



D6.1

Inventory & gap analysis of FAIR training materials

Work Package	6
Lead partner	ULUND
Status	Final
Deliverable type	Report
Dissemination level	Public
Due date	2019-08-31
Submission date	2019-10-02

Deliverable abstract

This deliverable summarizes the gap analyses undertaken to 1) clarify the needs & requirements of the ENVRI-FAIR project partner RIs for training on both general FAIR-related topics as well as related research data management (RDM) issues; and 2) map out existing materials (courses, webinars, texts) that may be leveraged by WP6.

The outcomes of these gap analyses will be used to inform the prioritization and planning of ENVRI-FAIR training and skills-building activities, including the preparation of new course materials and the design & scheduling of training events targeting different staff categories.



DELIVERY SLIP

	Name	Partner Organization	Date
Main Author	Margareta Hellström	ULUND	
Contributing Authors	Maria Johnsson Jacco Konijn Nicola Fiore Erwann Quimbert Damien Boulanger Garry Baker	ULUND UVA LIFEWATCH ERIC IFREMER IAGOS UK-RI	
Reviewer(s)	Zhiming Zhao Markus Stocker	UVA TIB	
Approver	Andreas Petzold	FZJ	2019-10-02

DELIVERY LOG

Issue	Date	Comment	Author
V 0.9	2019-09-16	Sent for review	Margareta Hellström
V 1.0	2019-09-30	Sent for approval	Margareta Hellström
V 1.0	2019-10-02	Approved, submitted	Andreas Petzold

DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the Project Manager at manager@envri-fair.eu, as well as to the lead author (margareta.hellstrom@nateko.lu.se).

TERMINOLOGY

A relevant project terminology is included in Appendix A. The latest version of the master list of the terminology is available at <http://doi.org/10.5281/zenodo.3465753>.

PROJECT SUMMARY

ENVRI-FAIR is the connection of the ESFRI Cluster of Environmental Research Infrastructures (ENVRI) to the European Open Science Cloud (EOSC). Participating research infrastructures (RI) of the environmental domain cover the sub-domains Atmosphere, Marine, Solid Earth and Biodiversity / Ecosystems and thus the Earth system in its full complexity.

The overarching goal is that at the end of the proposed project, all participating RIs have built a set of FAIR data services which enhances the efficiency and productivity of researchers, supports innovation, enables data- and knowledge-based decisions and connects the ENVRI Cluster to the EOSC.

This goal is reached by: (1) well defined community policies and standards on all steps of the data life cycle, aligned with the wider European policies, as well as with international developments; (2) each participating RI will have sustainable, transparent and auditable data services, for each step of data life cycle, compliant to the FAIR principles. (3) the focus of the proposed work is put on the implementation of prototypes for testing pre-production services at each RI, the catalogue of prepared services is defined for each RI independently, depending on the maturity of the involved RIs; (4) the complete set of thematic data services and tools provided by the ENVRI cluster is exposed under the EOSC catalogue of services.

TABLE OF CONTENTS

1 Introduction and background	4
1.1 About Work Package 6 and this deliverable	4
1.2 The motivation for a gap analysis	4
1.3 Identifying relevant FAIR training topics	4
2 Inventory of existing training materials on FAIR and related RDM topics	6
2.1 Materials and methods.....	6
2.2 Summary per topic category.....	8
2.2.1 General FAIR-related training topics	8
2.2.2 Research Data Management (RDM) training topics	8
3 Survey of ENVRI-FAIR partners' knowledge level and training needs	8
3.1 Materials and methods.....	8
3.1.1 Input from WP7 activities	8
3.1.2 The WP6 questionnaire	9
3.2 Summary per topic category.....	10
3.2.1 General FAIR-related training topics	10
3.2.2 Research Data Management (RDM) training topics	11
4 Implications for Task 6.1 activities	12
4.1 Training materials & resources.....	12
4.2 Training events	12
4.3 Target audiences for training	12
4.4 Training requirements according to priorities	13
4.5 Domain-specific needs.....	15
5 Summary & conclusions	15
6 Impact on the project	17
7 Impact on stakeholders	17
Acknowledgements	17
Appendix A: Glossary and terminology	18
Appendix B: Survey of RI training needs & requirements	19
Appendix C: Supplementary materials	23
Summary across RIs: current knowledge levels & priorities	23
General FAIR-related topics: knowledge status & priorities	25
RDM-specific topics: knowledge status & priorities.....	26
Availability of training materials relevant to sub-domain RIs	28

D6.1: INVENTORY & GAP ANALYSIS OF FAIR TRAINING MATERIALS

1 Introduction and background

1.1 About Work Package 6 and this deliverable

The objective of ENVRI-FAIR Work Package 6 is to provide training to ENVRI and key ENVRI stakeholder groups about the FAIR principles, how to implement them in RI services and data management activities at data centre level, how to evaluate the degree of implementation using FAIR metrics, as well as relevant legal and policy requirements.

This deliverable, D6.1, is the first output of Task 6.1, "Building training material for FAIR implementation". As outlined in the Description of Actions, this task shall concern itself with the collection, production and dissemination of training materials covering a wide range of themes:

- 1) Implementation of FAIRness of data (covering Acquisition, Collection, Quality Control, Storage, Harmonization of data formats and preparation of Data Management Plans), and possibly other types of digital assets (e.g., software, services and VREs) needed by ENVRI communities;
- 2) Implementation of FAIR metadata (covering Metadata standards, Provenance, Ontologies, Catalogues, Accessibility and Interoperability);
- 3) Implementation of FAIR data re-use (covering Curation, Long-term preservation/archiving and findability & accessibility); and
- 4) Legal and policy-related aspects of FAIR implementation.

1.2 The motivation for a gap analysis

In order to avoid duplication of effort, WP6 should build on materials and documentation from other initiatives, such as the Research Data Alliance, (RDA)¹, FORCE11², DataOne³ and GO-FAIR⁴, as well as existing knowledge in the ENVRI-FAIR partner RIs. To further streamline and facilitate both the planning of training events and the production of relevant resources, the first task of WP6 has been to undertake a gap analysis, the results of which are reported in this Deliverable.

The gap analysis has focused on two components: 1) an inventory of existing training materials and resources; and 2) a survey of the perceived needs & requirements for FAIR training of both established and emerging environmental and Earth science research infrastructures.

1.3 Identifying relevant FAIR training topics

The training materials gathered and/or developed by the WP should provide detailed and practical information of all FAIR principles, with special focus on addressing the needs of both established and developing ENVRI. The materials should help RIs to implement FAIR best practices in their data management, by building on existing practices but simultaneously expanding and updating these in order to enable data centres and portals to serve their respective end users with FAIR data products and services.

Which topic could then be considered as relevant to the intended audience? Based on the overall themes outlined above in Chapter 1, two main subject categories may be distinguished: "general FAIR-related" and "research data management-related". Based on discussions with representatives of the project partners, a number of concrete topics were

¹ <https://rd-alliance.org/>

² <https://force11.org/>

³ <https://dataone.org/>

⁴ <https://go-fair.org/>

identified; these are listed in Table 1 below. NOTE: in the following, we will refer to the individual topics using the codes G1-G7 and R1-R17, respectively.

Table 1: The training topics considered in the gap analysis. Note that the topic codes (G1-G7 and R1-R17) are frequently used in the report to refer to individual topics.

General FAIR-related training topics	
G1	Introduction to FAIR principles
G2	Metrics for FAIRness evaluation
G3	Performing a FAIRness self-assessment
G4	GDPR (General Data Protection Regulation ⁵) issues related to data sharing
G5	Basic Research Data Management (RDM)
G6	Writing technical documentation for services
G7	Other
Research Data Management (RDM) training topics	
R1	Access control (Authorization-Authentication-Identification, or AAI) methods
R2	API (Application Program Interface) design for data & metadata access
R3	Cataloguing - design & implementation
R4	Certification schemes for repositories (CoreTrustSeal)
R5	Cloud computing (Virtual Machines & containers) for data processing
R6	Data Management Plans
R7	Landing page design
R8	Licenses & policies for data use
R9	Linked Data and ontologies
R10	Metadata standards & schemas (including geospatial, instruments, variables)
R11	PID allocation & use (including citation support, bibliometry, provenance)
R12	Portal design & operation
R13	Provenance tracing
R14	Repository design, operation & sustainability
R15	Virtual Research Environments for data analysis (design & implementation)
R16	Workflow engines for automated data processing
R17	Other

Chapter 2 and Chapter 3 describe how the relevant information was collected and processed, and presents the outcomes of the analysis, for the training resources and RI needs and requirements, respectively. The implications of the gap analysis on the rest of the Task 6.1 activities are discussed in Chapter 4, with relevant conclusions presented in Chapter 5. Finally, the expected impacts on the project and on stakeholders are summarized in Chapters 6 and 7.

⁵ The General Data Protection Regulation 2016/679 is a regulation in EU law on data protection and privacy for all individual citizens of the European Union and the European Economic Area, see <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

2 Inventory of existing training materials on FAIR and related RDM topics

2.1 Materials and methods

Basic information searches were undertaken on the Internet via the search engine Google during the time period 2019-07-05 - 2019-08-28. The following search concepts and keywords were used:

- FAIR data
- FAIR data + training
- research data management + training
- data science skills
- data stewardship
- FAIR data + environmental science

The result of the information searches are presented in Table 2 below. The list does not pretend to be exhaustive, but rather indicates a snapshot of the current state concerning FAIR and RDM training. The WP6 team will continue to monitor external resources, with the intention to provide an up-to-date catalogue.

