute, department or project
Data archiving	Adequate support and data infrastructure for FAIR and long-term archiving of data of the institute, department or project

(Source: Adapted from: Towards FAIR data steward as profession for the lifesciences)

For each data steward role, is detailed the function description (responsibilities and tasks), competencies (knowledge, skills and abilities) and learning objectives, as can be found at the appendix 3 of the report.[27]

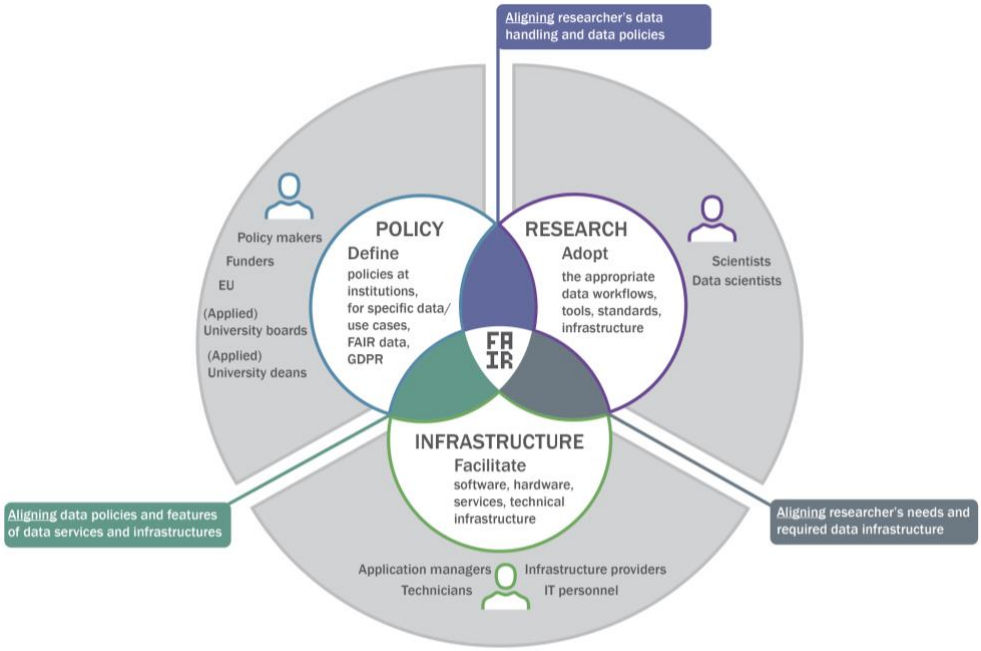

Figure 7. Implementation areas for data stewardship
(Source: Towards FAIR Data Steward as profession for the Lifesciences)

2.1.4. DeiC - Models for a Data Steward Education

The Danish e-Infrastructure Cooperation (DeiC) – www.deic.dk/en – and the Danish National Forum for Research Data Management (DM FORUM) published, in January 2020, the report "National

[27] S. Scholtens, M. Jetten, J. Böhmer, C. Staiger, I. Slouwerhof, M. van der Geest, C.W.G van Gelder (2019, October 5). Towards FAIR data steward as profession for the lifesciences. Report of a ZonMw funded collaborative approach built on existing expertise. Zenodo. http://doi.org/10.5281/zenodo.3471708

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Coordination of Data Steward Education in Denmark: Final report to the National Forum for Research Data Management (DM Forum)"[28], providing an 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. The aim is to inform future data stewardship education in Denmark.

Based on the conducted study, the report identifies four roles of the Data Steward, the Administrator, the Analyst, the Developer and the Agent of Change, and recommends Data Steward education to support them, preferably in an educational model that allows for experiential learning, collaboration with industry and innovative pedagogies that will enable a flexible education that fits with problem-based learning strategies and increase the employability of the graduates. The summary of project results and interpretations are illustrated by the four roles of a Data Steward and the three models for data stewardship education, as can be seen in Tables 7 and 8 below.

DeiC report proposes different models for Data Steward Education taking into account distinct educational and professional backgrounds. Each model illustrates the need for differently structured education, as follows:

- Students with a bachelor's degree in any field;
- Students with a PhD or equivalent experience in any field;
- Mature students need flexible continuing education in DS that will enable them to improve or supplement their knowledge and skills.

The following tables summarizes the four roles of a Data Steward and the three models for a data stewardship education.

Table 7. The 4 roles of a Data Steward

Roles of a Data Steward	
The Administrator	• Establish good practices in compliance and data privacy • Fast learner with a structured and analytical mindset • Focus on execution and seek challenges in strategic development • Implement solutions and educating end-users about them • Passion for policy and IT security • Positive attitude on cloud solutions • Risk assessments while having disciplinary knowledge • Team player with can-do attitude towards processes and operations
The Analyst	• Ensure data quality

[28] www.deic.dk/sites/default/files/Data%20Steward%20Education%20in%20Denmark_0.pdf

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

	• Enthusiasm in cloud solutions • Fast learner and innovative on building custom software and databases • Good at multitasking • Programming skills in statistical and data analysis • Seek challenges, have positive attitude towards reporting
The Developer	• FAIR principles advisor and good data planning and governance • Focus on collaboration and knowledge sharing to raise business awareness • Innovative thinking concerning master data management • Passionate about process optimization via good project management • Working in a team with compliance and data privacy experts trying to establish good practices
The Agent of Change	• Agile mindset and enthusiasm • Client and customer oriented, understanding both users, processes and operations • Developing user friendly procedures and guidelines • Educate users on ethics and the responsible conduct of research • Focus on execution of policy and strategy awareness • Passionate to implement solutions via project and change management

(Source: Adapted from: National Coordination of Data Steward Education in Denmark: Final report to the National Forum for Research Data Management (DM Forum))

Table 8. The 3 models for a data stewardship education

Models for a Data Stewardship Education		
Students with Bachelor degree *Directed corporate employment* • *One-year master programme* For students who fulfill requirements to basic programming skills, study skills, subject knowledge and academic language level. • *Two year master programme* For students who do not meet the requirement for programming and study skills follow a *pre-master's year before the master's* and then continue with the one-year master.	**Students with a PhD or equivalent** *Directed corporate or academic employment* For PhDs from any field at university faculties or part of research teams in industry. Educational programmes are a collaborative endeavor between faculties, library or knowledge centers, Centers for Information Security, Data Steward Community, Research coordinators, Project PI's, System Developers, Communication and Teaching teams, the Graduate School (for PhD training) and Human Research Ethics Committees.	**Continuing and professional education** *Directed corporate or academic employment* *Flexible master programme* • For students who have professional experience and wish to improve their DS skills but keep working full-time, or already have a master's degree in any field. • The programme is a part-time flexible master continuing education program with a prescribed period of study, that the student individually plans, and a total of 60 ECTS-points. • The student combines elements

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

• *Two-year candidatus* Traditional university candidatus combining theory, methods and internships. A dissertation in the area of DS is required.	• *Requirements:* Short courses, workshops, mentorships, online modules, summer school programmes and MOOCs, that immerse the student in the DS Community.	from established programmes, thus it is a requirement and DS courses, such as the one year master and two year candidatus, are available through the Open University.

(Source: Adapted from: National Coordination of Data Steward Education in Denmark: Final report to the National Forum for Research Data Management (DM Forum))

2.1.5. EDISON Data Science Framework (EDSF)

The EDISON Data Science Framework (EDSF), delivered as a result of the work developed by the EDISON project (http://edison-project.net/), provides a basis for the definition of the Data Science profession and enables the definition of the other components related to Data Science education, training, organisational roles definition and skills management, as well as professional certification.

The figure below illustrates the main components of the EDSF and their inter-relations that provide a conceptual basis for the development of the Data Science profession.

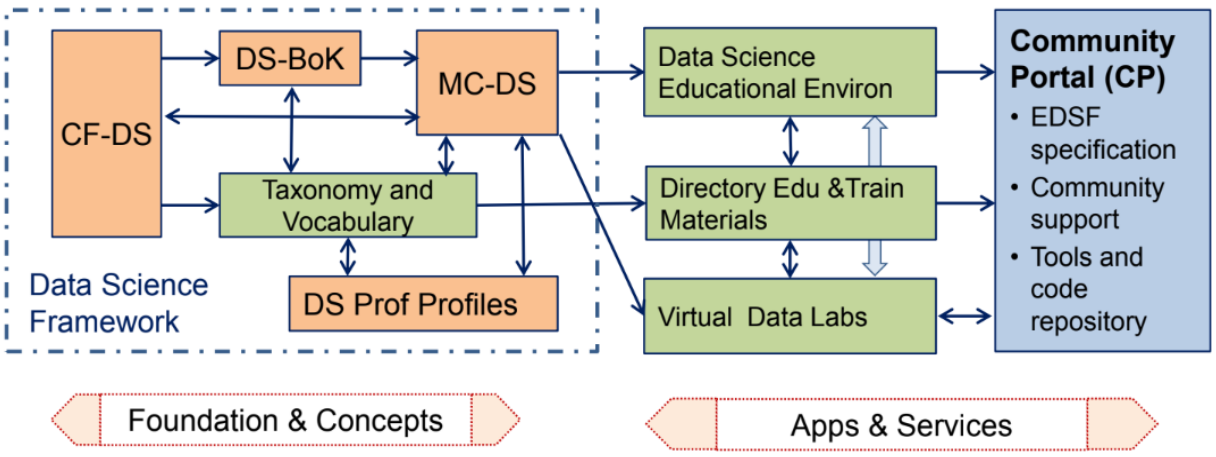

Figure 8. EDISON Data Science Framework components
(Source: EDISON Data Science Framework (EDSF))

The core components of the EDSF are the following:

- Data Science Competence Framework (CF-DS) (Part 1): http://doi.org/10.5281/zenodo.1044346

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

- Data Science Body of Knowledge (DS-BoK) (Part 2): http://doi.org/10.5281/zenodo.1044350
- Data Science Model Curriculum (MC-DS) (Part 3): http://doi.org/10.5281/zenodo.1044358
- Data Science Professional Profiles (DSPP) (Part 4): http://doi.org/10.5281/zenodo.1044364
- Data Science Taxonomies and Scientific Disciplines Classification

Data Science Competence Framework (CF-DS)[29]

The CF-DS serves as a foundation for the definition of other EDSF components. The presented CF-DS provides full and comprehensive view of the demanded Data Science competences, skills and knowledge comparing to the existing Data Science definitions that primarily cover the data analytics and software engineering competences while modern data driven enterprises and processes require advanced skills for heterogeneous data management and use of research methods to uncover full data value.

The identified Data Science Competence and Skills groups are:

- Core Data Science competences/skills groups defining profile of the Data Science related professional profiles:
  - Data Science Analytics (including Statistical Analysis, Machine Learning, Data Mining, Business Analytics, others)
  - Data Science Engineering (including Software and Applications Engineering, Data Warehousing, Big Data Infrastructure and Tools)
  - Domain Knowledge and Expertise (Subject/Scientific domain related)
- Additional common competence groups demanded by organisations:
  - Data Management and Governance (including data stewardship, curation, and preservation)
  - Research Methods for research related professions and Business Process Management for business related professions

Data Science Model Curriculum (MC-DS)[30]

The Model Curriculum (MC-DS) is built based on CF-DS and DS-BoK where Learning Outcomes[31] are defined based on CF-DS competences and Learning Units are mapped to Knowledge Units in DS-BoK. Three mastery (or proficiency) levels are defined for each Learning Outcome to allow for flexible curricula development and profiling for different Data Science professional profiles. The proposed

[29] EDISON Data Science Framework: Part 1. Data Science Competence Framework (CF-DS). https://zenodo.org/record/1044346

[30] Data Science Model Curriculum (MC-DS) (Part 3). http://doi.org/10.5281/zenodo.1044358

[31] Learning Outcomes: EDSF Release 2: Part 3. Data Science Model Curriculum (MC-DS). http://doi.org/10.5281/zenodo.1044358

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Learning outcomes are enumerated to have direct mapping to the enumerated competences in CF-DF.

