nilu

NILU report 7/2023

The FAIRness of ACTRIS Data Centre

The FAIRness of ACTRIS Data Centre

Authors:

Cathrine Lund Myhre[1)], Markus Fiebig[1)], Richard O. Rud[1)], Lucia Mona[2)], Claudio Dema[2)], Nicolas Pascal[3)], Patrice Henry[4)], Bénédicte Picquet - Varrault[5)], Guillaume Brissebrat[6)], Cathy Boonne[7)], Ewan O'Connor[8)], Simo Tukiainen[8)]

Cooperating institutions:

1) NILU, Norway
2) CNR-IMAA, Italy
3) AERIS/ICARE, France
4) CNES/AERIS, France
5) IPSL/LISA, France
6) CNRS, IPSL/AERIS, France
7) CNRS, IPSL/AERIS/ESPRI, France
8) FMI, Finland

NILU report 7/2023	ISBN: 978-82-425-3117-9 ISSN: 2464-3327	CLASSIFICATION: A – Unclassified (open report)
DATE 19.02.2024	SIGNATURE OF RESPONSIBLE PERSON Aasmund Fahre Vik, Acting CEO (sign.)	NUMBER OF PAGES 34

TITLE The FAIRness of ACTRIS Data Centre	PROJECT LEADER: Cathrine Lund Myhre NILU PROJECT NO. U-119001, E-122001
AUTHOR(S) Cathrine Lund Myhre[1)], Markus Fiebig[1)], Richard O. Rud[1)], Lucia Mona[2)], Claudio Dema[2)], Nicolas Pascal[3)], Patrice Henry[4)], Bénédicte Picquet - Varrault[5)], Guillaume Brissebrat[6)], Cathy Boonne[7)], Ewan O'Connor[8)], Simo Tukiainen[8)] [1)] NILU, Norway, [2)] CNR-IMAA, Italy, [3)] AERIS/ICARE, France, [4)] CNES/AERIS, France, [5)] IPSL/LISA, France [6)] CNRS, IPSL/AERIS, France, [7)] CNRS, IPSL/AERIS/ESPRI, France, [8)] FMI, Finland	QUALITY CONTROLLER Shridhar Jawak
REPORT PREPARED FOR European Commission, Directorate-General Research & Innovation; and The project coordinator: Forschungszentrum Julich GMBH (FZJ), Wilhelm Johnen Strasse, Julich 52428, Germany	CONTRACT REF. Grant Agreement No. 824068 – ENVRI-FAIR H2020-INFRAEOSC-2018-2020/H2020-INFRAEOSC-2018-2 Research Council of Norway Norwegian Grant Agreement No #32224.

ABSTRACT

The purpose of this report is to document the status and implementation of FAIRness within ACTRIS Data centre as of March 2023, developed over the period January 2019 – March 2023.

The report is an extended version of ENVRI-FAIR deliverable D8.4 due March 2023 and available through Zenodo: ENVRI-FAIR D8.4: The FAIRness of ACTRIS | Zenodo, only including the work until autumn 2022. This present report adds more information to the implementation of the FAIR principles by ACTRIS Data Centre over the period January 2019 – March 2023. In addition to D8.4, the present report provides a comprehensive external FAIRness assessment covering the entire period 2019 - 2023, along with an evaluation of the implementation in the years 2022 and the first half of 2023. It's important to note that the project deliverable only encompasses the period from 2019 to 2021.

NORWEGIAN TITLE

FAIRness av ACTRIS Data senter

KEYWORDS

FAIR data Data centre

ABSTRACT (in Norwegian)

Formålet med denne rapporten er å dokumentere status for implementeringen av FAIRness i ACTRIS Data senter frem til juni 2023, over perioden januar 2019 – mars 2023.

Rapporten er en utvidet versjon av ENVRI-FAIR D8.4 som ble levert i mars 2023 og er tilgjengelig via Zenodo: ENVRI-FAIR D8.4: The FAIRness of ACTRIS | Zenodo Den rapporten inkluderer kun arbeidet frem til høsten 2022. Denne rapporten som publiseres her legger til mer informasjon om implementeringen av FAIR-prinsippene av ACTRIS data senter i perioden januar 2019 – juni 2023. Sammenlignet med D8.4 dekker denne rapporten også fullstendig ekstern FAIRness-vurdering av implementeringen siden år 2019 inkludert også implementeringen i år 2022. Prosjektleveransen i Zenodo inkluderer kun året 2019-2021.

PUBLICATION TYPE: Digital document (pdf)	COVER PICTURE: Source: NILU

Citation: Myhre, C. L., Fiebig, M., Rud, R.O., Mona, L., Dema, C., Pascal, N., Henry, P., Picquet-Varrault, B., Brissebrat, G., Boonne, C., O'Connor, E., Tukiainen, S. (2024). The FAIRness of ACTRIS. (NILU rapport 7/2023). Kjeller: NILU.