Table 2: List of training resources and materials related to FAIR data principles and Research Data Management. Relevant ENVRI-FAIR WP6 training topics are indicated.

	Resource title	Organization; URL	Topic codes
T1	Author Guidelines	Enabling FAIR Data; http://www.copdess.org/enabling-fair-data-project/author-guidelines/	G1; G5
T2	Best practices for creating reusable Dryad data packages	DRYAD; https://datadryad.org/pages/reusabilityBestPractices	G4; G5; R1; R8; R13
T3	Data and the FAIR principles	ReproNim; http://www.repronim.org/module-FAIR-data/	G1; G4; G5; R9; R16
T32	Data FAIRness in Environmental & Earth Science Infrastructures	LifeWatch and ENVRIplus International Summer School, Lecce, Italy, July 2019; https://www.lifewatch.eu/web/guest/iss-data-fairness	G1; G2; G3; G5; R2; R3; R6; R8; R9; R10; R11; R12; R13; R16
T4	Data Management Expert Guide	CESSDA; https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide	G1; G5; R3; R4; R6; R9; R10
T33	Data Management in Environmental & Earth Science Infrastructures	LifeWatch and ENVRIplus International Summer School, Lecce, Italy, July 2018; http://www.servicecentrelifewatch.eu/web/guest/envriplus-lifewatch-iss-data-management	G5; R3; R5; R9; R10; R11; R13
T5	Data Management Planning	DTL - Dutch Tech centre for Life Sciences; https://www.dtls.nl/fair-data/research-data-management/data-management-planning/	G5; R6
T6	Data Management Training (DMT) Clearinghouse	Data Management Training (DMT) Clearinghouse; http://dmtclearinghouse.esipfed.org/	G1; G5; R6; R10
T7	Data science and reproducibility (course)	INCF Training Space; https://training.incf.org/course/data-science-and-reproducibility	G1; G3; G5

Table 2 continued

T8	Data Stewardship Wizard	Data Stewardship Wizard; https://ds-wizard.org/	G1; G5; R6
T9	DataONE Best Practices database	DataONE; https://www.dataone.org/best-practices	G5; R6; R10
T10	DataONE Education	DataON; https://www.dataone.org/Education	G5
T34	EGI Webinar Programme	EGI Foundation; https://wiki.egi.eu/wiki/EGI_Webinar_Programme	G5; R5; R15; R16
T11	Enabling FAIR Data – FAQs	Enabling FAIR Data; http://www.copdess.org/enabling-fair-data-project/enabling-fair-data-faqs/	G1; G5
T12	Essential 4 Data Support	Research Data Netherlands; https://datasupport.researchdata.nl/en/	G1; G5
T13	FAIR Data Training - DTL	DTL - Dutch Techcentre for Life Sciences; https://www.dtls.nl/fair-data/fair-data-training/	G1; G5
T14	FAIR Data Training - ANDS	ANDS - Australian National Data Service; https://www.ands.org.au/working-with-data/fairdata/training	G1; G3; G5; R1; R3; R4; R8; R9; R10; R11
T15	FAIR for Beginners	The DeIC e-science knowledge portal; https://vidensportal.deic.dk/en/FAIR	G1; G5
T16	FAIR Principles Training	PARTHENOS; http://training.parthenos-project.eu/sample-page/manage-improve-and-open-up-your-research-and-data/introduction-to-research-data-management/the-fair-principles/	G1; G4; G5; R4; R6; R8; R9; R10
T17	FAIR self-assessment tool	Australian Research Data Commons (ARDC); https://www.ands-nectar-rds.org.au/fair-tool	G1; G3
T18	FAIR Training	Phortos Consultants; http://www.phortosconsultants.com/Fair-training	G1; G2; G3; R3; R8; R9; R10; R13; R16
T19	FAIRsharing.org	FAIRsharing.org; https://fairsharing.org/	G1; G2; G3; G4; G5; G6; G7
T20	FAIR FAQ	GO FAIR; https://www.go-fair.org/faq/	G1
T21	Farm Data Train	DTL - Dutch Techcentre for Life Sciences; https://vimeo.com/215975839	G1; R1; R10; R11
T22	GO FAIR Training	GO FAIR; https://www.go-fair.org/training/	G1; G5
T23	How to make your data FAIR, Guide for Researchers	OpenAIRE; https://www.openaire.eu/how-to-make-your-data-fair	G1; G5
T24	Personal Health Data Train	Leiden University; https://vimeo.com/143245835	G1; R1; R10; R11
T25	RDM Starter Kit	GO FAIR; https://www.go-fair.org/resources/rdm-starter-kit/	G1; G5; R6; R8; R11
T26	TeSS: ELIXIR's Training Portal	ELIXIR; https://tess.elixir-europe.org/	G1; G5; R8; R9; R10; R16
T27	Training	DANS (Data Archiving and Networked Services); https://dans.knaw.nl/en/training-consultancy/training	G1; G5

Table 2 continued

T28	Training Resources - CESSDA	CESSDA; https://www.cessda.eu/Training/Training-Resources	G1; G5
T29	Training, Webinars & Workshops	Environmental Data Initiative (EDI); https://environmentaldatainitiative.org/events/training-webinars-workshops/	G1; G5; R2; R10; R15
T30	Webinars	EBI/EMBL; https://www.ebi.ac.uk/training/webinars	G5; R9; R13
T31	Webinars and Tutorials	NASA EARTH Data; https://earthdata.nasa.gov/learn/user-resources/webinars-and-tutorials	G5; R17

2.2 Summary per topic category

2.2.1 General FAIR-related training topics

The resources on FAIR in general often emanate from disciplinary research infrastructures, national data services or special initiatives. The resources are in various formats, ranging from guidelines to training modules and webinars. They mostly refer to the original FAIR data principles, and do not bring up the aspects of implementing the FAIR data principles in practise.

2.2.2 Research Data Management (RDM) training topics

The field of research data management has grown tremendously during the last years, as evidenced by the large number of, e.g., articles and books published on related ICT (Information and Communications Technology) topics written from an academic or industry point of view. However, the development of training resources relevant to enhancing the FAIR RDM skills and competences of ENVRI community data practitioners has not kept the same pace. Consequently, the availability of elaborated course materials covering actual FAIRness-promoting best practices and/or implementations of data management technologies is very limited. Indeed, among the more comprehensive "R topic" resources that we identified are two summer schools co-organized by the ENVRI-FAIR predecessor ENVRIplus – a good indicator for the importance of the work that WP6 is now undertaking.

3 Survey of ENVRI-FAIR partners' knowledge level and training needs

3.1 Materials and methods

3.1.1 Input from WP7 activities

During the period March-June 2019, Work Package 7 ("Common implementation and support"), with support from Work Package 5 ("Common requirements and testbed for (meta)data services, community standards and cataloguing"), asked the project partner RIs to complete two FAIR assessment questionnaires. The first of these was a generic survey of the current status of FAIRness implementation, while the other was more technical in nature and focused on details of the support for machine-actionable evaluation of FAIRness of data objects.

As a follow-up to the surveys, each of the four ENVRI-FAIR sub-domains (Atmosphere, Marine, Solid Earth and Ecosystem & Biodiversity, represented by WPs 8-11) held a workshop where the survey outcomes were discussed with representatives from WP5 and WP7. WP6 was not directly involved in these activities, but our sub-domain contact persons for training-related issues collected their colleagues' experiences of the applied FAIRness evaluation criteria and the methodology of self-assessment.

While it was somewhat difficult to directly convert these views and impressions into quantifiable information on training needs, they were nevertheless taken into account in the design of the WP6 questionnaire.

3.1.2 The WP6 questionnaire

The questionnaire "RI training needs & requirements" was distributed on 2019-05-24 to all ENVRI-FAIR RIs via the training issue contact persons identified by each infrastructure. The WP6 sub-domain contacts, representing each of the WPs 8-11, acted as area-specific coordinators. The questionnaire was structured into two main parts; FAIR training and RDM training with four different questions within each part, with an additional 7 questions related to the respondents as well as allowing for general comments. The questionnaire in its entirety is included in Appendix B⁶.

In total, we received 26 answers: all but two of the partner RIs responded, with some of the distributed infrastructures submitting more than one set of answers. The fact that the overall number of responses is low, together with the risk that the input from the respondents may be biased towards the individual respondents' views rather than be reflective of their organisations' opinions, combine to make it difficult to draw statistically significant conclusions. On the other hand, we have received quite detailed information regarding FAIR and RDM training needs and experiences from these persons. **Figures C-1** and **C-2** in **Appendix C** illustrate the variability in the knowledge levels and training need priorities reported by the 12 RIs that completed the questionnaire.