As stated in the report of MC-DS, the proposed Model Curriculum intends to provide guidance to universities and training organisations in the construction of Data Science programmes and individual courses selection that are balanced according to the requirements elicited from the research and industry domains.

Data Science Professional Profiles (DSPP)[32]

Regarding Professional Profiles, at the Data Science Professional Profiles (DSPP), are proposed a set of profiles placed in four top classification groups: Managers (for managerial roles); Professionals (for applications developers and for infrastructure engineers); Technicians and Associate professionals (for operators and technicians); and Clerical support workers (for data curators and stewards).

2.1.6. Data Science: Learning and training content (GFI)

The next case is a 2019 working paper by the *German Gesellschaft für Informatik* (GFI) [Society for Informatics].[33] The paper, written by the Society's working group on Data Science/Data Literacy, details recommendations for learning content in Data Science at master level and for continued professional development. The report includes an analysis and synthesis of five different studies and frameworks for data science skills, including EDSF.[34] It discusses each briefly and presents pros and cons.

The paper itself identifies 14 fields of competences, including basic and advanced mathematics (1, 2), informatics (3, 4), cryptography and security (5), data ethics and data privacy (6), data governance (7), data integration (8), data visualisation (9), data mining (10), machine learning (11), business

[32] Data Science Professional Profiles (DSPP) (Part 4). http://doi.org/10.5281/zenodo.1044364

[33] Gesellschaft für Informatik (2019). *Data Science: Lern- und Ausbildungsinhalte*. www.plattform-lernende-systeme.de/files/Downloads/Publikationen/GI_Arbeitspapier_Data-Science_2019-12_01.pdf

[34] The others besides EDSF being acatech (2018): *Pilotphase Nationales Kompetenz-Monitoring (NKM): Bericht: Data Science. Auswahl, Beschreibung, Bewertung und Messung der Schlüsselkompetenzen für das Technologiefeld Data Science.* (acatech DISKUSSION), München 2018. www.acatech.de/wp-content/uploads/2019/02/acatech_NKM_Data_Science_WEB-2.pdf;
Ridsdale, C. et al. (2015). *Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report*. http://hdl.handle.net/10222/64578;
Chapman, P. et al. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. ftp://ftp.software.ibm.com/software/analytics/spss/support/Modeler/Documentation/14/UserManual/CRISP-DM.pdf;
Data Assessment Solutions GmbH. (2014). *IT-Skills-Studie 2014: Big Data Projekte*. White-Paper, 2014.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

intelligence (12), domain-specific applications and communicating with experts (13), and implementing data science within an organisation (14).

To develop more specific profiles, each competence field was also associated with a different level of competence: (1) understanding, (2) using, and (3) analysing. This allows the report to construct several use cases for curriculum development. In the paper, three examples are given. First, a master programme for bachelor graduates in informatics, mathematics or statistics seeking to become data science experts (A). Second, a master programme targeting bachelor graduates with a degree in a specific domain (e.g. engineering, sciences, or social sciences and humanities) wishing to gain data science competences to apply them in a domain. This is split into two different profiles and recommended minimum curricula – one for a domain specific data science programme (e.g. Management & Data Science) (B1) and another for domain-specific master programmes with data science components (B2). And third, professional development programmes for professionals seeking to upskill themselves with data science competences (not included in Table 9 due to this report's focus on formal education). For each profile, the report lists the 14 competence fields and whether students should take mandatory or optional courses. Moreover, the suggested competence level is defined.

In relation to RDM, several competence fields are relevant – mainly data ethics and privacy, data governance, data integration, as well as domain-specific applications. The latter, domain-specific applications, includes a specific mention of research data management. According to the framework, master students of both types should acquire RDM skills at the highest level in mandatory courses. Table 9 provides an overview over the main recommendations at master level in competence fields related to research data management and FAIR data.

Table 9. Selection of recommended data management competences for master students

Profile	**A) Master Data Science**	**B1) Master Data Science + Domain**	**B2) Domain-specific master with data science competences**
Competence field	M=mandatory; O=Optional L1=understanding; L2=using; L3=analysing		
(6) Data ethics and data privacy			
Data ethics	M, L3	M, L2	O, L2
Legal frameworks (GDPR)	M, L2	M, L2	O, L2
Data Privacy & Data Compliance	M, L3	M, L2	O, L2
(7) Data Governance			
Data policy	M, L1	M, L1	O, L1
Metadata management	M, L2	M, L2	M, L2
Structures and responsibilities	M, L1	M, L1	O, L1

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

(8) Data integration			
-Data collection			
Instrumentation	M, L2	M, L2	O, L1
Logging	M, L2	M, L2	O, L1
Sensors	M, L2	M, L2	O, L1
Data sources	M, L2	M, L2	O, L1
-Data pipeline			
Data flow	M, L2	M, L2	M, L2
Infrastructure and tools	M, L2	M, L2	O, L1
ETL: extract, transform, load	M, L2	M, L2	O, L1
Structured and unstructured data	M, L2	M, L2	M, L2
-Data preparation			
Data quality, data curation etc.	M, L2	M, L3	M, L2
Data wrangling, transformation, cleaning; anomaly detection	M, L2	M, L3	M, L2
Basic labelling/aggregation, analytics, metrics, segmentation	M, L2	M, L3	M, L2
Feature selection/extraction, training sets etc.	M, L2	M, L3	M, L2
(13) Domain specific applications			
Domain specific practical experience	M, L1	M, L3	M, L3
Research data management	*M, L1*	*M, L3*	*M, L3*

(Source: www.plattform-lernende-systeme.de/files/Downloads/Publikationen/GI_Arbeitspapier_Data-Science_2019-12_01.pdf, Annex 1, authors' translation)

The framework presents an interesting exercise in explicitly combining data science with domain specific recommendations, and in providing non-data scientists with suggestions how to design programmes with data science components. In terms of RDM, however, some recommendations may overlap. Research data management as a competence, for instance, is not just a domain-specific competence but may also relate to several of the other competences listed in the framework (e.g. metadata management, data policy, legal frameworks, data curation, data cleaning). The report appears to not consider RDM as an activity that concerns the whole lifecycle of research data but as a domain-specific addition to more generic data science competences.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

2.1.7. Open Data Institute (ODI) Data Skills Framework

The Open Data Institute's (ODI) Data Skills Framework[35] published in early 2020 is a high-level collection of 26 different technical and non-technical skills, grouped into six categories (Fig. 9). These groups are:

- Business (3 skills)
- Management (5 skills)
- Leadership (4 skills)
- Foundation (4 skills)
- Engineering (4 skills)
- Analysis (7 skills)

The framework is aimed at organisational leaders, learning and development specialists or "teams aiming to develop skills and collaborate better" within private or public sector organisations. It intends to serve as an assessment tool to identify existing skills within an organisation and to, subsequently, address or fill gaps by developing training programmes. It includes two schematic examples of training exercises: a one-day workshop "Open Data in a Day", which addresses six basic skills required for Open Data; and a course titled "Strategic Data Skills" that covers 19 skills with a focus on "[deriving] insight from data and then find[ing] ways to present insights to leaders."

[35] Open Data Institute (2020). Data Skills Framework. http://theodi.org/wp-content/uploads/2020/05/2020-05-Data-Skills-Framework.pdf

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Figure 9. ODI Data Skills Framework
(Source: http://theodi.org/wp-content/uploads/2020/05/2020-05-Data-Skills-Framework.pdf)

In terms of synergies with research data management, several skills may link to elements of the research data lifecycle or FAIR data, such as standardising, cleaning and linking data as part of the "engineering" skills group. Others, such as governing access and working ethically may relate to accessing data, legal knowledge, data publishing and access levels within an RDM context. However, given that the skills are not further detailed within the framework, an in-depth comparison is not easily possible.

2.1.8. Belmont Forum - Data skills curricula framework

The Belmont Forum (www.belmontforum.org) is an international partnership that mobilizes funding of environmental change research and accelerates its delivery to remove critical barriers to sustainability. Forum members and partner organizations work collaboratively to meet this challenge by issuing international calls for proposals, committing to best practices for open data access, and providing transdisciplinary training. To that end, the Belmont Forum is also working to enhance the

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

broader capacity to conduct transnational environmental change research through its e-Infrastructure and Data Management initiative (e-I&DM - www.belmontforum.org/eidm/)[36].

The e-I&DM performed a skills gap survey[37], resulting in a Skills Gap Analysis which identified data complexity, lack of standards and exchange vocabulary, and data management as vital barriers to sharing and reusing data effectively.

- Skills Gap Analysis and Curricula for Data Intensive Digital Skills in Global Environmental Change Research: https://zenodo.org/record/1404652#.XlVRU0p7nVg

After further peer consultation, including a workshop, a curricula framework was proposed for the development of digital skills for data intensive global change research, being endorsed by the Belmont Forum in November 2017 (www.belmontforum.org/wp-content/uploads/2017/12/PR_Belmont_Forum_Endorses_Curricula_Framework.pdf).

The recommended Curricula Framework core modules are designed to standardize skills of domain scientists specifically to make data handling more efficient, research more reproducible and data more shareable – including visualisations for end-users. The five core skills emphasized include programming, particulars of environmental data, visualisation, data management and interdisciplinary data exchange[38].

A number of optional modules are suggested for more established researchers which would be useful introductions to widen their data skills and provide new options for examining and processing data, such as machine learning and object orientated programming. Two additional modules are briefings aimed at Principal Investigators, providing an overview of data management and skills.

Each module is targeted according to specific educational and professional backgrounds, starting from the first year PhD until Principal Investigators, as follows: A = first year PhD, B = final year PhD, C = postdoc, D = mid-career, Principal Investigators[39].

Full curricula framework is presented in the following report "Looking ahead: extended recommendations for digital skills in data-intensive global environmental change research" at www.belmontforum.org/wp-content/uploads/2019/10/Outcomes_AT4_2017_final.pdf.

The following table showcases the structure of "Data Skills Curricula Framework".

[36] www.belmontforum.org/about/

[37] The skills survey was sent to a range of Belmont Forum contacts and completed in November and December 2016 by 164 people who work in a range of institutions around the world. Of these, 76% were employed by government or universities and 81% came from either North America or Europe, with 6% of respondents from Africa.

[38] www.belmontforum.org/wp-content/uploads/2017/12/PR_Belmont_Forum_Endorses_Curricula_Framework.pdf

[39] www.belmontforum.org/wp-content/uploads/2019/10/Outcomes_AT4_2017_final.pdf

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Table 10. Data Skills Curricula Framework

<table>
<tr><td colspan="2">Data Skills Curricula Framework

Curricula goals:
● Data handling more efficient
● Research more reproducible
● Ease data sharing</td></tr>
<tr><td colspan="2">Curricula structure:</td></tr>
<tr><td>Core modules</td><td>1. Programming for data intensive research (for those who already code) (A)
2. Environmental data: expectation and limitations (A)
3. Introduction to visualising environmental data (A and B)
4. Data management (A and B)
5. Interdisciplinary data exchange (B, C and D)</td></tr>
<tr><td>Optional modules</td><td>● Software development ideas for scientific coding (C and D)
● Object-orientated programming (C and D)
● Introductory data science topics (C and D)
○ Databases
○ Machine Learning
● Data organisation (A to D)
○ Workflow
○ Code sharing facilities</td></tr>
<tr><td>Principal investigators</td><td>DMPs and data repositories
● Prioritising data management
○ Funder requirements
○ DMPs as living documents
○ Resouring
Overview of skills for data intensive research
● Team roles
○ Software engineering is a specialism
○ How to recognise contributions of “non-publishers”</td></tr>
<tr><td colspan="2">Targets:
A = first year PhD, B = final year PhD, C = postdoc, D = mid-career, Principal Investigators</td></tr>
</table>

(Source: Adapted from: Looking ahead: extended recommendations for digital skills in data-intensive global environmental change research)

FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

e-Infrastructures & Data Management Toolkit

Another relevant output from the Belmont Forum is the e-Infrastructures & Data Management Toolkit (https://bfe-inf.github.io/toolkit/index.html), which provides training and educational resources for data discovery, management, and curation across the globe, in support of an international collaborative effort to enable open access to scientific data.