NILU's ISO Certifications: NS-EN ISO 9001 and NS-EN ISO 14001. NILU's Accreditation: NS-EN ISO/IEC 17025.

Preface

The implementation of FAIRness within ACTRIS DC has benefitted largely from the project ENVRI-FAIR and funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824068. In addition the work at NILU, we also gain advantages from the ACTRIS-Norway project, which is funded by the Research Council of Norway under grant number 32224.

The overarching goal of ENVRI-FAIR was to advance the findability, accessibility, interoperability, and reusability (FAIRness) of the data and services offered by the ENVRI Cluster research infrastructures and to connect them to the emerging European Open Science Cloud. Figure 1 shows the ENVRI landscape and natural compartments with all the RIs within the environmental domain indicated at the bottom. The lower panel shows the atmosphere RIs in more detail, taken from Petzold et al (2023)[1].

[1] Petzold, A., Bundke, U., Hienola, A., Laj, P., Lund Myhre, C., Vermeulen, A., Adamaki, A., Kutsch, W., Thouret, V., Boulanger, D., Fiebig, M., Stocker, M., Zhao, Z., and Asmi, A.: Opinion: New directions in atmospheric research offered by research infrastructures combined with open and data-intensive science, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-1423, 2023.

Figure 1: The landscape of environmental research infrastructures in Europe (ENVRI), overall in the upper panel, and in atmospheric sciences in a nutshell in the lower panel, taken from Petzold et al, 2024[1].

Contents

Table of Contents

Summary

The ACTRIS Data Centre has made substantial progress towards FAIRness, both within the data centre and with respect to common solutions with other environmental research infrastructures (RIs). This report documents the solutions and achievements from the last years, and outlines the way forward for ACTRIS.

The purpose of this report is to document the status and implementation of FAIRness within the ACTRIS Data centre as of March 2023, developed over the period between January 2019 and March 2023.

The report is an extended version of ENVRI-FAIR deliverable D8.4 due March 2023 and available through Zenodo: ENVRI-FAIR D8.4: The FAIRness of ACTRIS | Zenodo. The report accessible through Zenodo includes only the work until autumn 2022. At that stage, many tasks were in final implementation stage, and a full assessment of the FAIRness of the data centre was under development. This present report presents updated information on the implementation of the FAIR principles by ACTRIS Data Centre over the period January 2019 – March 2023. Compared to "*ENVRI-FAIR D8.4:* The FAIRness of ACTRIS" on Zenodo, the present report also covers a full external FAIRness assessment of the implementation over the year 2019 including also the implementation in year 2022 and first half of 2023. The project deliverable D8.4 includes only the period 2019-2021.

1 Context and introduction to the ENVRI community

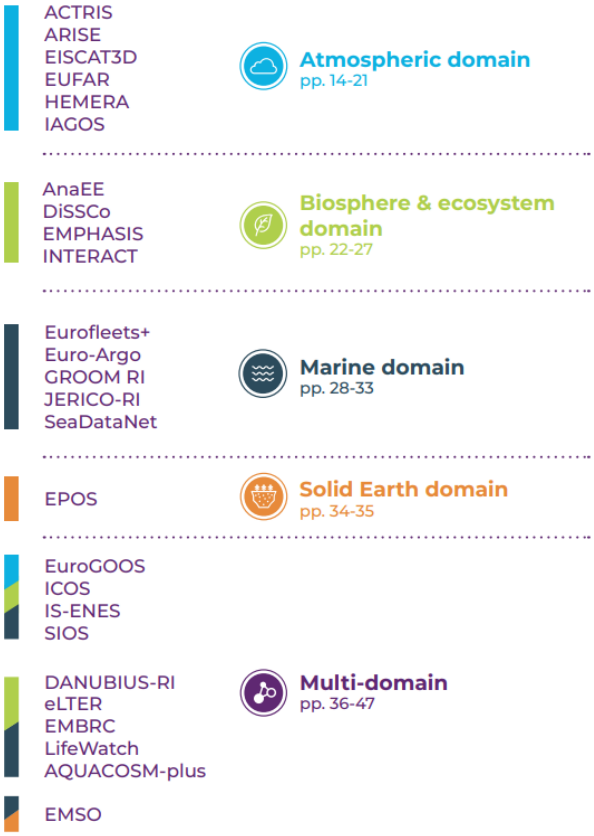


Figure 2: Participating research infrastructures (RI) of the environmental domain cover the subdomains Atmosphere, Marine, Solid Earth and Biodiversity / Ecosystems and thus the Earth system in its full complexity, and associated RIs. Taken from https://envri.eu/home-envri-fair/