We also note that the design of some of the questions in the survey was unclear, leading to somewhat incomplete, and in some cases contradictory answers. The main problem occurred as a consequence of questions Q5 and Q9 combining two questions in one, as exemplified by Q5: "Which general FAIR-related topics are personnel in your RI knowledgeable about, or have received training on previously?". Here, we expected respondents to select at least two options, one from the set {"No knowledge", "Some knowledge", "Very knowledgeable"} and one from the other option set {"Have undergone (some) training", "Able to train others!"}, or to indicate that they didn't have any relevant information by selecting "Don't know". In order to make it possible to extract and visualise representative response frequency histograms and calculate priority rankings, responses were adjusted as indicated by Table 3.

⁶ The collected responses to the questionnaire contain information that is personal and/or sensitive in nature, and can therefore not be disseminated in the original form. The response database will however remain accessible to reviewers until the end of the ENVRI-FAIR project, after which it will be permanently deleted. (Contact manager@envri-fair.eu for assistance.)

Table 3: Translation table applied to clarify responses to questions Q5 and Q9.

Original response	Interpreted/modified response
ONLY "Have undergone (some) training"	"Some knowledge, Have undergone (some) training"
ONLY "Able to train others"	"Very knowledgeable, Able to train others"
"No knowledge, Don't know"	"No knowledge"
"No knowledge, Some knowledge"	"Some knowledge"
"No knowledge, Some knowledge, Very knowledgeable"	"Some knowledge"
"No knowledge, Very knowledgeable"	"Some knowledge"
"Some knowledge, Very knowledgeable"	"Very knowledgeable"
"No knowledge, Have undergone (some) training"	"Some knowledge, Have undergone (some) training"
"Some knowledge, Very knowledgeable, Have undergone (some) training"	"Very knowledgeable, Have undergone (some) training"

3.2 Summary per topic category

3.2.1 General FAIR-related training topics

All respondents have declared to have *Some knowledge* or be *Very knowledgeable*, when it comes to the general FAIR topics. Some respondents (12 out of 26 persons) stated they or their colleagues had attended some previous training, and in some cases (7 of 26) they declared that their organisations would be able to contribute with trainers for selected topics. The distribution of knowledge in the sub-topics differs among the respondents, but Introduction to FAIR (G1) and Basic RDM (G5) are the sub-topics in which the respondents have the highest knowledge level.

Following up on the needs for training the sub-topics getting most marks as "High Priority" were: Metrics for FAIRness evaluation (G2), Performing a FAIRness self-assessment (G3), GDPR issues related to data sharing (G4). The sub-topic Basic RDM (G5) was rated as "Low priority" for training needs. This indicates that most of respondents have basic knowledge both in the FAIR principles and in research data management (RDM), but that they have a need to deepen into the measurement and assessment of FAIR and into the issues of GDPR. **Figure 1** below illustrates the status and needs related to topic G4. The corresponding histograms for all "G" topics are shown in **Appendix C, Figure C-3**

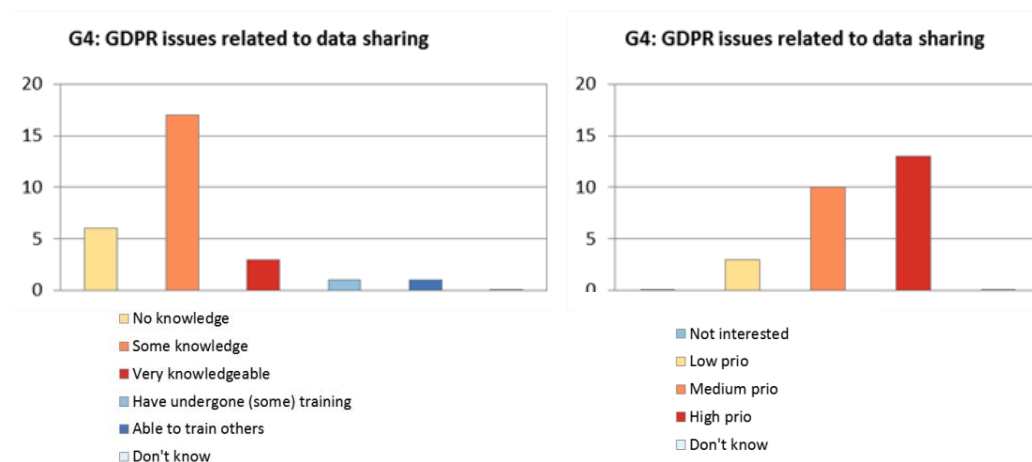


Figure 1: Histograms summarizing all responses related to (left) the current level of knowledge and training experience in the RIs, and (right) the corresponding estimated priority, for the "general FAIR" training topic G4.

Figure 4 below shows an overview of the availability of resources across the "G" sub-topics. Comparing knowledge and training needs in FAIR General topics with our inventory on existing training materials, there are adequate resources covering the sub-topics G1-G5, but only two for G6 - Writing technical documentation for services.

Concerning the issue of which specific RI staff categories are in need of FAIR related training, the opinions are quite scattered. For the roles "RI managers & administrators" and "Data managers/IT specialists" it is usually estimated as 1-5 persons per research infrastructure, while for the role "Researcher in your RI" the estimates vary considerably, with some RIs responding more than 10 persons, some 1-5 persons and some "don't know".

3.2.2 Research Data Management (RDM) training topics

Judging from the responses, the level of knowledge concerning the advanced Research Data Management topics is again quite varied. Some respondents (12 out of a total of 26) declared that personnel in their RIs have attended some previous training, and 8 out of 26 stated that they or their colleagues are able to train others in these technology-related Research Data Management sub-topics.

For the sub-topics Application Programmable Interfaces (R1) and Data management plans (R6), all the respondents declared to have *Some knowledge* or to be *Very knowledgeable*. Regarding the other topics, the current knowledge level was reported as relatively high, but some organisations (especially RIs that are now in their establishment phase) estimated that they had low or even very low knowledge of a majority of the proposed sub-topics.

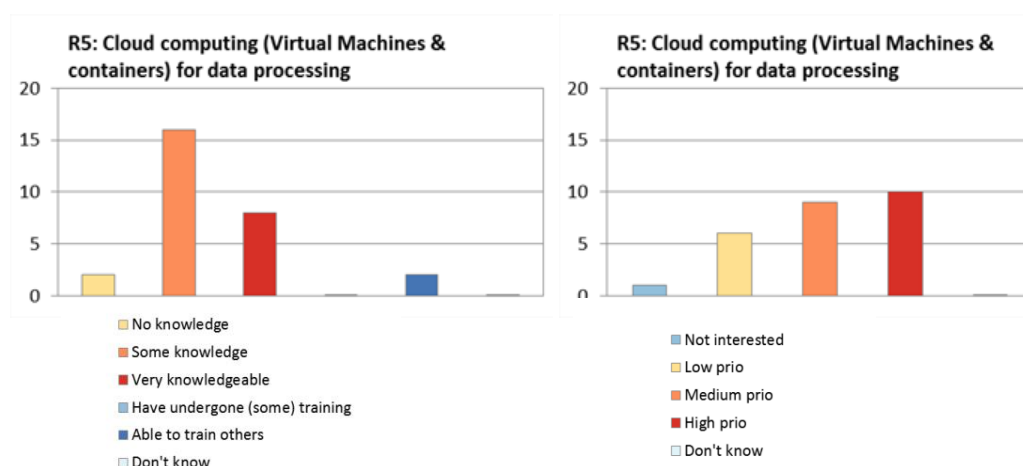


Figure 2: Histograms summarizing all responses related to (left) the current level of knowledge and training experience in the RIs, and (right) the corresponding estimated priority, for the "RDM-specific" training topic R5.

Regarding the needs of RDM training, the following topics were rated "High priority": Cloud computing for data processing (R5), Licences & policies for data use (R8), Provenance tracing (R13) and Workflow engines for automated data processing (R16). These sub-topics are quite specialized, which may indicate that the respondents have quite a good knowledge on basics of RDM, but need to deepen their understanding of specific tasks or components. **Figure 2** illustrates the current status and expected future needs for topic R5. The corresponding histograms for all "R" topics are shown in **Appendix C, Figures C-4 and C-5**.

Comparing the needs of RDM training with the inventory list (**Table 2**) there are at least some training resources covering the prioritized needs. The "R" topics with the fewest available resources are R1 (2), R2 (2), R5 (1), R7 (0), R12 (1), R14 (0) and R15 (2). (See also **Figure 4** below.)

When it comes to the number of people in a research infrastructure needing RDM-related training, there is about the same trend as in the training needs for FAIR related training, see **Figure 3** below.

4 Implications for Task 6.1 activities

4.1 Training materials & resources

The training resources are often in the form of guidelines or self-study modules. Webinars are also getting more and more common as a training form, and many organisations are listing their recorded webinars, sometimes tagged with subjects. The training resources need to be tested and evaluated in a systematic way, and this will be organised later by the WP6 team. We note that in several cases, respondents to the questionnaire indicated that their organisation's level of knowledge on specific topics is very high, and that they are willing and prepared to help train personnel from other RIs.

The ENVRI Community Training Platform (CTP) is an important dissemination point for training and information about training resources. Initially developed during the ENVRIplus project, the platform is now undergoing a major update under the leadership of LifeWatch ERIC.

4.2 Training events

Organising training events of different types is an integral part of WP6 activities. Within this framework, both webinars and face-to-face training events will be offered, both those targeting all project members and others with a sub-domain-specific audience in mind. In the latter case, co-location with sub-domain meetings could be an option, while topics of cross-domain interest may be addressed in conjunction with e-g- ENVRI-FAIR annual meetings or other Earth Science -related conferences and workshops.