2.2. Analysis of skills frameworks

The in-depth presentation of different skills frameworks in Chapter 2.1 already reveals a heterogeneous landscape. The overall objective of this additional analysis is to identify potential strengths and weaknesses related to their use or adaptation in higher education.

Following the discussion in Chapter 1.1, we compared and analysed the different skills frameworks along a set of characteristics (Annex 1).

---

*All results are displayed in **Annex 1***

Table: Comparison of skills frameworks.

***Qualification level**: Does the framework relate to different qualification levels, in particular first, second and third cycle of higher education (Bachelor, Master, Doctorate)?*

One framework covers undergraduate education (EDSF). Three frameworks are designed for education at Master level (EDSF, DEIC, GFI). The remainder are targeted at professionals or PhD level.

***Professional group:** What professional group is the framework mainly addressing (e.g. data stewards, data scientists, researchers)?*

The different frameworks cover a wide range of professional groups in the wide spectrum of research data management/stewardship, data science, and data infrastructures. There is a mix of frameworks aiming to support RDM and data science in an academic research context and frameworks targeting the private sector. Some frameworks combine recommendations for graduates working in the academic or private sector.

***Main field/domain**: Is the framework related to a particular field or domain?*

Most frameworks target data science or data stewardship in a more generic fashion (FAIR4S, DEIC, EDSF, GFI, ODI) implying that specific disciplinary practices are not strongly addressed. Two data science frameworks (EDSF, GFI) include linkages to domain research but don't further specify these competences or skills. Two frameworks (Elixir, Belmont) address domains, life sciences on the one

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

hand, and environmental sciences on the other. However, both are not targeted at the Bachelor or Master level. FOSTER can be considered a separate case as it covers a broad range of open science skills.

***Elements**: What main elements does the framework consist of? E.g. model curricula, body of knowledge.*

Overall, a large array of supporting research is present in the sample. Several frameworks are based on research on professional functions or job descriptions (FAIR4S, EDSF, ELIXIR, DEIC). EDSF is certainly the most extensive framework as it contains a framework, the body of knowledge, model curricula and other resources. FAIR4S stems from reports developed in the EOSCpilot project and includes detailed skills and role profiles.

***Learning outcomes or other aspects supporting course development:** Does the framework define learning outcomes, model curricula or other insights that help users in designing training programmes or courses?*

Model curricula are included in DEIC, EDSF, GFI, and Belmont. FOSTER is linked to an extensive collection of training resources. FAIR4S includes skills groups and competences, including recommendations for different levels and function groups. ELIXIR identifies competencies (knowledge, skills and abilities) and learning objectives. ODI includes high level recommendations for training and education courses.

***Addresses FAIR data:** Does the framework make an explicit mention of the FAIR principles/FAIR data?*

FAIR4S in particular is structured around the research data lifecycle and FAIR data stewardship. FAIR4S, FOSTER, ELIXIR, DEIC, and EDSF mention FAIR. ODI, GFI, and Belmont do not explicitly refer to FAIR data.

***Inclusion of research data lifecycle/RDM**: Does the framework include research data management skills based on different stages of the research data lifecycle?*

All frameworks address data management at varying levels. FAIR4S, ELIXIR, and DEIC were explicitly developed to address RDM skills. EDSF includes a specific data management skills group and addresses several stages of the data lifecycle. GFI and ODI include several aspects which are usually seen as part of the data lifecycle, but do not provide a holistic perspective on it. GFI mentions RDM as a discipline-specific skill without further details.

***Addresses Open Science**: Does the framework relate to other skills or elements related to Open Science?*

FOSTER in particular is a comprehensive framework/taxonomy that includes RDM as a component in broader Open Science skills (open access, evaluation, policies etc). FAIR4S, ELIXIR, EDSF and DEIC

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

refer to Open Science skills more generically, without specific definition beyond RDM and FAIR data. GFI, ODI and Belmont do not further refer to Open science.

Examples for practical implementation/usage of the resource: *Are there known examples of real life usage of the framework?*

Identifying whether a framework has been used in a specific context is a difficult task. Using the approach in this report, it was not possible to establish a reliable measure of the frequency at which higher education institutions employed the frameworks for curriculum or programme design. This lack of evidence, however, does not allow the conclusion that the frameworks are not used within this context.

With regard to the utilisation of the different frameworks across each other, the authors are aware that terms4FAIRskills, for instance, is built on FAIR4S. EDSF was also consulted in the development of FAIR4S, ELIXIR and DEIC. It is also possible that these frameworks have an indirect impact, e.g. through guidelines or frameworks at the national level.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

3. Overview and analysis of education and training programmes and initiatives

Mirroring the previous section, Chapter 2 briefly introduces different education and training programmes and initiatives. We cover bottom-up initiatives such as RDA, e-infrastructures such as OpenAIRE, international initiatives like GO-FAIR and CODATA, projects such as terms4FAIRskills, and a set of Europe-based research infrastructures developing more domain-specific RDM. For each programme or initiative we provide a summary of their main objectives, elements and coverage of skills and competences. The main goal of this additional section and analysis is to identify potential assets related to their use or adaptation in higher education.

3.1. Research Data Alliance

The Research Data Alliance (RDA) has as its mission to "build [...] the social and technical bridges to enable the open sharing and re-use of data", as stated in their Webpage (www.rd-alliance.org/about-rda).

RDA has a grass-roots, inclusive approach covering all data lifecycle stages, engaging data producers, users and stewards, addressing data exchange, processing, and storage. It has succeeded in creating the neutral social platform where international research data experts meet to exchange views and to agree on topics including social hurdles on data sharing, education and training challenges, data management plans and certification of data repositories, disciplinary and interdisciplinary interoperability, as well as technological aspects[40].

Webinars

Among their activities, RDA promotes several webinars[41] and events open to the community, covering the main topics related to research data management.

RDA is a pragmatic and agile organization, and through the work of its Working and Interest Groups, directly and logically tackles numerous data infrastructure challenges.

Interest Groups

Professionalising Data Stewardship IG

One of the RDA Interest Groups in formation is the "Professionalising Data Stewardship IG" (www.rd-alliance.org/groups/professionalising-data-stewardship).

[40] www.rd-alliance.org/get-involved/value-rda-you

[41] www.rd-alliance.org/plenaries-events/webinars/past-webinars

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

The idea to create an interest group was raised in a Birds of a Feather (BoF) session during the 14th Plenary (Professionalizing data stewardship in the RDA community), and a first informal group of interested people was established. At the 15th Virtual Plenary, a second BoF session was organised (Challenges in professionalising data stewardship: how can the RDA contribute to solving them?). They are currently in the process of submitting the charter to acquire endorsement as a formal Interest group.

The outputs from both sessions (Slides and session notes) are available for consultation at:

- Professionalizing data stewardship in the RDA community www.rd-alliance.org/professionalizing-data-stewardship-rda-co
- Challenges in professionalising data stewardship: how can the RDA contribute to solving them? www.rd-alliance.org/challenges-professionalising-data-stewardship-how-can-rda-contribute-solving-themmmunity

This Interest Group will focus on the following challenges: a business case for data stewardship, data stewardship terminology, the integration of data stewardship across an organisation, job profiles for data stewards, training for data stewards, career tracks for data stewards, networking and knowledge exchange, and certification.

Education and Training on handling of research data IG

Another relevant interest group is the "Education and Training on handling of research data IG" (www.rd-alliance.org/groups/education-and-training-handling-research-data.html), created in the context of increasing volumes of data created by researchers and the strengthening of requirements for research data management and data sharing, that has created an increasing demand for a new and evolving set of competencies and skills for researchers who create and use the data, and the growing cadre of professionals who support them.

The objective of this IG is the exchange of information about existing developments and initiatives and promotion of training/education to manage research data throughout the data lifecycle. Concretely, it will make the case for creating taxonomies of the skills required by different group of data management specialists/professionals and elaborating reference models as a basis to:

- enable the setting of quality standards for appropriate education and training programmes aimed at researchers and the professionals that support them, at all career stages;
- encourage the recognition of data skills amongst employees, employers, and professional bodies;
- prepare the ground for practical applications applying these standards in educational environments.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

3.2. OpenAIRE

OpenAIRE, the European Open Science Infrastructure, for open scholarly and scientific communication, among their activities aims to foster the culture change for Open Science practice, through the Open Science Helpdesk service, a comprehensive resource that covers all Open Science publishing issues, developed through an integrated approach with all OpenAIRE stakeholders, contributing to perform an effective helpdesk support system. The main goal of the Open Science Helpdesk is to run a distributed and multifaceted support system, to assist the OpenAIRE stakeholders in all matters related to Open Science, which includes the direct support to institutions, researchers and project coordinators from a network of National Open Access Desks (NOADs)[42]. OpenAIRE NOADs are responsible to facilitate train-the-trainers approaches for all stakeholders, synthesising national training initiatives, and ensuring training for researchers on Open Science topics, FAIR data, Copyright and IPR issues.

OpenAIRE is providing to the community a set of support materials and training activities targeted at a range of stakeholders, such as researchers, content providers, research managers, funders, research communities and innovators to acquire the skills and competencies on the practical implementation of Open Science.

The support and training materials available through the OpenAIRE portal (www.openaire.eu/support) targets different stakeholders and can be found in several formats, such as guides for researchers on Open Science related topics (Research Data Management, Legal Issues dealing with research data), Factsheets on H2020 mandates and Open Science policies, Use Cases narratives with scenarios of Open Science services offered by OpenAIRE targeting different users, materials from the Webinars and Workshops on various Open Access and Open Science topics, and a Toolkit for Policy Makers to assist the design and adoption of Open Science policies aligned with EU developments in the field.

The following main topics are covered:

- How to comply with H2020 mandate: for research data and publications
- Research Data Management
    - How to identify and assess RDM costs
    - How to make data FAIR
    - How to find a trustworthy data repository
    - How to create a DMP for H2020 projects
    - Data formats for preservation
    - How to deal with non-digital data
    - How to deal with sensitive data

[42] www.openaire.eu/contact-noads

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

- Raw data, backup and versioning
- Legal Issues
  - Protection of research data by intellectual property rights
  - How to license research data
  - How to reuse research data

Community of practice for training coordinators and managers

In May 2018 OpenAIRE launched a Community of Practice (CoP)[43] for training coordinators and managers that currently include over 60 members representing major e- and research infrastructures and national training initiatives. The CoP is an informal network to share training experiences initiated by a group of people who coordinate training programmes of research and e-infrastructures aiming to map out the training activities of various pan-European, EOSC-related initiatives and strengthen their training capacity. The monthly calls and face-to-face sessions at DI4R, Open Science FAIR and EOSC stakeholders forum provided a good platform to discuss how to make training materials FAIR, open badges for training activities, how to measure impact of training, build training skills and competencies more effectively and other topics. CoP contributed to the launch of EOSC Skills and Training Working Group and many CoP members are part of the EOSC working group. This Community of practice is in fact a relevant initiative for synergies regarding the development of the FAIR data science curricula for HEI within FAIRsFAIR project activities.

3.3. GO-FAIR

Some of GO FAIR Implementation Networks (IN)[44] - a consortium committed to defining and creating specific materials and tools as elements of the Internet of FAIR Data and Services (IFDS)[45] - among their objectives, the implementation of the FAIR data principles is one of the key elements of their activities, promoting the development of training materials for specific research communities and domain areas[46]. Focusing on the training area, there are three IN in preparatory phase (Season Schools, Training Curriculum, Training Frameworks), which could be relevant to follow, as they could develop useful resources for the community, helping to promote the development of competences around FAIR data principles.

[43] www.openaire.eu/cop-training

[44] www.go-fair.org/implementation-networks/

[45] www.go-fair.org/resources/internet-fair-data-services/

[46] www.go-fair.org/implementation-networks/overview/

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

3.3.1. Data Stewardship Competence Centers Implementation Network (DSCC-IN)

One of the current GO-FAIR Implementation Networks is the Data Stewardship Competence Centers (DSCC), composed mainly by individual national DSCC chapters that can reflect local setup needs and ensure a functioning network of DSCCs on the national or regional level[47].