Organisations external to ENVRI-FAIR are also offering regular training events such as courses, programs, and workshops, and there are several examples of this in the list of resources. As FAIR and RDM continue to increase in terms of usage and attention, we foresee many training events to appear in the near future. The WP6 team will continue to monitor upcoming events and regularly communicate information about these to the ENVRI-FAIR community, e.g., via e-mails, the calendar maintained by the project office, and – as soon as it becomes operational – the new Community Training Platform.

4.3 Target audiences for training

A very important aspect that must be considered when both during development of training resources and the planning of training events is the personnel category for which the training is meant. Here, we believe it is meaningful to consider at least four different groups: RI

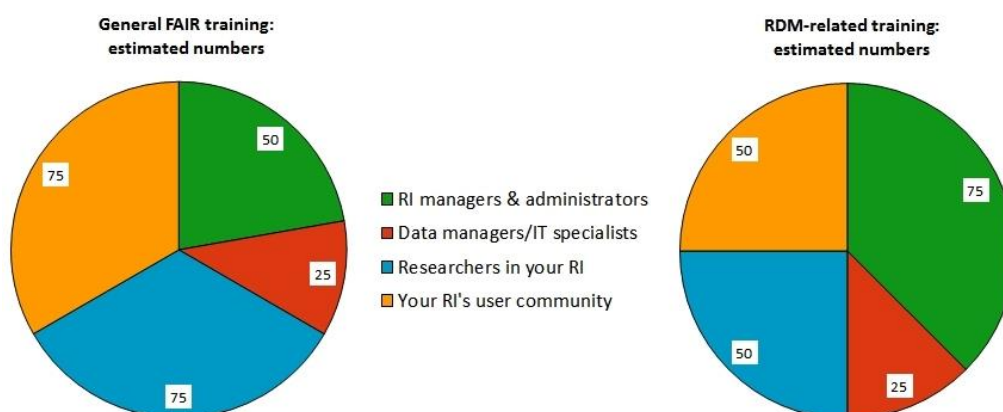


Figure 3: Preliminary estimate of total numbers (summed over all sub-domains) of individuals that will require FAIR-related training during the ENVRI-FAIR project, divided according to staff category.

managers & administrators, ICT experts & data centre staff, researchers working in the RIs, and end users of RI services and data including junior scientists. Obviously, the current level

of knowledge must be taken into account, both when recommending self-study materials and carrying out specific learning activities such as lectures and webinars. Our survey asked respondents to roughly indicate the number of people in those different staff categories that could need training. **Figure 3** summarizes the outcome, which should however only be seen as a very preliminary indicator.

4.4 Training requirements according to priorities

In the questionnaire, respondents were asked to attach priorities to the topics as laid out in Chapter 2.2.1 to identify from a 'demand perspective' what RI's considered as the most needed training opportunities for their communities. We analysed the answers to those questions and selected the training topics with a score⁷ over 70% for medium or high priority as relevant for the first activities of WP6 in actually offering options to participate in training. The level of these training activities then depends on the level of knowledge as indicated in the questionnaire as well, and analysed in Chapter 2.2.

As analysed, eight topics were identified as having high priority and WP6 will focus on offering training for these:

- G2: Metrics for FAIRness evaluation
- G3: Performing a FAIRness self-assessment
- G4: GDPR issues related to data sharing
- G6: Writing technical documentation for services
- R5: Cloud computing (Virtual Machines & containers) for data processing
- R8: Licences & policies for data use
- R13: Provenance tracing
- R16: Workflow engines for automated data processing

Figure 4 below shows the coverage of all FAIR and RDM topics by the identified training resources. For all of these topics, external parties are offering at least some training assets. However, we note that in the case of R5, this sub-topic is only covered by the T34 resource – materials developed for the 2019 edition of the LifeWatch & ENVRIplus summer school. Here, the search for complementary materials from external sources should be expanded in order to broaden the scope of skills-building offerings.

A number of considerations and risks need to be taken into account:

- The overall number of responses to the questionnaire (26) is relatively low considering the typically large communities behind the RIs, both in terms of technical and management staff working in the RI itself or in distributed nodes and the scientists using the RI or providing data or services. Although it isn't unreasonable to assume that questionnaire respondents will indeed have consulted with colleagues and co-workers, from a statistical standpoint we only have 26 data points. This of course lowers the statistical power and significance of the analysis, but we assume that at least the set of training priorities will contribute considerably to the needs of the overall ENVRIFAIR community;
- Although we found that there are available training offerings for these eight prioritized topics, in some cases the number of externally produced training resources is low (say, less than 3). This poses a risk that the specific content of the materials isn't tailored to the needs of the RI's or the type of data we handle and offer. This would require development of such training from scratch;

⁷ The score was calculated for each sub-topic as the ratio of ("Medium prio" + "High prio") to the total number of valid responses, i.e. discarding "Don't know".

- In case WP6 needs to develop training, the ENVRI-FAIR Summer School under planning is an excellent opportunity to do so – especially for the Cloud Computing topic (R5).

All respondents

	G1	G2	G3	G4	G5	G6	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	Covered topics:
	0.54	0.85	0.81	0.88	0.46	0.72	0.60	0.62	0.58	0.69	0.73	0.50	0.48	0.73	0.68	0.65	0.65	0.46	0.84	0.60	0.64	0.81	
T1	X				X																		2
T2				X	X		X							X					X				5
T3	X			X											X							X	5
T4	X				X			X	X		X				X	X							7
T5					X	X																	2
T6	X				X							X				X							4
T7	X		X		X																		3
T8	X				X							X											3
T9					X							X				X							3
T10					X																		1
T11	X				X																		2
T12	X				X																		2
T13	X				X																		2
T14	X		X		X		X		X	X				X	X	X	X						10
T15	X				X																		2
T16	X			X	X					X		X		X	X	X							8
T17	X		X																				2
T18	X	X	X					X						X	X	X			X			X	9
T19	X	X	X	X	X	X																	6
T20	X																						1
T21	X																						1
T22	X				X																		2
T23	X				X																		2
T24	X				X							X		X			X						5
T25	X				X							X		X			X						5
T26	X				X							X		X			X						5
T27	X				X																		2
T28	X				X																		2
T29	X				X			X								X						X	5
T30					X										X				X				3
T31					X																		1
T32	X	X	X		X			X	X			X		X	X	X	X	X	X	X		X	14
T33					X			X							X	X	X		X				6
T34					X						X										X	X	4
Number of resources:	26	3	6	4	30	2	2	2	5	3	1	9	0	8	8	9	6	1	5	0	2	4	

Figure 4: Availability of training resources for general FAIR (G1-G6) and RDM-specific (R1-R16) topics, with focus on the eight medium (light red columns) and high (light blue columns) priority areas identified in the analysis of all responses. White crosses in dark blue and dark red cells identify relevant training resources. Numbers along the bottom and the right side indicate the total numbers of found resources and covered topics per resource. Row 2 contains the all-respondents priority scores for each topic.

4.5 Domain-specific needs

The same type of priority analysis that was done for all responses was also performed for the four sub-domains. Table 4 below summarises the outcome across all suggested training topics. In addition, Figures **C-6**, **C-7**, **C-8** and **C-9** in **Appendix C** contain the sub-domain equivalents of the resource-topic matrix of Figure 4.

Some differences in prioritisation are indicated between the “all” analysis reported above and the individual sub-domains. However, it is very important to note that the number of respondents per sub-domain lies between 6 and 11, so the statistical significance is even lower, and the personal preferences and needs of the respondents could bias the outcome for the entire sub-domain. Thus, the variations should not be over-analysed.

For this reason, WP6 proposes to focus on those topics that will be useful for several individual sub-domains and/or score relatively high for all respondents:

- *Atmosphere*: R9 Linked data and ontologies
- *Ecosystem & Biodiversity*: R9 plus R4 Certification schemes for repositories (CoreTrustSeal)
- *Marine*: R4 plus R10 Metadata standards & schemas (including geospatial, instruments, variables)
- *Solid Earth*: R4

WP6 will consult the sub-domains first about the actual needs, before going into detailed planning of activities. Sub-domains should check internally with a wider audience about this before ENVRI-FAIR WP6 undertakes action.

5 Summary & conclusions

This deliverable is not to be considered as a final, complete and closed document. WP6 will continue to update the lists of training opportunities, training experts within the project and training needs in the consortium. Furthermore, we plan to send out a second questionnaire around Month 36 in order to gauge the changes in the project partners’ self-assessment of their knowledge level with respect to the “general FAIR” and “RDM-related” training topics

Through the questionnaire we have at an early stage in the ENVRI-FAIR project managed to collect valuable opinions regarding the current knowledge and the training needs of FAIR and research data management (RDM). Although there was a relatively small number of respondents to the questionnaire, 26 persons, their answers are very detailed and indicate clearly the current state of FAIR and RDM.

Looking on what is currently available on the market in terms of training resources we may conclude that there exist external training resources within sub-topics that are rated “high priority” by the respondents in the questionnaire. There is about 1-3 external training resources for each high priority sub-topic, which is not so much, but at least something to start-up with for ENVRI-FAIR.