This network structure in the form of a GO-FAIR Implementation Network IN aims to establish a broader consensus on the requirements for sustainable DSCC infrastructures and on professional data stewardship skills.

The overall aim of this IN is networking and establishing pragmatic cooperation, reuse of assets and convergence between national and international DSCCs at universities or other research performing institutions (RPI). The DSCC IN partners are committed to mutual support for continuous and collaborative improvements to research data services, especially in their implementation of the FAIR principles. Strong emphasis will be put on joint development of skills, requirements, services and the underlying protocols, formats, templates, and standards that will drive interoperability and reuse through practical convergence, especially across geographical and disciplinary borders.

The DSCC-IN Manifesto is available at:
www.go-fair.org/wp-content/uploads/2020/01/DSCC-IN-Manifesto-20191210.pdf

3.4. FOSTER Open Science

The FOSTER portal (www.fosteropenscience.eu), as previously mentioned, is an e-learning platform that delivers training resources on Open Science and related topics, targeting several stakeholders, from early-career researchers, to data managers, librarians and research administrators. In order to meet their needs, the existing materials are extended from basic to more advanced-level resources[48].

As main activities, FOSTER project (delivered face-to-face training events, blended and e-learning courses; developed efforts for the consolidation of an Open Science trainers network involving the disciplinary communities of humanities, social sciences and life sciences; and created advanced-level training resources including a multi-module Open Science training toolkit and an Open Science training handbook, both being detailed below.

The Open Science training toolkit[49] is composed of a set of courses on Open Science (OS), which answer some of the most common questions about putting OS into practice, including practical tips on getting started with OS as well as providing information on discipline specific tools and resources. This toolkit is composed by the following courses:

[47] www.go-fair.org/implementation-networks/overview/dscc/

[48] www.fosteropenscience.eu/about

[49] Open Science training toolkit. www.fosteropenscience.eu/toolkit

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

- What is Open Science?
- Best practices
- Managing and Sharing Research Data
- Open Source Software and Workflows
- Data Protection and Ethics
- Open Licensing
- Open Access Publishing
- Sharing Preprints
- Open Peer Reviews
- Open Science and Innovation
- Use Open Data in Teaching
- Assessing the FAIRness of data

The Open Science Training Handbook[50] is an open knowledge and educational resource oriented to practical teaching, focused to show how to spread Open Science most effectively, and aimed to support educators of Open Science. It is presented as a key resource and a first step towards developing Open Access and Open Science curricula. Supporting and connecting an emerging Open Science community that wishes to pass on their knowledge as multipliers, the handbook will enrich training activities and unlock the community's full potential. Furthermore, this resource is already translated into 4 different languages (available in English, Portuguese, Spanish and Italian, under translation to French, Greek and German).

3.5. CODATA

CODATA (https://codata.org/), the Committee on Data of the International Science Council (ISC), aims to promote global collaboration to advance Open Science and to improve the availability and usability of data for all areas of research. CODATA supports the principle that data produced by research and susceptible to be used for research should be as open as possible and as closed as necessary. CODATA works also to advance the interoperability and the usability of such data, meaning that research data should be FAIR (Findable, Accessible, Interoperable and Reusable).

One of their working groups, the FAIR Data Training working group (https://codata.org/initiatives/working-groups/fair-data-training/), aims to advance education and training in the science of data and data stewardship. Important activities include the CODATA-RDA School of Research Data Science[51] and the regular CODATA China Training Workshops. As well as convening such training activities, CODATA is concerned to assist with advancing and coordinating

[50] Open Science Training Handbook. www.fosteropenscience.eu/content/open-science-training-handbook

[51] https://codata-rda-datascienceschools.github.io/

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

understanding of training requirements and formalising competence frameworks for data science and data stewardship.

The CODATA-RDA School of Research Data Science is a programme of short courses equipping young researchers globally with core data skills, which curriculum covers the following main areas[52]:

- Principles and practices of Open Science
- FAIR data and using data ethically
- An introduction to research data management and the research data lifecycle
- How to prepare a data management plan
- Introduction to persistent identifiers and licensing
- A grounding in how to compile bibliographies
- How to find data more efficiently
- How to publish data
- Software and data carpentry
- How to use Github
- Introduction to theory and practice of visualisation
- Introduction to machine learning
- Grounding in use of computational research infrastructures

CODATA and RDA are partners in FAIRsFAIR. The latest instances of the CODATA-RDA School of Research Data Science with two interconnected tracks for data scientists and data stewards are supported by FAIRsFAIR.

3.6. terms4FAIRskills

terms4FAIRskills is a community initiative (https://terms4fairskills.github.io/) related to CODATA and the outputs generated by EOSCpilot (specifically FAIR4S) which aims to create a formalised terminology that describes the competencies, skills and knowledge associated with making and keeping data FAIR. It builds on FAIR4S with the intention to refine the definitions and terminology used in FAIR4S. When mature, this terminology will apply to a variety of use cases, including:

- To assist with the creation and assessment of stewardship curricula;
- To facilitate the annotation, discovery and evaluation of FAIR-enabling materials (e.g. training) and resources;
- To enable the formalisation of job descriptions and CVs with recognised, structured competencies.

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.

[52] ibid.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

3.7. LEARN project

The LEARN project (2015-2017) was aimed to create resources to help Research Performing Institutions to manage their research data. Among its outputs, the LEARN project produced a "Toolkit of Best Practice for Research Data Management"[53], from which we highlight two case studies on training and skills development on Research Data Management.

Training early career researchers

The purpose of this case study was to look at the role of Research Performing Organisations in skills development for early career researchers, and to set those needs in the context of the growing importance of research data and the emerging role of the data steward. Also highlights the need for research data training, identified as a high priority from several research universities, in order to deploy the needed training for data stewardship, on generic skills and subject-specific. As an example of training activities, is presented the League of European Research Universities (LERU) doctoral summer school as a model for the training seminars needed to deliver generic skills and subject-specific insights into the activity of data stewardship, having as main topics, the FAIR data principles, Data Management Plans, Licensing, Open Science and infrastructures.

Training subject librarians in Research Data Management

This case study showcases the activity of University College London (UCL) Library Services on support and training for academics, researchers, students and staf, focusing on the Research Data Support Officers which coordinate the Research Data Management (RDM) advocacy and support across the institution. In order to inform and train the library staff on Research Data Management and Data Management Plans, a set of workshops were conducted (2015 and 2016). Subsequent to formal RDM training, librarians were given the opportunity to apply their knowledge by actively contributing to a new Working Group, the Library RDM Working Group, aiming to support the two Research Data Support Officers with discipline-specific knowledge and essential staff resources for short-term projects, creating discipline-specific resources to help researchers throughout their research projects. Such resources include RDM guidance, metadata standards, data repositories and ethics guidelines, and also a second project to design a course template to introduce research students to RDM. It was planned to continue with the RDM training within UCL Library Services to ensure that subject liaison and site librarians' knowledge stays up to date.

[53] LEARN Toolkit of Best Practice for Research Data Management. http://learn-rdm.eu/wp-content/uploads/RDMToolkit.pdf

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

3.8. Research Infrastructures

3.8.1. ELIXIR

ELIXIR is an intergovernmental organisation that brings together life science resources from across Europe. These resources include databases, software tools, training materials, cloud storage and supercomputers, so that researchers can more easily find, analyse and share data, exchange expertise, and implement best practices.

Regarding training materials, as an example of an ELIXIR resource, TeSS (https://tess.elixir-europe.org/) is the online training portal that gathers life science training materials and training courses from across Europe.

One of the activities that this organisation carries out, is the ELIXIR's Training Platform (https://elixir-europe.org/platforms/training), aimed to develop professional skills for managing and exploiting data, namely:

- Establish sustainable training infrastructure supported and adopted by all ELIXIR Nodes;
- Deliver a coherent, high-quality and impactful ELIXIR training program;
- Building on and complementing training programmes in the 23 ELIXIR Nodes.

Despite ELIXIR's focus on life sciences, their training materials regarding research data management and FAIR data principles are considered relevant for other domain areas.

Regarding the FAIR principles, ELIXIR has identified the need to build capacity in FAIR data stewardship, and then has contributed to improve the competences on this area, developing and delivering training activities in data stewardship for stakeholders within research data management, as follows:

- Researchers: create, manage, analyse and share FAIR data;
- Data Stewards: provide support for research teams, data repositories, data services;
- Reviewers and funders: assess DMPs, assess FAIRness of datasets, data resources in publications, grant proposals and reports;
- Trainers: teach how to apply FAIR principles and how to perform good Data Stewardship.

As an example, ELIXIR-NL organised a training[54] event on the technical aspects of data stewardship, which learning outcomes were the following:

- Explore the ELIXIR Data stewardship wizard[55] and see how one can employ and adopt it towards a community or research project;

[54] https://www.dtls.nl/courses/elixir-nl-training-technical-data-stewardship/

[55] https://ds-wizard.org/

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

- Learn about the principles of data modeling and making data FAIR by means of linked data;
- How to approach their own data and can set first steps on the FAIRification process;
- How FAIR data can be used in computational workflows and how to gather provenance data while working with data assets.

3.8.2. CESSDA

CESSDA Consortium of European Social Science Data Archives (https://www.cessda.eu/) – has as part of its mission to facilitate teaching and learning in the social sciences. Their Training Working Group is responsible for supporting continuous learning and training of its Service Provider staff and the social science user community, designing various training materials and resources for finding, managing and preserving data.

These materials (guides and webinar recordings), listed in a dedicated portal page (www.cessda.eu/Training), are focused on training for data producers, data users, data professionals and CESSDA Service Provider staff, covering the following main topics: discovering and using data, managing research data, preserving data and using CESSDA's tools and services.

An example of a training resource is the *Data Management Expert Guide*[56], designed by European experts to help social science researchers make their research data FAIR and to create and update their DMPs.

3.8.3. SSHOC Training

SSHOC (Social Sciences & Humanities Open Cloud) (www.sshopencloud.eu) is a project funded by the EU framework programme Horizon 2020 and unites 20 partner organisations and their 27 associates in developing the social sciences and humanities area of EOSC.

One of the SSHOC main goals is to maximise the efficiency and effectiveness of data re-use. SSHOC will apply Open Science practices and FAIR principles to data management. Some of their activities to achieve this goal, are the SSHOC training and educational resources (www.sshopencloud.eu/training), focused on providing training, advice, and educational resources, aiming to enable data producers, data users, and data professionals to gain maximum benefit from the SSH area of the EOSC.

Within their training activities, in 2020 and 2021, SSHOC will run regular webinars and workshops on themes including:

- Data Protection and the GDPR
- Data Stewardship and RDM in theory and practice

[56] www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

- Data Citation principles and practice
- Data Science for the Social Science and Humanities
- Data Science for Heritage Science
- Text Mining for the Social Sciences and Digital Humanities

And also provides a Training Toolkit (https://training-toolkit.sshopencloud.eu/), a resource for social sciences and humanities trainers, that provides an inventory of educational materials covering topics such as research data management, FAIR data, Open Science, programming and didactics. The Training Toolkit covers materials directed at different target audiences including researchers, service providers and data stewards and trainers themselves.

3.8.4. ENVRI Community Training Platform

The ENVRI community (https://envri.eu/) is a community of Environmental Research Infrastructures, projects, networks and other diverse stakeholders interested in environmental Research Infrastructure matters.

The ENVRI community is right now supported by the ENVRI-FAIR project (http://envri-fair.eu/) building FAIR Research Infrastructure services and connecting the community to European Open Science Cloud.

One of the project work packages - Training for & by the ENVRI community - aims to provide training to ENVRIs and key ENVRI Community stakeholder groups about:

- the FAIR principles;
- how to implement them in RI services and data management activities at data center level;
- how to evaluate the degree of implementation using FAIR metrics;
- relevant legal and policy requirements.

Courses are available for the community at the ENVRI Training platform (https://training.envri.eu/), from which the following are an example of courses related to the adoption of FAIR data principles:

- Towards ENVRI Community International Winter School DATA FAIRness (https://training.envri.eu/course/view.php?id=46)

Resources from ENVRI summer school on the use of FAIR data within the Environmental and Earth sciences research community. The course focuses on the creation and reuse of FAIR data and services in the Environmental and Earth sciences. It is built as a five-day summer school where leading scientists will address various topics from different perspectives.

- International Summer School Data FAIRness in Environmental & Earth Science Infrastructures: theory and practice (https://training.envri.eu/enrol/index.php?id=43)

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

3.8.5. ExPaNDS

The European Open Science Cloud (EOSC) Photon and Neutron Data Service (ExPaNDS) project (https://expands.eu/) aims to expand, accelerate and support the data management and data services provided through the EOSC for major national Photon and Neutron Research Infrastructures (PaN RIs) in delivering world-leading science.

As stated in their webpage, ExPaNDS will make the majority of PaN RIs data 'open' following the FAIR principles (Findable, Accessible, Interoperable, Reusable) according to the user's needs, and to harmonise efforts to migrate facility's data analysis workflows to EOSC platforms enabling them to be shared in a uniform way.

ExPaNDS therefore seeks to, among other activities, enable EOSC services and to provide coherent FAIR data services to the scientific users of national Photon and Neutron sources.

In order to promote a faster adoption of best practices by an enlarged number of scientific users, one of the ExPaNDS Work Packages, focused on training activities through EOSC platform, will organise in cooperation with EOSC related activities, workshops and deliver training materials through the e-learning platforms made available on the EOSC[57]. These training will address:

- Service provision
- User training in FAIR principles
- Data stewardship
- Data management
- Data analysis services integrated into the EOSC

3.8.6. DARIAH

The Digital Research Infrastructure for the Arts and Humanities (DARIAH) (www.dariah.eu) aims to enhance and support digitally-enabled research and teaching across the arts and humanities[58]. As stated in their webpage, training and education is central to DARIAH mission, which results in the development of training and education activities for their community. Several training resources are provided in the Training and Education page (www.dariah.eu/activities/training-and-education/), coming from different platforms and projects, namely: DARIAH Campus (https://campus.dariah.eu/), DARIAH Teach (https://teach.dariah.eu/), PARTHENOS Training (https://training.parthenos-project.eu/), and DH Course Registry (https://dhcr.clarin-dariah.eu/). Although the majority of these training materials are focused on the Digital Humanities (DH) domain, the development of skills on RDM, FAIR Principles and Open Science related topics are also a concern.

[57] https://expands.eu/work-packages/

[58] www.dariah.eu/about/dariah-in-nutshell/

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

An example is the Winter School "Shaping new approaches to data management in arts and humanities"[59] in November 2019. The main objective was to introduce scientific and academic communities in the arts and humanities to the principles and practices of responsible research and Open Science. It was composed by seven sessions on data in the humanities (specificity of data in the Humanities, assessing the FAIRness of data, FAIR principles, depositing data), data and software citation practices, copyright and (open) licensing, data management plans, publishing practices and open research notebooks. A full synthesis of this Winter School is published on the DARIAH-Campus platform (https://campus.dariah.eu/) and can be found in open access at https://campus.dariah.eu/resource/ws2019.

3.8.6. PARTHENOS Project

PARTHENOS Project (www.parthenos-project.eu), as stated on their webpage, stands for "Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies". One of their activities is the development of training modules and resources for researchers, educators, managers, and policy makers who want to learn more about research infrastructures and related issues and topics, being available at PARTHENOS training platform (https://training.parthenos-project.eu/). Despite of these resources being focused mainly on topics related to digital humanities, digital heritage and research infrastructures, one of the training modules is focused on RDM, FAIR Principles and Open Science, being comprised by ten sections specifying different domains, as presented below:

Training module: "Manage, improve and open up your research and data"[60]

This module looks at emerging trends and best practices in RDM, quality assessment and Intellectual Property Rights issues, through the following main sections:

- Introduction to Research Data Management
- The FAIR Principles
- Managing Cultural Heritage Assets
- Reviews of practices within Cultural Heritage Institutions
- Data Management Planning
- Data Quality Assessment
- The Core Trust Seal
- Ethics and Research
- Open Data, Open Access and Open Science
- Research Infrastructures and Data Policy

[59] https://desirschool.sciencesconf.org/

[60] http://training.parthenos-project.eu/sample-page/manage-improve-and-open-up-your-research-and-data/

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

According to the information in the module webpage, the learning outcomes are the following:

- Understand and describe the FAIR Principles and what they are used for,
- Understand and describe what a Data Management Plan is, and how they are used,
- Understand and explain what Open Data, Open Access and Open Science means for researchers,
- Describe best practices around data management,
- Understand and explain how Research Infrastructures interact with and inform policy on issues around data management.

3.8.7. PaNOSC

The PaNOSC project, Photon and Neutron Open Science Cloud (www.panosc.eu), has as its mission to contribute to the realization of a data commons for Neutron and Photon science, providing services and tools for data storage, analysis and simulation, for the many scientists from existing and future disciplines using data from photon and neutron sources. This project is working to make FAIR data a reality, developing common policies, strategies and solutions in the area of FAIR data policy, data management and data services, having as one of its objectives to make scientific data produced at Europe's major Photon and Neutron sources fully compatible with the FAIR principles[61].

One the PaNOSC work Packages - Data Policy and Stewardship - is about Data Policies and Data Management Plans. Among its objectives, one of them is to propose a new framework for Data Policies taking into account the FAIR principles. As stated in the PaNOSC website, FAIR principles will be studied in depth with the help of organisations specialised in interpreting and applying FAIR principles and GDPR to scientific research data (GO-FAIR, FAIRsFAIR). A set of guidelines will be produced by all PaNOSC RIs, taking into account the specific needs of the Photon and Neutron community[62].

The new PaNOSC data policy framework, according to the information in its webpage, will be developed in line with the EOSC activities on data policy harmonization and shall ensure that FAIR principles are applied as broadly as possible. Within this process, guidelines will be published for DOIs, long term archiving, and legal aspects linked to FAIR data, GDPR and data federation, as well the development of a Data Management Plan template for experiments performed at the PaNOSC research infrastructures[63].

[61] www.panosc.eu/about-panosc/

[62] www.panosc.eu/work-packages/work-package-2-data-policy-and-stewardship/

[63] www.panosc.eu/data/panosc-data-policy-framework/

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

3.9. National initiatives

Several projects and initiatives are being developed at the national level to support RDM, FAIR data and Open Science training and data stewardship education. Within the purpose of this report we have supplemented the analysis carried out for competence frameworks and training programmes with a brief analysis of a sample of relevant national training initiatives. The main purpose of this complementary analysis aimed to explore potential assets related with the use of any competence frameworks or mention of training providers and also related with the topics coverage. The following initiatives were identified and analysed in the context of this additional collection of information and does not intend any complete representation of EU Member States initiatives.

Denmark

Data Stewardship Education in Denmark

Zenodo community: www.zenodo.org/communities/ds-edu-dk

This community collection provides access to results from the National Coordination of Data Stewardship Education in Denmark[64]. The work started in March 2019 and is sponsored by the National Forum for Data Management. The initiative has analysed and summarised the state of data stewardship training in Denmark (DeiC) and aims to implement the recommendations in practice in form of a nationally-coordinated data stewardship programme at master-level.

- Wildgaard, Lorna, Vlachos, Evgenios, Nondal, Lars, Larsen, Asger Væring, & Svendsen, Michael. (2020, January 31). *National Coordination of Data Steward Education in Denmark: Final report to the National Forum for Research Data Management* (DM Forum) (Version 1). Zenodo. http://doi.org/10.5281/zenodo.3609516

- D.B. Deutz, M.C.H. Buss, J. S. Hansen, K. K. Hansen, K.G. Kjelmann, A.V. Larsen, E. Vlachos, K.F. Holmstrand (2020). *How to FAIR: a Danish website to guide researchers on making research data more FAIR.* https://doi.org/10.5281/zenodo.3712065

Finland

Open Science learning course

Course: https://findocnet.fi/course/view.php?id=136

This course is designed as an introduction to Open Science and Research principles and practices for doctoral students and other research active members new to Open Science and Research. The

[64] www.deic.dk/datamanagement/aktiviteter/tvaerinstitionelle

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

resource introduces principles, concepts, and new ways of doing research with the view that at every step the learning can be related to the research the learner is working on.

This online course, targeted for PhD students, is focused and structured by four main sections, Open Science, Open Access, Open Data and Open Research Process. Course has a clear structure, being defined modules for each section, as well the learning outcomes.

Germany

DINI/nestor research data working group

Since 2014, the joint working group on research data of DINI e.V.[65] and nestor[66] (DINI/nestor-AG Forschungsdaten) has served as a platform for the exchange of knowledge and experiences among professionals and experts in the German-speaking countries and has played an active role in supporting, advancing and coordinating efforts related to research data and research data management.[67] It is regularly organising workshops on relevant topics targeted at persons working in data-related roles[68] and has a number of sub-working groups dedicated to specific topics, including one on training.[69]

The aim of this training sub-working group is to develop training models and -materials as well as to collect existing training materials, with a strong focus on data stewards, data managers and related roles. The ultimate goal is to design and establish a certified RDM training course for this target group, including a sustainable business model. As a first step in this endeavour, the group is developing a modular training concept based on existing materials, e.g. the concept for a train-the-trainer course (see next paragraph for more detail) developed by the FDMentor project (2017–2019).[70] Working group outputs so far include a collection of training resources[71] and a metadata schema for research data management training materials.[72]

---

[65] "Deutsche Initiative für Netzwerkinformation", an association aiming to improve and develop information infrastructures in the higher education sector, https://dini.de/dini/ueber-uns/

[66] "network of expertise in long-term storage of digital resources in Germany", www.langzeitarchivierung.de/Webs/nestor/EN/nestor/Ueber_uns/ueber_uns_node.html

[67] www.langzeitarchivierung.de/Webs/nestor/EN/Arbeitsgruppen/AG_Forschungsdaten/ag_forschungsdaten_node.html;jsessionid=30840D0988966609922946E073815971.internet571, https://dini.de/ag/dininestor-ag-forschungsdaten/

[68] www.forschungsdaten.org/index.php/AG_Forschungsdaten

[69] www.forschungsdaten.org/index.php/UAG_Schulungen/Fortbildungen

[70] www.forschungsdaten.org/index.php/FDMentor

[71] https://rs.cms.hu-berlin.de/uag_fdm/pages/home.php

[72] Biernacka, Katarzyna, Peters, Karsten, Danker, Sarah Ann, Engelhardt, Claudia, Helbig, Kerstin, Hendriks, Sonja, … Ziedorn, Frauke. (2020, May 4). Metadata Schema for Research Data Management Training Materials. Zenodo. http://doi.org/10.5281/zenodo.3784238,
Poster (in English): https://doi.org/10.5281/zenodo.3466377

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

FDMentor train-the-trainer course on research data management

The FDMentor train-the-trainer concept[73] is a detailed plan for a two-day train the trainer course on research data management. It covers not only different areas related to RDM, but also didactics and introduces a variety of exercises and methods which enable the participants to design and deliver the content in an engaging and entertaining way.

The following RDM-related topics are covered:

- Digital research data, research data lifecycle, RDM (incl. The FAIR principles)
- Research data policies
- Data management plans
- Order and structure
- Documentation and metadata
- Storage and backup
- Long-term preservation
- Data security and access control
- Data publication
- Reuse of research data
- Legal aspects
- Institutional infrastructure

To date, the course has been delivered 12 times (with 10-12 participants each), 5 more trainings are planned in 2020 (currently, the material is only available in German).

Activities on (federal) state level

There are RDM initiatives in place in several of the German federal states, varying greatly in terms of funding (volume and source), organisational structure and activities.[74] Many of them tackle the issue of RDM training in one form or another – although in some cases not much information is available online about these activities. Noteworthy examples are the FOKUS project from Hessen, *eHumanities – interdisziplinär* from Bavaria and the platform *forschungsdaten.info* which originated from the RDM initiative of Baden-Württemberg.

In the FOKUS[75] project (2017–2019), five Hessian universities collaborated to develop (mainly) specialised RDM training units for a number of disciplines - most of them in the form of modules that can be integrated into existing discipline-specific courses, some designed as stand-alone workshops.