The list of training resources is relatively short, but we concentrated our search on finding resources focused on training of FAIR or RDM, and in this aspect the field is very much under development. We foresee that there will be many training resources developed in the near future, in particular as both the European Union and many national research funders are focusing on FAIR and RDM. The WP6 team will follow this development closely.

Table 4: Priority scores of the proposed training topics for all respondents and split per sub-domain. Blue and red colour indicate medium-range and high priority, respectively. Orange cells in sub-domain rows represent additional prioritised training topics not identified in the 'all respondents' analysis.

General FAIR topics G1-G6								
Domain	G1	G2	G3	G4	G5	G6		
All	0.54	0.85	0.81	0.88	0.46	0.72		
Atm	0.17	0.66	0.66	0.83	0.17	0.66		
E&B	0.64	0.82	0.73	0.91	0.64	0.73		
Mar	0.62	1.00	1.00	0.87	0.50	0.75		
Sol	0.54	0.85	0.81	0.88	0.46	0.72		
RDM-specific topics R1-R8								
Domain	R1	R2	R3	R4	R5	R6	R7	R8
All	0.60	0.62	0.58	0.69	0.73	0.50	0.48	0.73
Atm	0.50	0.50	0.66	0.50	0.50	0.17	0.33	0.50
E&B	0.82	0.64	0.64	0.73	0.64	0.82	0.18	0.73
Mar	0.50	0.75	0.87	1.00	1.00	0.75	0.62	0.62
Sol	0.60	0.62	0.58	0.69	0.73	0.50	0.48	0.73
RDN-specific topics R9-R16								
Domain	R9	R10	R11	R12	R13	R14	R15	R16
All	0.68	0.65	0.65	0.46	0.84	0.60	0.64	0.81
Atm	0.83	0.66	0.50	0.33	1.00	0.33	0.50	0.66
E&B	0.73	0.64	0.64	0.45	0.73	0.45	0.54	0.73
Mar	0.62	0.87	0.75	0.50	0.87	0.75	0.75	0.87
Sol	0.68	0.65	0.65	0.46	0.84	0.60	0.64	0.81

The topic codes (G1-G5 and R1-R16) are explained in Table 1.

"All": sum over all sub-domains;

"Atm": Atmosphere (WP8 members)

"E&B": Ecosystem & Biodiversity (WP11 members)

"Mar": Marine sub-domain (WP9 members)

"Sol": Solid Earth sub-domain (WP10 members)

The main conclusion from the outcomes of the questionnaire is the focus on eight initial topics for training on FAIR and RDM, as mentioned in 4.3. WP6 proposes to the RI's work on development in one way or the other these eight topics. WP6 will take into account a quality check on the level and content of the training as offered on the market (and presented in Table 2). Also the knowledge level of the ENVRI community will be taken into account when developing training as well as an analysis of the target audience.

A second conclusion from the questionnaire is the need for additional training for individual sub-domains in the ENVRI-FAIR project. WP6 has a list of potential training needs and will consult the sub-domains about needs and wish lists for additional training.

Finally, the questionnaire indicated that within the project, the consortium or more widely the ENVRI community quite some training expertise is available. WP6 will in the near future consult the RI's in more detail about persons available, willing and capable of contributing to specific training activities.

6 Impact on the project

This deliverable is considered by WP6 as a starting point for organizing the training activities in the ENVRI-FAIR project. It is based -as said- on a limited number of responses, certainly considering the amount of people directly or indirectly involved in the project, but it offers a clear starting point to which both WP6 and the sub-domains in the project can agree on. The conclusions and steps forward for the immediate future as presented in section 5 are consulted with the sub-domains and supported by them and give WP6 staff a clear way to deliver useful training to the project as well as valuable input for the training platform as will be developed in Task 6.1

7 Impact on stakeholders

WP6 is aware that the community of ENVRI related RI's is much bigger than those immediately involved in the ENVRI-FAIR project, both in terms of individual staff and RI's as such. The analysis of target audiences and the more detailed dialogue that WP6 will have with the sub-domains should ensure that the full ENVRI community will benefit from the activities in WP6.

Acknowledgements

We gratefully acknowledge the valuable contributions that we received from the representatives of the research infrastructures across all the ENVRI-FAIR sub-domains.

Appendix A: Glossary and terminology

NOTE: The latest version of the master list of the glossary is available at <http://doi.org/10.5281/zenodo.3465753>.

The following is a list of acronyms and terms used in this deliverable:

AAI - Authorization-Authentication-Identification

API - Application Program Interface

CTP - Community Training Platform

Data Centre - a large group of networked computer servers typically used by organizations for the remote storage, processing, or distribution of large amounts of data

ENVRI Community - Environmental Research Infrastructures community

ENVRIplus - Cluster Project for the ENVRI community 2015-2019

ENVRIs - Environmental Research Infrastructures

FAIR - Findable, Accessible, Interoperable and Reusable

FORCE11 - Future of Research Communications and e-Scholarship

GDPR - General Data Protection Regulation

RDA - Research Data Alliance

RDM - Research Data Management

RI - Research Infrastructure

Sub-Domain - A defined thematic set of Research Infrastructures within ENVRI working on either the Atmosphere, Marine, Solid Earth or Ecosystems/Biodiversity Research

Webinar - a seminar conducted over the Internet

Appendix B: Survey of RI training needs & requirements

This appendix is an off-line copy of the Google Forms-based questionnaire "ENVRI-FAIR WP6 training needs & requirements" that was active May 24-June 28, 2019.

All questions marked with "*" were compulsory, and for the "array questions" (Q5, Q6, Q8, Q9, Q10 and Q12), at least one option had to be selected for each row.

INTRODUCTION

Welcome to this survey where we are collecting information about the status of knowledge in the project member RIs about FAIR-related concepts in general, as well as specific technologies related to FAIR research data management!

The responses we collect will be used for the gap analysis of Task 6.1, where we want to map out and better understand the project partners' needs and requirements for training on both general issues of FAIR and more technical Research Data Management (RDM) topics. The analysis, done in close collaboration with the sub-domain WPs, will be reported in deliverable D6.1 (due at the end of August).

We expect that the survey will be completed at least once for each ENVRI-FAIR member RI, and that the collection of the input required to complete the survey will be collected by the person(s) designated as training contact persons for their respective RI.

If you have questions or need help to fill out the form, we suggest that you get in touch with the WP6 representative of your RI's sub-domain, or with the WP6 co-leads.

Thank you very much in advance for your help!

Maggie, Jacco & Nicola (the WP6 co-leads)

Q0: Are you OK with your responses being stored digitally, and used in the ENVRI-FAIR project context? *

You can only access the rest of the questionnaire if you answer "Yes" to this question. Note that the collected information will be permanently deleted at the end of the ENVRI-FAIR project. You can also contact the ENVRI-FAIR WP6 co-chairs at any time and ask for your a copy of the data you submitted, and/or request that it be removed.

BASIC CONTACT INFO

This information will only be used internally in WP6, for example if we need to contact you for clarifications or additional information.

Q1: Your name *

Q2: Your e-mail *

Q3: Research infrastructure *

If your RI is distributed over many institutes, or covers many sub-domains, please fill out a response for each of these separately! In this case, label the different responses by giving them names like "ICOS (Marine)" or "EPOS (INGV)".

Q4: Sub-domain(s) of your RI *

Select all that apply

Atmosphere

Ecosystem/biodiversity

Marine

Solid Earth

GENERAL FAIR-RELATED TRAINING

Q5: Which general FAIR-related topics are personnel in your RI knowledgeable about, or have received training on previously? *

Please select all appropriate options for each of the suggested topics. If you can't answer, please select the "Don't know" option!

Options: "No knowledge", "Some knowledge", "Very knowledgeable", "Have undergone (some) training", "Able to train others!", "Don't know"

Topics:

- Introduction to FAIR principles
- Metrics for FAIRness evaluation
- Performing a FAIRness self-assessment
- GDPR issues related to data sharing
- Basic Research Data Management (RDM)
- Writing technical documentation for services

Q6: Which general FAIR-related training should WP6 provide to your RI? *

Please rate the priority for each of the suggested topics. If you can't answer, please select the "Don't know" option!

Options: "Not interested", "Low prio", "Medium prio", "High prio", "Don't know"

Topics:

- Introduction to FAIR principles
- Metrics for FAIRness evaluation
- Performing a FAIRness self-assessment
- GDPR issues related to data sharing
- Basic Research Data Management (RDM)
- Writing technical documentation for services

Q7: Any other FAIR-related issues WP6 should provide?

If we forgot something, please type in your suggestions (with priority ranking) here!

Q8: How many people in your RI need general FAIR training? *

Please give a rough estimate for the period 2019-2020. NOTE: This is *not* a sign-up sheet - we just want to get a first impression of what our training target audiences look like!

Options: "None", "1-5", "6-10", "More than 10", "Don't know"

Categories:

- RI managers & administrators
- Data managers/IT specialists
- Researchers in your RI
- Your RI's user community

RESEARCH DATA MANAGEMENT (RDM) TRAINING

Q9: Which RDM-related topics are personnel in your RI knowledgeable about, or have received training on previously? *

Please select all appropriate options for each of the suggested topics. If you can't answer, please select the "Don't know" option!