[73] Dolzycka, Dominika, Biernacka, Katarzyna, Helbig, Kerstin, & Buchholz, Petra. (2019). *Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement* (Version 2.0). Zenodo. http://doi.org/10.5281/zenodo.2581292

[74] For an overview, see www.forschungsdaten.info/fdm-im-deutschsprachigen-raum/

[75] www.uni-marburg.de/de/forschung/kontakt/eresearch/projekte-und-netzwerke/fokus

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

The material was tested and evaluated in 16 university courses in 2018 and 2019 and is available online.[76] (For an analysis of some of these courses, see chapter 4).

The three-year project *eHumanities – interdisziplinär,*[77] running since January 2018, is aiming to advance research data management practices in the Digital Humanities and Social Sciences. Part of its activities is to develop modular training materials on RDM topics. So far, a number of video tutorials[78] and three online courses have been developed. The courses are in German, the titles and topics below have been translated into English for this report:

Table 11. Online courses developed by eHumanities – interdisziplinär

Research software and data formats in the humanities and social sciences[79]	**Data management plans & RDMO**[80]	**Search for and reuse research data**[81]
• Introduction (Overview of research software and data formats, uses of research software, how to choose research software) • R • IIIF & Mirador • Git & GitHub • Regular Expressions • Graph databases and SPARQL	• Introduction to DMPs • RDMO for users • RDMO for administrators	• Finding an appropriate data source • Searching the data source • Evaluate search results • Reuse data properly (citation, licences)

[76] Becker, Henrike, Dorn, Christian, Einwächter, Sophie, Klein, Benedikt, Krähwinkel, Esther, Mehl, Sebastian, … Werthmüller, Julia. (2019). Forschungsdatenkurse für Studierende und Graduierte: Lehr- und Schulungsmaterialien zur Nachnutzung (Version v1.0). Zenodo. https://doi.org/10.5281/zenodo.3381974

[77] www.fdm-bayern.org/ehumanities-interdisziplinaer/

[78] Available through the Youtube channel of Forschungsdatenmanagement Bayern: https://www.youtube.com/channel/UC5CCijOlCLxMKO4PIjE5-fg

[79] Forschungssoftware und Datenformate in den Geistes- und Sozialwissenschaften. www.studon.fau.de/lm2993840.html

[80] Datenmanagementpläne und RDMO. https://www.studon.fau.de/lm2993053.html

[81] Forschungsdaten suchen & nachnutzen. https://www.studon.fau.de/lm2994018.html

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Netherlands

Research Data Netherlands

Research Data Netherlands (RDNL, a collaboration of 4TU, DANS and SURFsara).

- Essentials 4 Data Support https://datasupport.researchdata.nl/en/. This course has educated more than 300 data supporters / data stewards in the Netherlands on FAIR data, Open Science and RDM. The course has been running over 10 years. It consists of two face-to-face course days, exercises and online content. The online content was updated in 2019[82].

National Coordination Point Research Data Management (LCRDM): www.lcrdm.nl/en

LCRDM is the National Coordination Point Research Data Management in the Netherlands, a network of experts in the field of RDM, connecting policy and practice. It maintains a Zenodo community to provide access to all outputs and products created by their Working Groups (2015-2017) and Task Groups (2018 - onwards) at: https://zenodo.org/communities/lcrdm-output-and-products/.

National Platform Open Science: www.openscience.nl/en

The focus for the Platform is to create acceleration regarding the three key areas[83] of the National Plan Open Science (open access, (re)use of data and evaluation systems), as described at www.openscience.nl/en/national-platform-open-science.

Helis Academy

The Helis Academy (https://helisacademy.com/en) strives to optimise the Life Sciences and Health sector in Flanders and the Netherlands. The Helis Academy focuses on the development and coordination of appropriate training, internships and traineeships in the Life Sciences and Health sector.

One of the domains targeted by Helis Academy is "Data analysis & stewardship" (https://helisacademy.com/en/data-analysis-stewardship), focusing on four specific topics: FAIR data stewardship, omics data analysis, statistics and machine learning. Regarding FAIR data stewardship, HELIS is offering the following 3 day course:

[82] The competences are described: https://datasupport.researchdata.nl/en/about-the-course/competencies.
[83] www.openscience.nl/en/national-platform-open-science/the-key-areas

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Table 12. Modules and objectives of the Helis FAIR data stewardship course

Helis FAIR data stewardship course: Digital data scholarship for the life sciences	
Modules	Module 1: Why use a data management plan and what for? Module 2: Mastering the data chaos during the research phase. Module 3: Boost your visibility and get cited for your data!
Objectives	Gaining knowledge on: ● Research data management Life Cycle ● FAIR principles ● FAIR data stewardship Practical experience in: ● Data management planning ● Cleaning data (Step 1 of making data FAIR) ● Semantic interoperability between data ● Archiving and publishing of data ● Persistent identifiers and their use cases

(Source: Helis FAIR data stewardship course. www.aanmelder.nl/data-stewardship-course)

3.10. Analysis of education and training programmes

The education and training programmes summarised in Chapter 3 are important initiatives advancing skills in a variety of scientific domains or professions. They contribute to skills of researchers, data stewards, infrastructure personnel and other groups.

Mirroring the approach in the previous chapter, the different education and training programmes were compared and analysed the different programmes using a set of dimensions (Annex 2).

---

*All results are displayed in **Annex 2***

Table: Comparison of education and training programmes.

***Use of skills or competence frameworks:** Does the education or training programme refer to any underlying skills or competence framework?*

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

The majority of initiatives do not specify if a framework has been used or consulted for the development of training or education activities. An exception is the ELIXIR training programme, which uses insights from the "Towards FAIR data stewardship in the life sciences report". terms4FAIRskills, although not a training programme in the strictest sense, is based on the FAIR4S framework.

Professional group covered: *Is the education or training programme targeting a specific group, such as researchers, data stewards/managers, research support staff, other non-academic professionals etc.?*

The different programmes in the sample a wide range of scientific communities, in particular the programmes based in specific research infrastructures (ELIXIR, CESSDA, SSHOC, ENVRI, ExPaNDS, DARIAH, PARTHENOS, PaNOSC). RDA, OpenAIRE, GO-FAIR, FOSTER, CODATA, terms4FAIRskills, as well as LEARN represent more broad and/or generic initiatives seeking to develop training and skills for RDM - although RDA groups in specific domains can be understood as contribution to skills and competence development in their respective communities.

In relation to the professional groups and roles addressed, the programmes in the sample show a tendency towards professionals in RDM-related roles (e.g. data stewards, project managers, librarians, infrastructure staff) or researchers starting with early-career researchers/doctoral candidates. Bachelor or master-level education does not appear as priority across the sample.

Learning outcomes or other elements supporting course development: *Is there information available about learning outcomes or other elements that could support others in building courses based on the documents?*

With the exception of ELIXIR, the research infrastructure training programmes do not specify learning outcomes in the documentation available to the authors. Other training programmes provide main topics covered on the curriculum (CODATA) or specify that they intend to develop curricula in the future (GO-FAIR, terms4FAIRskills). RDA and OpenAIRE, being large broad initiatives, did not further specify learning outcomes. The FOSTER courses, based on the FOSTER taxonomy, defined learning outcomes.

Topics covered: *Does the education or training programme explicitly address topics such as RDM, FAIR data, or Open Science?*

Given the sampling method, almost all training programmes addressed RDM and/or FAIR data. Some included broader topics in Open Science (OpenAIRE, FOSTER, RDA, SSHOC, DARIAH, PARTHENOS) or data science (CODATA).

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

4. Institutional practices

The main objective of FAIRsFAIR WP7 is to advance the integration of FAIR RDM skills in higher education curricula. To complement the analysis of existing skills and competence frameworks in Chapter 2 with use cases and examples from practice, we collected via convenience sampling data about 27 courses and modules offered by universities and other higher education institutions that address RDM and FAIR within formal higher education and carried out a brief quantitative analysis.[84] The purpose of this exercise is to enrich the analysis done in Chapter 4 with data on programme-level and institutional approaches to data-related skills and the support identification of possible gaps.

The sample contains 27 different learning units, courses or full programmes addressing RDM and, to some extent, FAIR skills (Annex 3). Based on the discussion in Chapter 1.2 and in particular the results of the Universities UK report *Making the most of data: Data skills training in English universities* (2015), which pointed towards challenges stemming from the isolation of data-related skills training from the rest of the curriculum and the need to relate data training to disciplinary practice, these 27 cases were analysed in terms of connection to domain or discipline, their qualification level, and other factors. Notably, 13 cases were taken from a regional project in the German state of Hessen, which sought to advance RDM training within universities and produced reusable learning resources available on Zenodo ("FOKUS" project).[85]

---

*Results are displayed in **Annex 3***

Table: Institutional practices

The sample consists of a wide range of different cases. Some are individual learning units which are embedded in a specific course -– in several instances on research methods. Others are fully in-person or blended training courses delivered over a duration of several weeks. Three cases are Massive Open Online Courses (MOOCs) developed by universities. Another three cases are self-standing programmes at postgraduate level which award university certificates or Master degrees for RDM or library and information professionals upon completion.

One discernible difference can be found between the collection of courses from the aforementioned German project and the rest. 12 of the 13 FOKUS learning units were part of disciplinary or domain-

[84] This was done with a convenience sample of courses and modules known to the project partners as well as information gathered online and through FAIRsFAIR D7.1.

[85] Becker, Henrike, Dorn, Christian, Einwächter, Sophie, Klein, Benedikt, Krähwinkel, Esther, Mehl, Sebastian, … Werthmüller, Julia. (2019). *Forschungsdatenkurse für Studierende und Graduierte: Lehr- und Schulungsmaterialien zur Nachnutzung* (Version v1.0). Zenodo. https://doi.org/10.5281/zenodo.3381974

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

specific courses – either in research training or introductory courses. 6 out of 13 were also targeted at Bachelor students; one solely at Master students; the remainder at doctoral candidates or early-career researchers (ECR). Of the other 14 cases, only one was devised for the Bachelor level; 2 were at least 'open' for Bachelor students; 3 were at Master level, and the remainder for doctoral candidates or ECRs.

Moreover, 12 out of the 13 FOKUS cases specifically had a disciplinary focus. This is contrasted with only one of the other 14 cases (including the 3 cases which were self-standing programmes at postgraduate level), which consisted of a generic part and a more topical part on biodiversity data. In contrast, at least 7 FOKUS cases were embedded in a Bachelor, Master or doctoral training programme; whereas only one of the other cases was remarked to be 'mandatory' in the two programmes that it was originally developed for.

On the other hand, with the exception of one case it could not be established to what extent the FOKUS cases included a specific assessment beyond practical exercises during each course – likely resulting from their nature as learning units that would be part of a full course or module. This compares with 9 of the remaining 14 cases with a specific assessment exercise (e.g. a written assignment). The 3 identified MOOCs did not mention assessment, in addition to 2 cases with unclear information.

Finally, we investigated if the publicly available documentation of each case included specific reference to the FAIR principles, e.g. as part of the course description or session plans. In total, this could be verified in 12 cases - out of which only two draw from the FOKUS cases. However, this does not exclude that FAIR is not fully addressed; many cases also included reference to different stages of the research data lifecycle.[86]

This preliminary and very heterogeneous data allows only limited observations:

- At Bachelor and Master level, FAIR appears to be scarcely taught. Yet, there are practical cases, examples and experiences how to teach RDM which could be expanded to make reference to FAIR.
- RDM is, often, being taught in standalone courses or learning units embedded into existing methods courses.
- RDM is being taught in very different formats and contexts.
- Outside of specific projects such as FOKUS which target Bachelor and Master-level education, RDM training broadly appears to be targeted at doctoral candidates and ECRs, in a generic fashion.

[86] This may be a result of the FOKUS sample stemming largely from different academic years between 2017 and 2019, at the time FAIR may not have been as widespread as at the time of writing this report.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

5. Summary

This deliverable presents an overview of different skills and competence frameworks, training and skills programmes and a selection of use cases at institutional and research infrastructure/national level.

Main findings

With regard to **skills and competence frameworks**, the report shows that a variety of resources is available for data science and more data stewardship. Domain-specific resources on data management skills seem scarce however, with the exception of ELIXIR and, to some extent, Belmont. Moreover, these frameworks are not specifically designed for use at Bachelor or Master level. Data science frameworks such as EDSF or GFI provide links to domain-specific practices but fall short of providing detailed information. With the objective of supporting the integration of RDM skills training in higher education curricula in mind, there appears to be a gap of resources that provide information for programme designers about discipline specific practices and supporting material in form of frameworks and model courses or curricula. Moreover, the identification of relevant use cases at Bachelor or Master level for the frameworks prove difficult, which is a challenge for an approach that would highlight good practices.

With regard to **training programmes**, the analysed cases display a 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. The analysed training programmes ranged from more generic data stewardship or data science initiatives (GO-FAIR, CODATA, RDA) to research infrastructures servicing different domains and fields (e.g. ELIXIR, CESSDA, SSHOC, ENVRI, ExPaNDS, DARIAH, PARTHENOS, PaNOSC).

The analysis of **institutional practices** of RDM training at Bachelor, Master or doctoral level confirms that the practical use and impact of the frameworks is unclear. They also highlight the difficulty in collecting detailed information and identifying cases in which RDM is embedded at BA/Master level. The bulk of 'positive' cases where RDM was addressed at Bachelor or Master level stems from a targeted project advancing exactly these practices and sharing results and learning material. Outside of these cases, RDM training appears targeted at doctoral or postdoctoral level. A speculative reason for this dominant approach could be that academic research data management is mostly required at these stages - therefore higher education institutions focus the use of resources at training early-career researchers with necessary skills. It would also, implicitly, resonate with the observation that RDM is not yet common disciplinary practice. Focussing efforts on ECRs could, in the long run, lead to a "trickle-down" effect which sees RDM slowly emerge as a standard element of, for instance, undergraduate research methods training.

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Outlook

Overall, the information compiled in this report highlights that existing frameworks that address bachelor or master-level education mainly stem from data science initiatives, with frameworks targeting master-level education for data stewards in development. Outside of targeted projects, RDM education and training with a domain focus appeared mostly at a later stage, the target groups are (early-career) researchers or professionals with roles related to data stewardship. However, the wide range of existing training in more specific communities given through European-level research infrastructures shows that disciplinary RDM and/or FAIR data cultures exist or are emerging. However, this has apparently not yet translated, on the one hand, into a more formal definition of disciplinary RDM/FAIR data skills and competences in skills and competence frameworks or, on the other hand, a wider integration of RDM skills in bachelor or master-level curricula.

In other words, this implies that more community-driven action on RDM skills is needed to anchor RDM and FAIR skills in higher education curricula, organised e.g. at European or national level through European research infrastructures or societies which are active in advancing learning and teaching in a given discipline. This is a task beyond the immediate reach of FAIRsFAIR or most EOSC-related actions. Moreover, targeted projects or programmes supporting the integration of RDM or FAIR skills in bachelor or master programmes are helpful in supplying resources for their practical implementation and dissemination. Such activities could be informed by the plethora of resources and activities identified in this briefing - and others not represented here. There is a high potential for collaboration and mutual learning across institutional, disciplinary or national boundaries.

With regard to FAIRsFAIR, this briefing is a step advancing the objective to develop a competence framework for FAIR data by February 2021, to continue targeted stakeholder engagement throughout the duration of the project, and to develop training resources for universities that advise on RDM/FAIR skill integration in higher education curricula.

In the context of EOSC, this deliverable also aims to supply the Skills & Training Working Group with background information to reach their objective to “Take stock of existing work [...] and identify and prioritize open science and digital skill sets for EOSC, targeting researchers (for open science and RDM practices), service providers (for providing their services to EOSC), and policy makers (on how to align with open science and data policies)”[87] and the proposed objective of the EOSC European Partnership that seeks to advance the teaching of Open Science and data stewardship skills.

[87] www.eoscsecretariat.eu/working-groups/skills-training-working-group

FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Bibliography

Ashley, Kevin (2016) *Developing Skills for Managing Research Data and Software In Open Research*. Wellcome Trust. https://dx.doi.org/10.6084/m9.figshare.4133916

Becker, Henrike, Dorn, Christian, Einwächter, Sophie, Klein, Benedikt, Krähwinkel, Esther, Mehl, Sebastian, … Werthmüller, Julia. (2019). *Forschungsdatenkurse für Studierende und Graduierte: Lehr- und Schulungsmaterialien zur Nachnutzung* (Version v1.0). Zenodo. https://doi.org/10.5281/zenodo.3381974

Biernacka, Katarzyna, Peters, Karsten, Danker, Sarah Ann, Engelhardt, Claudia, Helbig, Kerstin, Hendriks, Sonja, … Ziedorn, Frauke. (2020, May 4). Metadata Schema for Research Data Management Training Materials. Zenodo. http://doi.org/10.5281/zenodo.3784238,

Carretero, S.; Vuorikari, R. and Punie, Y. (2017). *DigComp 2.1: The Digital Competence Framework for Citizens with eight profi-ciency levels and examples of use*, EUR 28558 EN, doi:10.2760/38842

Chapman, P. et al. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. ftp://ftp.software.ibm.com/software/analytics/spss/support/Modeler/Documentation/14/UserManual/CRISP-DM.pdf;

Data Assessment Solutions GmbH. (2014). *IT-Skills-Studie 2014: Big Data Projekte*. White-Paper, 2014.

Dolzycka, Dominika, Biernacka, Katarzyna, Helbig, Kerstin, & Buchholz, Petra. (2019). Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement (Version 2.0). Zenodo. http://doi.org/10.5281/zenodo.2581292

Einwächter, Sophie G., Esther Krähwinkel, and Frederik Ostsieker. (2020). Lessons Learned: Thesen zur FDM-Kompetenzausbildung: Erkenntnisse aus dem Vernetzungstreffen der vom BMBF geförderten Projekte EeFDM Jena, FDMentor, FOKUS, PODMAN Und UniLLAB am 30. und 31. Januar 2019 in Marburg“. *Bausteine Forschungsdatenmanagement*, Nr. 1 (April), 8-15. https://doi.org/10.17192/bfdm.2020.1.8101

Gesellschaft für Informatik (2019). *Data Science: Lern- und Ausbildungsinhalte*. www.plattform-lernende-systeme.de/files/Downloads/Publikationen/GI_Arbeitspapier_Data-Science_2019-12_01.pdf

Ingrid Dillo, Marjan Grootveld, Simon Hodson, & Sara Pittonet Gaiarin. (2020). *Second Report of the FAIRsFAIR Synchronisation Force (D5.5)* (Version 1.0). Zenodo. https://doi.org/10.5281/zenodo.3953979

FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

McCaffrey, Ciara, Meyer, Thorsten, Riera Quintero, Clara, Swiatek, Cecile, Marcerou-Ramel, Nathalie, Gillén, Camilla, … Egerton, Frank. (2020, March 10). *Open Science Skills Visualisation*. Zenodo. http://doi.org/10.5281/zenodo.3702401

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

Open Data Institute (2020). Data Skills Framework. http://theodi.org/wp-content/uploads/2020/05/2020-05-Data-Skills-Framework.pdf

Pontika, Nancy; Knoth, Petr; Cancellieri, Matteo and Pearce, Samuel (2015). Fostering Open Science to Research using a Taxonomy and an eLearning Portal. In: iKnow: 15th International Conference on Knowledge Technologies and Data Driven Business, 21-22 Oct 2015, Graz, Austria. https://doi.org/10.1145/2809563.2809571

Ridsdale, C. et al. (2015). *Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report*. http://hdl.handle.net/10222/64578;

Scholtens, S., Jetten, M., Böhmer, J., Staiger, C., Slouwerhof, I., Geest, M. van der, Gelder, C.W.G van (2019, October 5). Towards FAIR data steward as profession for the lifesciences. Report of a ZonMw funded collaborative approach built on existing expertise. Zenodo. http://doi.org/10.5281/zenodo.3471708

Stoy, Lennart, Saenen, Bregt, Davidson, Joy, Engelhardt, Claudia, & Gaillard, Vinciane. (2020). *D7.1 FAIR in European Higher Education* (Version v1.0_draft). Zenodo. https://doi.org/10.5281/zenodo.3629683

Universities UK (2015). *Making the most of data: Data skills training in English universities*, p. 22. www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2015/making-the-most-of-data-training-skills-english-universities.pdf

Vuorikari, R., Punie, Y., Carretero Gomez S., Van den Brande, G. (2016). *DigComp 2.0: The Digital Competence Framework for Citizens. Update Phase 1: The Conceptual Reference Model*. Luxembourg Publication Office of the European Union. EUR 27948 EN. doi:10.2791/11517

Whyte, A., Ashley K. (2017). *D7.1: Skills landscape analysis and competence model*. https://eoscpilot.eu/sites/default/files/eoscpilot-d7.1.pdf.

Wilkinson, M., Dumontier, M., Aalbersberg, I., *et al.* (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018, pp. 1-9. https://doi.org/10.1038/sdata.2016.18

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Annex 1 – Table: Comparison of skills frameworks

Framework/ Resource	Qualification level covered	Professional group covered	Main field / domain	Main elements	Defines Learning Outcomes or other elements that support development of courses	Addresses FAIR data specifically	Aspects from research data life cycle (do they address different RDM aspects)	Addresses Open Science	Examples for practical implementation / usage of the resource
FAIR4S	Professional and PhD level	Data manager, data curator, data librarian, data service engineer, data science engineer, research service developer, research software engineer.	FAIR data stewardship	Group of skills / competences for each research lifecycle stage; Identification of professional groups sharing stewardship responsibility.	No	Yes	Yes. Skills are organised around the research life cycle.	Yes, but generically.	Basis for terms4FAIRskills initiative. Informed other frameworks such as ELIXIR and DEIC.
FOSTER	Professional and PhD level	Doctoral Students, Researchers, Research Project Managers, Knowledge Managers, Librarians, Funding Agencies	Open Science	Open Science Taxonomy; Open Science training toolkit; Open Science Training Handbook oriented to practical teaching; Open Science Learning objectives.	Yes	Yes	Yes	Yes. Courses answer some of the most common questions about putting Open Science into practice.	Referenced by several learning courses
Towards FAIR data steward in the life sciences	Professional and PhD level	Data steward policy (Policy makers, Funders, University boards, University deans), Data steward research (Researchers, Scientists, Data scientists), Data steward infrastructure (Application managers, Technicians, Infrastructure providers, IT personnel)	Data steward for Life sciences	Data Steward roles, function descriptions (responsibilities and tasks), competencies (knowledge, skills and abilities) and learning objectives.	Yes	Yes	Yes	Yes. Knowledge about the FAIR data principles and the principles of Open Science, including knowledge about implementation and how to be compliant to the FAIR principles.	n/a

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

DEIC	3 models: - for students with Bachelor degree (1 year master programme; 2 year master programme; 2 year candidatus); - for students with PhD - for continuing and professional education	- students with a bachelor's degree in any field and targeted to the employability in the corporate sector; - students with a PhD or equivalent experience in any field, aspiring to work as a DS in a HEI; - mature students, who have professional experience and wish to improve their DS skills.	Data Stewardship	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.	No learning outcomes. Defines the roles of a data steward and the modules for each DS education model.	Yes	Yes, but generic.	Addresses some OS skills but not specifically defined/described.	n/a
EDSF	Bachelor, Master	- Managers (for managerial roles); - Professionals (for applications developers and for infrastructure engineers); - Technicians and Associate professionals (for operators and technicians); - Clerical support workers (for data curators and stewards)	Data Science	- 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	Yes	Yes, but not as a main competence.	Yes. Includes for example the development and implementation of a data management strategy for data collection, storage, preservation, and availability for further processing.	Yes, but generically.	Used in the development of FAIR4S, ELIXIR and DEIC
Data Science: Learning and training content (GFI)	Master level and professional learners	Students in data science or a specific domain with data science components; professionals	Data science; data science basic for other disciplines/domains	Curriculum recommendations; entry requirements for curricula; comparison of other skills frameworks	No learning outcomes; but defines modules and specific topics at different learning levels as part of model curricula	No	Mentions separate aspects of research data lifecycle, such as data ethics and privacy, data collection, data governance, data integration, as well as domain-specific applications, which include RDM (without further specification).	No	n/a