Options: "No knowledge", "Some knowledge", "Very knowledgeable", "Have undergone (some) training", "Able to train others!", "Don't know"

Topics:

- Access control (Authorization-Authentication-Identification, or AAI) methods
- API (Application Program Interface) design for data & metadata access
- Cataloguing - design & implementation
- Certification schemes for repositories (CoreTrustSeal)
- Cloud computing (Virtual Machines & containers) for data processing
- Data Management Plans
- Landing page design
- Licences & policies for data use
- Linked Data and ontologies
- Metadata standards & schemas (including geospatial, instruments, variables)
- PID allocation & use (including citation support, bibliometry, provenance)
- Portal design & operation
- Provenance tracing
- Repository design, operation & sustainability
- Virtual Research Environments for data analysis (design & implementation)
- Workflow engines for automated data processing

Q10: Which RDM-related training should WP6 provide to your RI? *

Please rate the priority for each of the suggested topics. If you can't answer, please select the "Don't know" option!

Options: "Not interested", "Low prio", "Medium prio", "High prio", "Don't know"

Topics:

- Access control (Authorization-Authentication-Identification, or AAI) methods
- API (Application Program Interface) design for data & metadata access
- Cataloguing - design & implementation
- Certification schemes for repositories (CoreTrustSeal)
- Cloud computing (Virtual Machines & containers) for data processing
- Data Management Plans
- Landing page design
- Licences & policies for data use
- Linked Data and ontologies
- Metadata standards & schemas (including geospatial, instruments, variables)
- PID allocation & use (including citation support, bibliometry, provenance)

Portal design & operation

Provenance tracing

Repository design, operation & sustainability

Virtual Research Environments for data analysis (design & implementation)

Workflow engines for automated data processing

[Q11: Any other RDM-related issues WP6 should provide?](#)

If we forgot something, please type in your suggestions (with priority ranking) here!

[Q12: How many people in your RI need RDM-related training? *](#)

Please give a rough estimate for the period 2019-2020. NOTE: This is *not* a sign-up sheet - we just want to get a first impression of what our training target audiences look like!

Options: "None", "1-5", "6-10", "More than 10", "Don't know"

Categories:

RI managers & administrators

Data managers/IT specialists

Researchers in your RI

Your RI's user community

COMMENT IS FREE!

Here you can add any other comments or suggestions you may have on what WP6 should be working on.

[Q13: Please share any thoughts you have on training & skills building in ENVRI-FAIR](#)

Appendix C: Supplementary materials

This appendix contains supplementary materials that illustrate details of the gap analysis work.

Summary across RIs: current knowledge levels & priorities

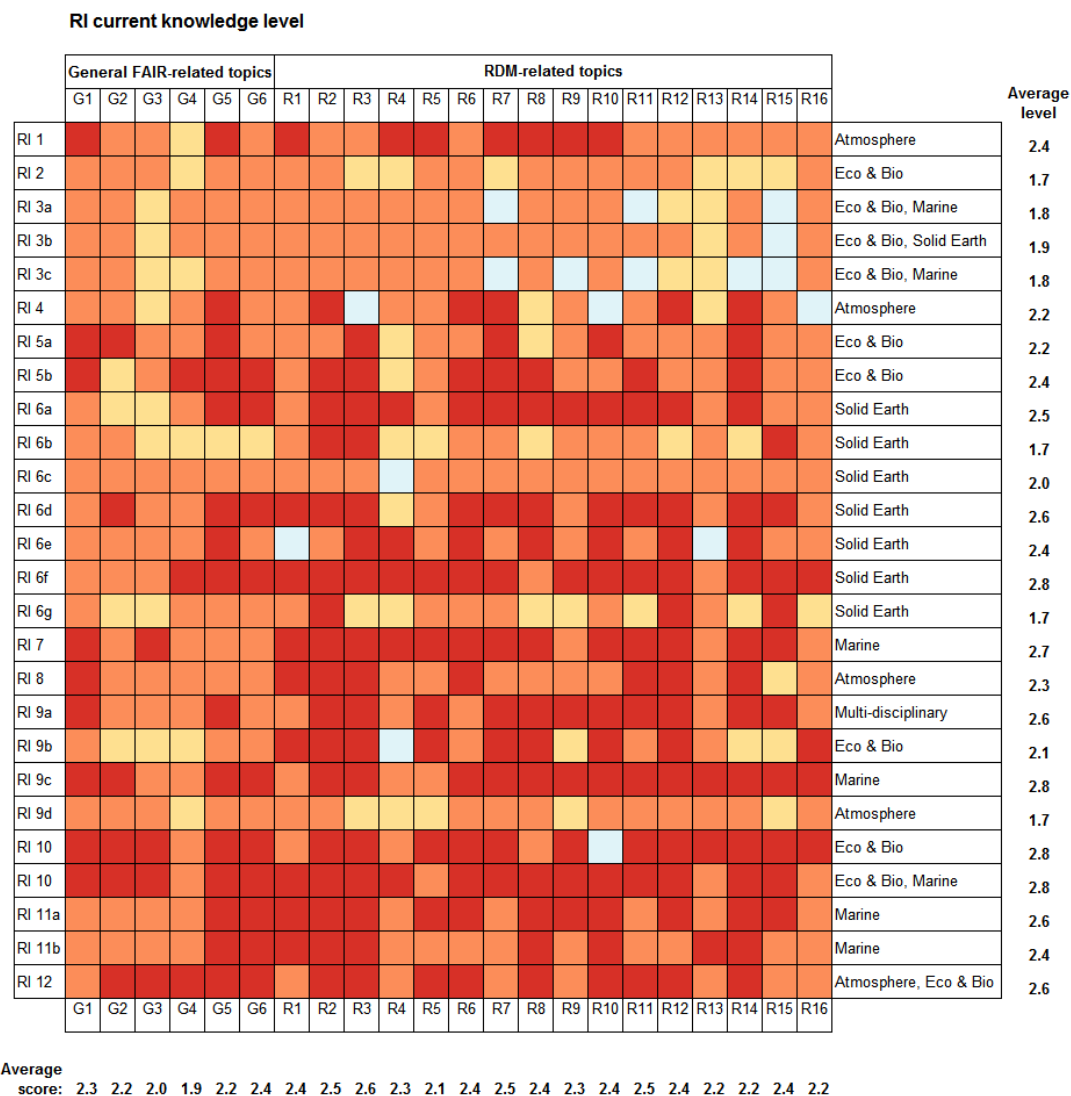


Figure C-1: The current knowledge level for all 12 Research Infrastructures that responded to the survey. The listing is anonymised, but the sub-domain(s) covered by the RI is indicated. The colour scheme is the same as used in the left column of Figure C-3. (Dark red: very knowledgeable; orange: some knowledge; yellow: no knowledge; light grey: don't know.)

RI training priorities

	General FAIR-related topics						RDM-related topics										Average prio								
	G1	G2	G3	G4	G5	G6	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10		R11	R12	R13	R14	R15	R16		
RI 1	Yellow	Dark Red	Dark Red	Dark Red	Yellow	Orange	Orange	Orange	Orange	Orange	Yellow	Yellow	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Yellow	Orange	Orange	Atmosphere	2.0
RI 2	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Light Blue	Dark Red	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Light Blue	Light Blue	Yellow	Orange	Eco & Bio	2.5
RI 3a	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Eco & Bio, Marine	2.6
RI 3b	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Eco & Bio, Solid Earth	2.5
RI 3c	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Eco & Bio, Marine	2.5
RI 4	Dark Red	Orange	Yellow	Orange	Yellow	Light Blue	Yellow	Yellow	Orange	Dark Red	Yellow	Yellow	Dark Red	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Light Blue	Yellow	Orange	Atmosphere	1.8
RI 5a	Yellow	Dark Red	Dark Red	Dark Red	Yellow	Dark Red	Dark Red	Dark Red	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Eco & Bio	1.9
RI 5b	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Eco & Bio	2.7
RI 6a	Orange	Dark Red	Dark Red	Dark Red	Yellow	Orange	Orange	Orange	Orange	Orange	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Solid Earth	1.8
RI 6b	Yellow	Yellow	Dark Red	Dark Red	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Dark Red	Orange	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Solid Earth	2.0
RI 6c	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Dark Red	Orange	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Solid Earth	2.4
RI 6d	Yellow	Dark Red	Dark Red	Dark Red	Dark Red	Light Blue	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Solid Earth	2.2
RI 6e	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Light Blue	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Solid Earth	1.8
RI 6f	Yellow	Dark Red	Dark Red	Dark Red	Orange	Orange	Orange	Orange	Orange	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Solid Earth	1.5
RI 6g	Yellow	Orange	Orange	Dark Red	Yellow	Orange	Dark Red	Orange	Light Blue	Light Blue	Orange	Light Blue	Orange	Orange	Light Blue	Orange	Orange	Orange	Orange	Light Blue	Orange	Orange	Orange	Solid Earth	1.5
RI 7	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Marine	2.3
RI 8	Yellow	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Atmosphere	1.8
RI 9a	Light Blue	Dark Red	Dark Red	Dark Red	Light Blue	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Multi-disciplinary	1.7
RI 9b	Dark Red	Light Blue	Light Blue	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Light Blue	Dark Red	Orange	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Eco & Bio	1.0
RI 9c	Dark Red	Dark Red	Dark Red	Dark Red	Yellow	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Marine	1.9
RI 9d	Yellow	Orange	Orange	Dark Red	Dark Red	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Atmosphere	2.0
RI 10	Orange	Orange	Orange	Dark Red	Orange	Orange	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Eco & Bio	1.3
RI 10	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Dark Red	Eco & Bio, Marine	2.4
RI 11a	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Light Blue	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Light Blue	Dark Red	Orange	Orange	Orange	Orange	Dark Red	Marine	1.7
RI 11b	Dark Red	Dark Red	Dark Red	Dark Red	Orange	Orange	Orange	Orange	Orange	Dark Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Marine	1.7
RI 12	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Atmosphere, Eco & Bio	0.6

Average score: 1.7 2.1 2.4 2.4 2.1 1.9 2.0 1.9 1.8 1.8 2.0 1.9 1.7 1.8 2.0 2.0 2.1 1.8 1.9 2.0 1.8 2.1

Figure C-2: Estimated training needs for all 12 Research Infrastructures that responded to the survey. The listing is anonymised, but the sub-domain(s) covered by the RI is indicated. The colour scheme is the same as used in the right column of Figure C-3. (Dark red: high priority; orange: medium priority; yellow: low priority; light blue: not interested; light grey: don't know.)

General FAIR-related topics: knowledge status & priorities

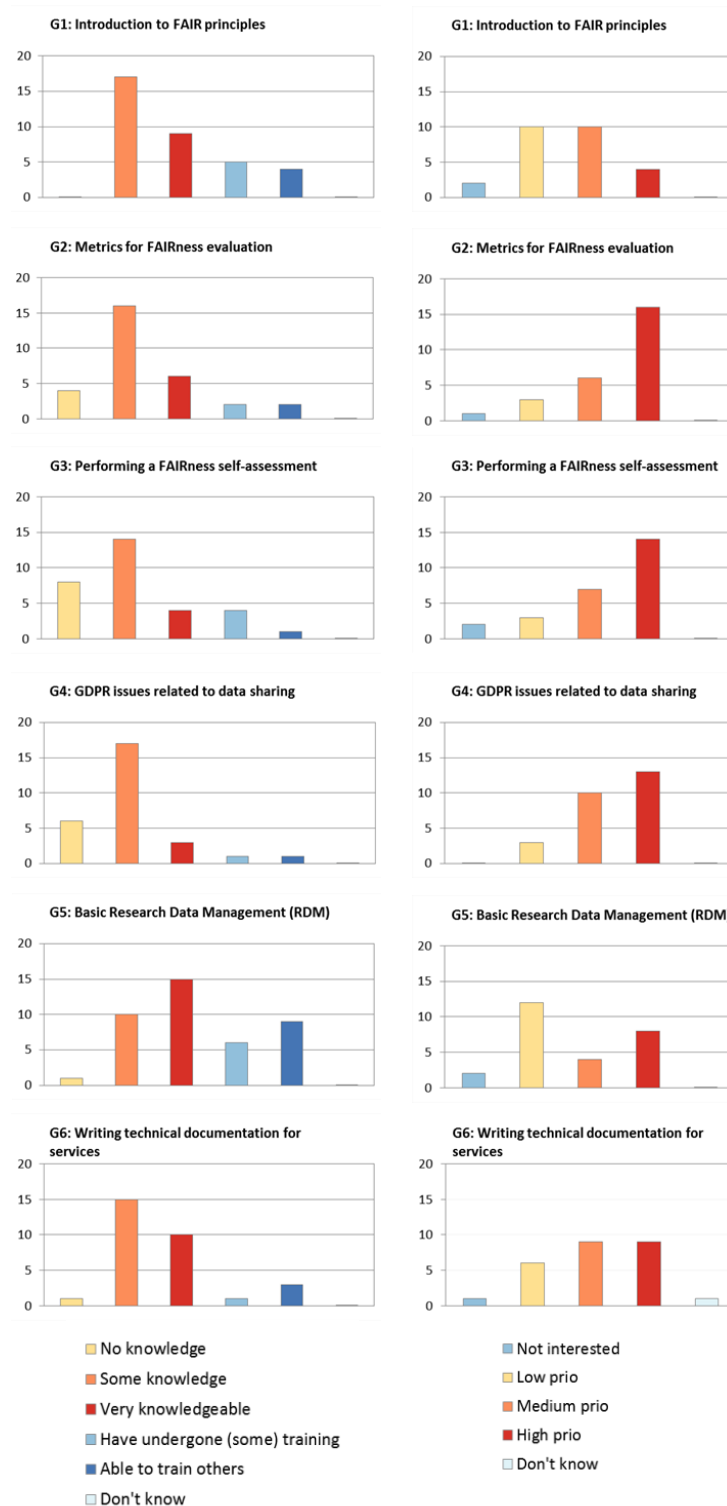


Figure C-3: Histograms summarizing all respondents' estimates of (left) the current knowledge level and training experience, and (right) priorities for training during the ENVRI-FAIR project period. See the respective legend for explanations of the colour schemes.

RDM-specific topics: knowledge status & priorities

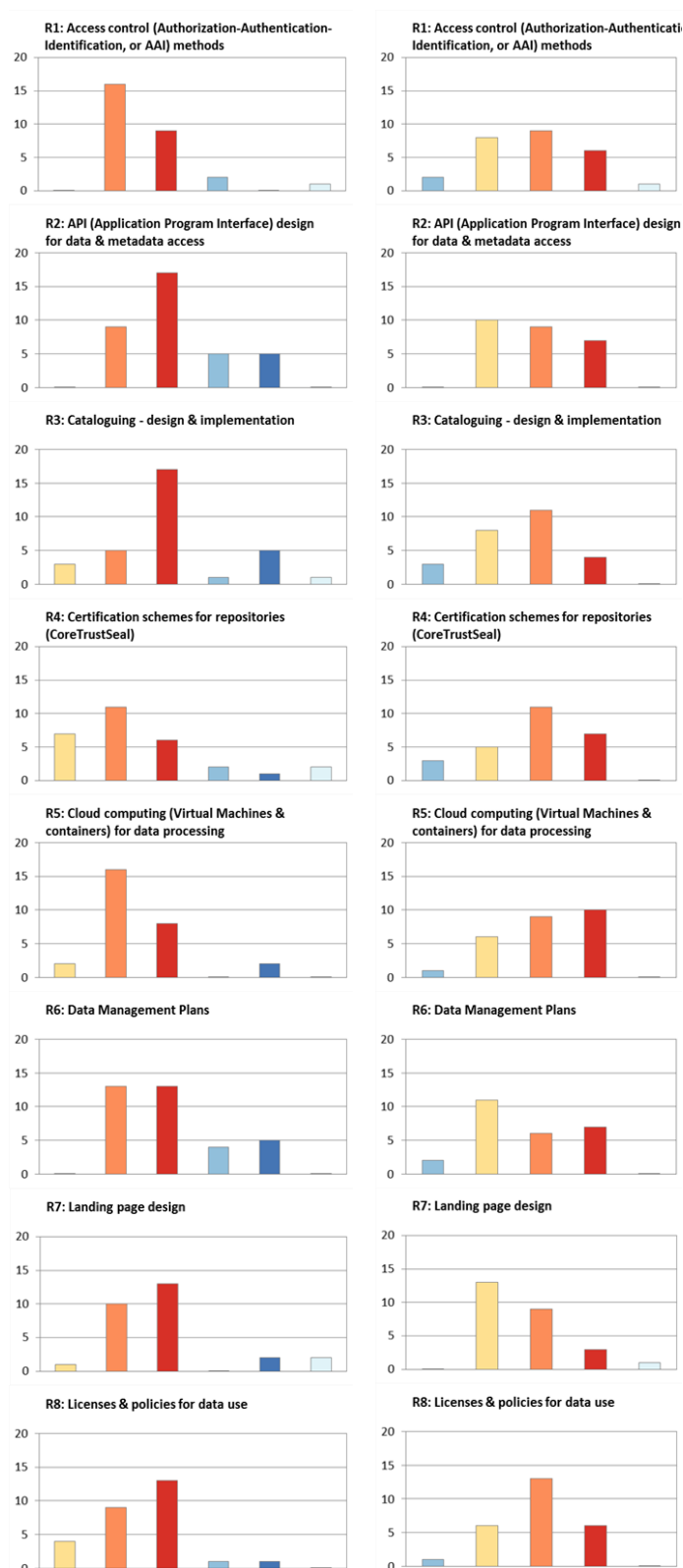


Figure C-4: Research Data Management-related topics R1-R8, all responses. Left: perceived current level of knowledge and training experience. Right: estimated priority of training activities. For legend information, see Figure C-3.