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

ODI Data Skills Framework	Other: specifies data-related skills in a company / organisation setting	Professionals & organisations	Data science in a business / organisation context	Includes examples for training courses	No learning outcomes;	No	Includes several skills related to RDM (e.g. standardising, cleaning and linking data, governing access) but no comprehensive view of RDM.	No	n/a
Belmont Forum: Data skills curricula framework	Professional and PhD level	PhD students, mid-career Researchers, Principal Investigators	Data skills in Global Environmental Change Research	Curricula goals, Curricula structure (core and optional modules, and for Principal Investigators)	No learning outcomes. Defines the curricula modules and specific related topics, as well the time length of each module.	No	Yes. 1 Core module about data management (DMPs, Reproducibility and organising data, Metadata, Citation and publishing, ethics, data transparency, PIDs, Data sharing, preservation, licensing and trusted repositories, …). For Principal Investigators, a module on DMPs and data repositories.	No	n/a

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Annex 2 – Table: Comparison of education and training programmes

Initiative	Mention of any framework / reference framework	Professional group covered	Defines Learning Outcomes or other elements that support development of courses	Topics coverage
RDA	No	All stakeholders involved in RDM activities	No	Addresses all topics related with RDM, FAIR data and Open Science.
OpenAIRE	No	Researchers, data managers, repository managers, research support staff, funders, policy makers	No	Addresses all topics related with RDM, FAIR data and Open Science.
GO-FAIR	No	Individuals, institutions and organisations dealing with FAIR data principles.	Not yet. 3 Implementation Networks in the training area are in preparatory phase (Season Schools, Training Curriculum, Training Frameworks). The DSCC-IN has as one of its objectives, to participate in the development of e-learning tools, courses and curricula for professional data stewards.	FAIR data, Open Science
FOSTER	No	Doctoral Students, Researchers, Research Project Managers, Knowledge Managers, Librarians, Funding Agencies	Yes	Addresses all topics related with RDM, FAIR data and Open Science.
CODATA	No	Early Career Researchers	No. Presents the main areas covered on the curriculum.	FAIR data, RDM, Data Science
terms4FAIRskills	No	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.	No. When finished will assist with the creation and assessment of stewardship curricula.	FAIR data
ELIXIR	Yes. Towards FAIR data steward in the life sciences	Life Sciences domain. Researchers, Data Stewards, Reviewers and funders, Trainers.	Yes	RDM, FAIR data
CESSDA	No	Data producers, data users, data professionals, CESSDA Service Provider staff	No	RDM, FAIR data
SSHOC	No. Mention that will develop new support materials based on existing resources	Social Sciences and Humanities. Data producers, data users, data professionals.	No	RDM, FAIR data, Open Science, Data science
ENVRI	No	Environmental and Earth sciences. ENVRI Community.	No	RDM, FAIR data
ExPaNDS	No	Photon and Neutron Research Infrastructures.	No	RDM, FAIR data

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

DARIAH	No	Digital Humanities. Researchers.	No	RDM, FAIR data, Open Science
PARTHENOS	No	Digital humanities and digital heritage. Researchers, educators, managers, policy makers	No	RDM, FAIR Principles and Open Science
PaNOSC	Will propose a new framework for Data Policies taking into account the FAIR principles	Neutron and Photon science.	No	RDM, FAIR data

FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Annex 3 – Table: Institutional practices

University	Country	Title	Level	Domain	Part of specific programme	Duration	ECTS / Other credits applied	Addresses RDM	Addresses FAIR	Learning outcomes	Delivery	Assessment	Other aspects	URL / Source
University College Cork	Ireland	Graduate Information Literacy Skills	PhD & Research Masters	Cross-domain / interdisciplinary	No	20 hours	Yes	Yes	Not specified	Yes	In person	Yes	RDM embedded in wider research method context.	https://libguides.ucc.ie/PG6009
Tampere University	Finland	Research Process: Managing Research Information	PhD, ECR	Cross-domain / interdisciplinary, with specialisation in 4 areas	No	4 weeks	Yes (2 ECTS)	Yes	Not specified	Yes	Online	Yes	RDM embedded in wider research method context.	https://www.tuni.fi/studentsguide/curriculum/course-units/otm-3f4fa64a-cb59-4b73-9eb9-c14a984d7600?year=2019
University of Twente	Netherlands	Research Data Bootcamp	PhD	Cross-domain / interdisciplinary	No	26 hours	Yes (1 ECTS)	Yes	Not specified	yes	Blended	Unspecified	Focus on DMPs	https://www.utwente.nl/en/courses/1000227/data-management-bootcamp/#content
Bielefeld University	Germany	Research Data Management	Open, but majority master-level	Cross-domain / interdisciplinary	No, but compulsory in 2 programmes	22.5 hours	Yes	Yes	Yes	Yes, competence based	Blended	Yes	Hands-on courses	https://doi.org/10.5334/dsj-2019-038
University of Turku	Finland	The Basics of the Research Data Management	PhD, ECR	Cross-domain / interdisciplinary	Not specified	not specified	Yes (3 ECTS)	Yes	Yes	Yes	In person	Yes		https://zenodo.org/record/3889895#.XuyuLudS9ZW, https://zenodo.org/record/3692225#.XuyuJedS9ZV
University of Tromso	Norway	Take Control of your PhD Journey: from (P)reflection to Publishing	PhD, ECR	Cross-domain / interdisciplinary	Not specified	3 days	Yes (2 ECTS)	Yes	Not specified	Yes	In person	Yes	RDM embedded in wider research method context.	https://uit.no/utdanning/emner/emne/541884/gen-8001
Leiden UAS	Netherlands	FAIR Data Stewardship (basis)	Bachelor	Cross-domain / interdisciplinary	No	6 days/24 hours	Not specified	Yes	Yes	Unspecified	In person	Yes		https://www.hsleiden.nl/nascholingen/bioscience-en-diagnostiek/fair-data-stewardship-nl/index.htm, https://www.go-fair.org/2020/02/16/elective-course-introduction-to-fair-data-stewardship/
University of Tartu	Estonia	Research Data Management and Publishing	PhD	Cross-domain / interdisciplinary, 2nd part on biodiversity data	Not specified	Not specified	No	Yes	Yes	Unspecified	Online, self-paced	No		https://www.openaire.eu/blogs/research-data-management-e-course-in-collaboration-1-2, https://sisu.ut.ee/andmeh

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

														aldus/open-science?lang=en
Danish National Forum for Data Management	Denmark	Research Data Management	PhD, researchers, research support, but open	Cross-domain / interdisciplinary	No	Not specified	No	Yes	Yes	Yes	Online, self-paced	No		https://vidensportal.deic.dk/en/RDMELearn
University of Cambridge	United Kingdom	Research Data Management	Open	Cross-domain / interdisciplinary	Not specified	Not specified	No	Yes	Not specified	Unspecified	Online, self-paced	No		https://libguides.cam.ac.uk/research-skills/rdm
Université de Montpellier	France	Gestion des données de la science – Scientific Data Management	University Diploma	Cross-domain / interdisciplinary	Self-standing programme	186 hours	Not specified	Yes	Yes	Yes	In person	Yes		https://sdm.edu.umontpellier.fr/
Potsdam UAS, HU Berlin	Germany	Digitales Datenmanagement	Master	Cross-domain / interdisciplinary	Self-standing programme	2 yrs, 2000 hours	Yes (120 ECTS)	Yes	Yes	Yes	In person	Yes		https://www.ddm-master.de/
University of Vienna	Austria	Certificate course “Data Librarian"	Master	Cross-domain / interdisciplinary	Self-standing programme	250 hours	Yes (10 10 ECTS)	Yes	Yes	Yes	Not specified	Yes	Self-standing certificate study programme for data librarians	https://www.postgraduatecenter.at/en/programs/communication-media/data-librarian/
TU Delft	Netherlands	Open Science: Sharing your research with the world	Bachelor and higher	Cross-domain / interdisciplinary	Self-standing programme	16 hours	No	Yes	Yes	Yes	Online, self-paced	Unspecified	MOOC on open science in general with module on RDM	https://ocw.tudelft.nl/courses/open-science-sharing-research-world/?view=course-materials
TU Darmstadt	Germany	Einführung in das Forschungsdatenmanagement	Bachelor, Master	Chemistry	Not specified	1.5 hours	Not specified	Yes	Not specified	Yes	In person	Unspecified	One learning unit in a full course	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
Giessen University	Germany	Einführung in das Forschungsdatenmanagement	Bachelor	Food science	Yes, embedded in specific module	1 hour	Yes (as part of full course)	Yes	Not specified	Yes	In person	Unspecified	One learning unit in a full course	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
Marburg University	Germany	Einführung in das Forschungsdatenmanagement	Master	Education science	Yes, embedded in specific module	1 hour	Yes (as part of full course)	Yes	Not specified	Yes	In person	Unspecified	One learning unit in a full course on research methods	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
Frankfurt University	Germany	Workshop zu Umgang mit Forschungsdaten in der Film- und Medienwissenschaft	Bachelor, Master	Film studies	Not specified	1.5 hours	Not specified	Yes	Not specified	Yes	In person	Unspecified	One learning unit in a full course	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
Frankfurt University	Germany	Research Data Management - Basics for Graduates	PhD	Film studies	Not specified	2 hours	Not specified	Yes	Not specified	Yes	In person	Unspecified	Part of research training workshops for graduate students	https://zenodo.org/record/3381975#.Xx7E0edS9KJ

FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Marburg University	Germany	Forschungsdatenmanagement in den Geisteswissenschaften	Bachelor	Humanities	Yes, embedded in specific module	1.5 hours	Yes (as part of full course)	Yes	Not specified	Yes	In person	Unspecified	One learning unit in a full course	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
Frankfurt University	Germany	ForschungsDatenManagement für Nachwuchswissenschaftler*Innen	PhD, ECR	Biology, geosciences	No	3.5 hours	Not specified	Yes	Yes	Yes	In person	Unspecified	Part of research training workshops for graduate students	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
Frankfurt University	Germany	Forschungsdatenmanagement für Promovierende und Postdocs der Biologie und der Geowissenschaften im Bereich der Biodiversitäts-, Klima- und Umweltforschung	PhD, ECR	Environmental sciences	No	7 hours	Not specified	Yes	Yes	Yes	In person	Unspecified	Part of research training workshops for graduate students	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
University of Applied Science Fulda	Germany	Forschungsdatenkurse für Studierende und Graduierte	PhD	Cross-domain / interdisciplinary	No	6 hours	Not specified	Yes	Not specified	Yes	In person	Unspecified	Part of research training workshops for graduate students	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
TU Darmstadt	Germany	Einführung in das Forschungsdatenmanagement	Bachelor	Computer science	Yes, embedded in specific module	1.5 hours	Yes (as part of full course)	Yes	Not specified	Yes	In person	Unspecified	One learning unit in a full course on research methods	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
Giessen University	Germany	Einführung in das Forschungsdatenmanagement in der Veterinärmedizin	PhD	Veterinary sciences	Yes, embedded in specific module	1.5 hours	Not specified	Yes	Not specified	Yes	In person	Unspecified	One learning unit in a full course on research methods	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
Giessen University	Germany	Einführung in das Forschungsdatenmanagement	Bachelor	Environmental sciences	Yes, embedded in specific module	3 hours	Yes (as part of full course)	Yes	Not specified	Yes	In person	Unspecified	One learning unit in a full course	https://zenodo.org/record/3381975#.Xx7E0edS9KJ
Marburg University	Germany	Forschungsdatenmanagement in den Wirtschaftswissenschaften	Bachelor, Master	Economics	Yes, embedded in specific module	6 hours	Yes (3 ECTS)	Yes	Not specified	Yes	In person	Yes	Workshop offered as part of research methods training	https://zenodo.org/record/3381975#.Xx7E0edS9KJ

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558