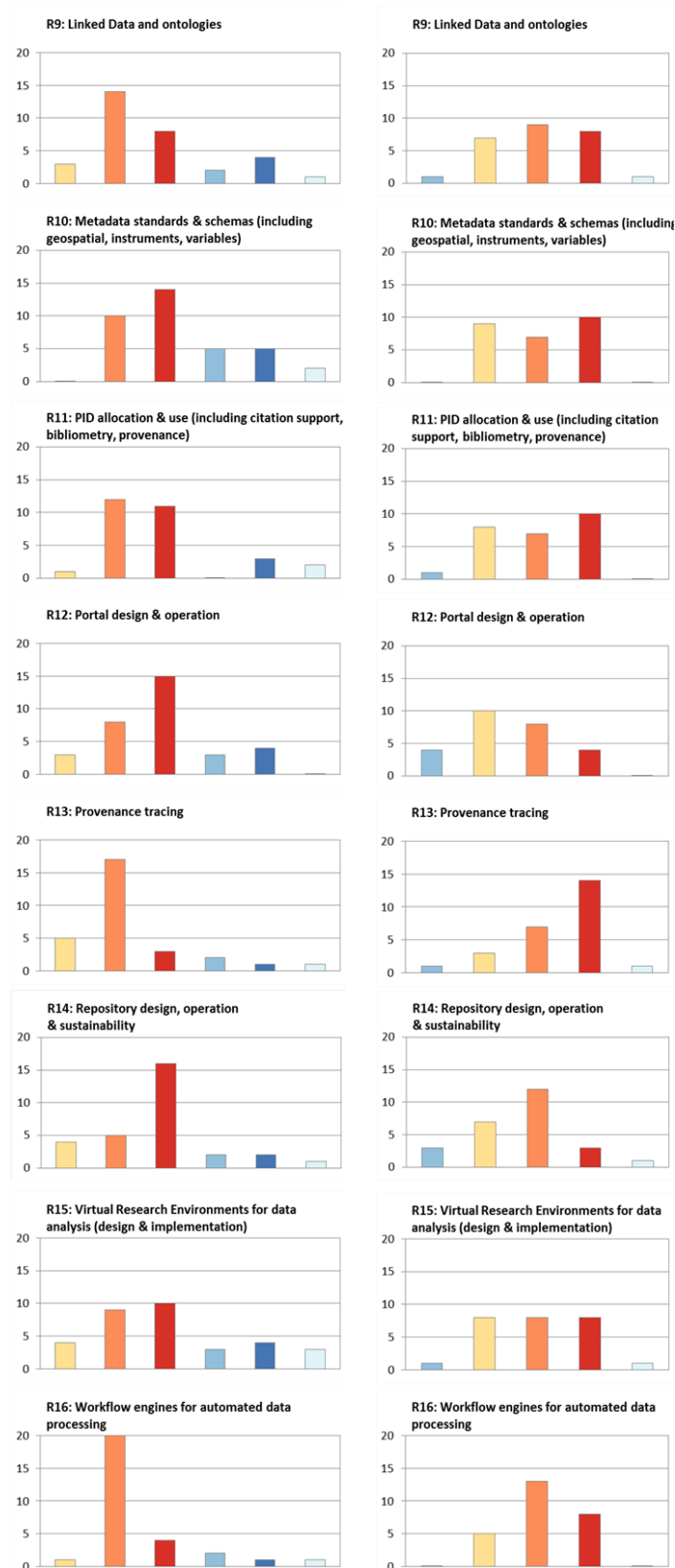


Figure C-5: Same as Figure C-4, but for Research Data Management-related topics R9-R16.

Availability of training materials relevant to sub-domain RIs

Atmosphere

	G1	G2	G3	G4	G5	G6	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16
	0.17	0.66	0.66	0.83	0.17	0.66	0.50	0.50	0.66	0.50	0.50	0.17	0.33	0.50	0.83	0.66	0.50	0.33	1.00	0.33	0.50	0.66
T1	X				X																	
T2				X	X		X							X					X			
T3	X			X	X										X							X
T4	X				X			X	X		X				X	X						
T5					X	X																
T6	X				X							X				X						
T7	X		X		X																	
T8	X				X							X										
T9					X							X				X						
T10					X																	
T11	X				X																	
T12	X				X																	
T13	X				X																	
T14	X		X		X		X	X	X					X	X	X	X					
T15	X				X																	
T16	X			X	X				X		X			X	X	X						
T17	X		X																			
T18	X	X	X					X						X	X	X			X			X
T19	X	X	X	X	X	X																
T20	X																					
T21	X																					
T22	X				X																	
T23	X				X																	
T24	X				X							X		X			X					
T25	X				X							X		X			X					
T26	X				X							X		X			X					
T27	X				X																	
T28	X				X																	
T29	X				X			X									X					X
T30					X										X				X			
T31					X																	
T32	X	X	X		X			X	X			X		X	X	X	X	X	X			X
T33					X			X							X	X	X		X			
T34					X						X										X	X

Figure C-6: Availability of training resources for general FAIR (G1-G6) and RDM-specific (R1-R16) topics with medium or high priority for Atmosphere sub-domain RIs. Yellow topic code cells indicate a difference in priorities vis-à-vis the All domains analysis. White crosses in dark blue and dark red cells identify relevant training resources. Light red and light blue columns indicate the topics with medium and high priority identified in the analysis of the complete set of responses. Row 2 contains the Atmosphere sub-domain priority scores for each topic.

Marine

	G1	G2	G3	G4	G5	G6	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	
	0.62	1.00	1.00	0.87	0.50	0.75	0.50	0.75	0.87	1.00	1.00	0.75	0.62	0.62	0.62	0.87	0.75	0.50	0.87	0.75	0.75	0.87	
T1	X				X																		
T2				X	X		X							X					X				
T3	X			X	X										X								X
T4	X				X				X	X		X			X	X							
T5					X	X																	
T6	X				X							X				X							
T7	X		X		X																		
T8	X				X							X											
T9					X							X				X							
T10					X																		
T11	X				X																		
T12	X				X																		
T13	X				X																		
T14	X		X		X		X		X	X				X	X	X	X						
T15	X				X																		
T16	X			X	X					X		X		X	X	X							
T17	X		X																				
T18	X	X	X						X					X	X	X			X				X
T19	X	X	X	X	X	X																	
T20	X																						
T21	X																						
T22	X				X																		
T23	X				X																		
T24	X				X							X		X				X					
T25	X				X							X		X				X					
T26	X				X							X		X				X					
T27	X				X																		
T28	X				X																		
T29	X				X			X								X						X	
T30					X										X				X				
T31					X																		
T32	X	X	X		X			X	X			X		X	X	X	X	X	X				X
T33					X				X						X	X	X		X				
T34					X						X											X	X

Figure C-7: Same as Figure C-6, but for Marine sub-domain RIs.

Solid Earth

	G1	G2	G3	G4	G5	G6	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	
	0.54	0.85	0.81	0.88	0.46	0.72	0.60	0.62	0.58	0.69	0.73	0.50	0.48	0.73	0.68	0.65	0.65	0.46	0.84	0.60	0.64	0.81	
T1	X				X																		
T2				X	X		X							X					X				
T3	X			X	X										X								X
T4	X				X				X	X		X			X	X							
T5					X	X																	
T6	X				X							X				X							
T7	X		X		X																		
T8	X				X							X											
T9					X							X				X							
T10					X																		
T11	X				X																		
T12	X				X																		
T13	X				X																		
T14	X		X		X		X		X	X				X	X	X	X						
T15	X				X																		
T16	X			X	X				X		X			X	X	X							
T17	X		X																				
T18	X	X	X						X					X	X	X			X				X
T19	X	X	X	X	X	X																	
T20	X																						
T21	X																						
T22	X				X																		
T23	X				X																		
T24	X				X							X		X			X						
T25	X				X							X		X			X						
T26	X				X							X		X			X						
T27	X				X																		
T28	X				X																		
T29	X				X			X								X						X	
T30					X										X				X				
T31					X																		
T32	X	X	X		X			X	X			X		X	X	X	X	X	X	X			X
T33					X				X						X	X	X		X				
T34					X						X										X		X

Figure C-8: Same as Figure C-6, but for Solid Earth sub-domain RIs.

Ecosystem & Biodiversity

	G1	G2	G3	G4	G5	G6	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	
	0.64	0.82	0.73	0.91	0.64	0.73	0.82	0.64	0.64	0.73	0.64	0.82	0.18	0.73	0.73	0.64	0.64	0.45	0.73	0.45	0.54	0.73	
T1	X				X																		
T2				X	X		X							X					X				
T3	X			X	X										X								X
T4	X				X				X	X		X			X	X							
T5					X	X																	
T6	X				X							X				X							
T7	X		X		X																		
T8	X				X							X											
T9					X							X				X							
T10					X																		
T11	X				X																		
T12	X				X																		
T13	X				X																		
T14	X		X		X		X		X	X				X	X	X	X						
T15	X				X																		
T16	X			X	X					X		X		X	X	X							
T17	X		X																				
T18	X	X	X						X					X	X	X			X				X
T19	X	X	X	X	X	X																	
T20	X																						
T21	X																						
T22	X				X																		
T23	X				X																		
T24	X				X							X		X				X					
T25	X				X							X		X				X					
T26	X				X							X		X				X					
T27	X				X																		
T28	X				X																		
T29	X				X			X								X						X	
T30					X										X				X				
T31					X																		
T32	X	X	X		X			X	X			X		X	X	X	X	X	X	X			X
T33					X				X						X	X	X		X				
T34					X						X											X	X

Figure C-9: Same as Figure C-6, but for Ecosystem & Biodiversity sub-domain RIs.