ENVIR- FAIR is the connection of the ESFRI Cluster of Environmental Research Infrastructures (ENVRI) to the European Open Science Cloud (EOSC). ENVRI is a community of environmental research infrastructures working together to observe the Earth as one system. ENVRI collaborate to provide multidisciplinary environmental in situ data, tools, and harmonised services that can be used by anyone for free (see also ENVRI)

The Horizon 2020 project ENVRI-FAIR has been crucial for the implementation of FAIRness in the environmental domain. Participating RIs of the environmental domain cover the subdomains Atmosphere, Marine, Solid Earth and Biodiversity / Ecosystems and thus the Earth system in its full complexity, see Figure 2 for sub-domains and associated RIs. NILU led the work with the implementation in the atmospheric sub-domain, and also the progress of FAIRness within ACTRIS DC.

2 ACTRIS Data Centre – a brief introduction

ACTRIS DC is a distributed data centre. The utilisation of numerous measurement methodologies and instruments within ACTRIS leads to a considerable diversity in the collected data . To encompass these requirements, ACTRIS DC is organized in six Units, with clear links and procedures for interaction between the data centre Units, National Facilities (NFs) and topical centres (TCs). There are five units with complementary topic expertise and one unit with integrating activities (DVAS) and the main entrance for users. However, the development of DVAS and its FAIRness relies on the contributing DC units. Accordingly, the development in all units is included here. In this report the implementation and progress within each unit is presented.

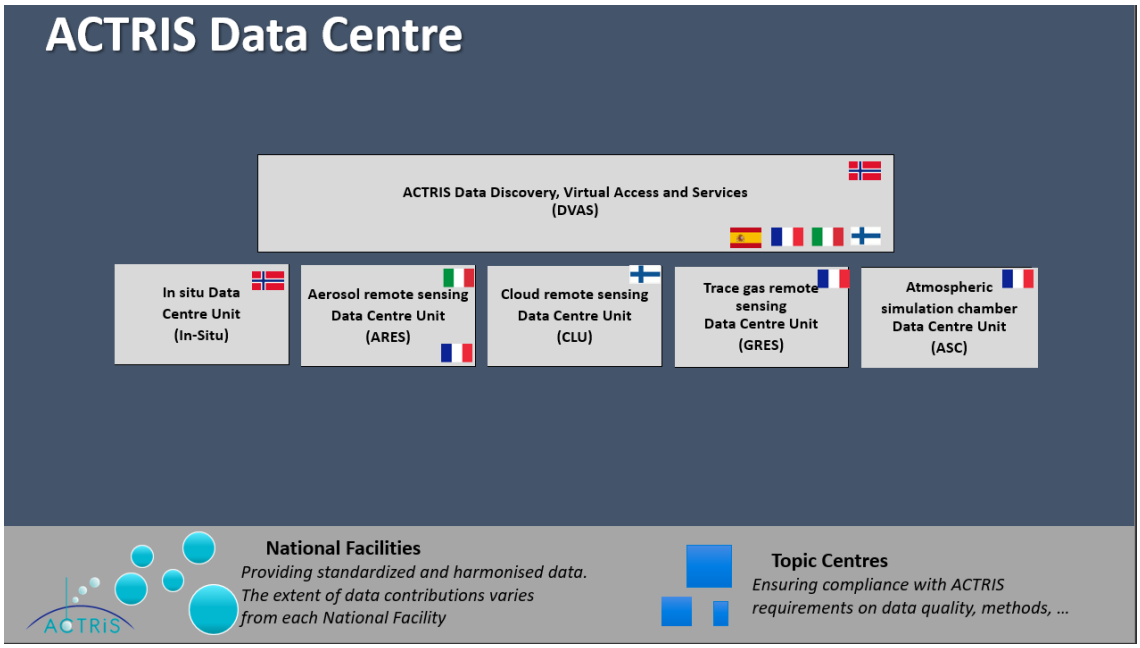


Figure 3: The architecture of the ACTRIS Data Centre involves units distributed across Norway, France, Italy and Finland, with close collaboration among all units and their integration in DVAS.

More information on the ACTRIS Data Centre is available from the ACTRIS web; and in the data management plan.

3 Background and the starting point for ACTRIS data centre joint implementation of FAIRness

3.1 Gap analysis

The RIs of the ENVRI atmospheric subdomain conducted a comprehensive analysis of the FAIRness of their data management, in several steps.

The first step was an analysis of the FAIRness of the data centre, data curation and management in 2018, in relation to the proposal writing phase of ENVRI-FAIR. This analysis resulted in a FAIRness matrix, guiding the tasks and development of the ENVRI-FAIR project.

Follwing the start of the project in January 2019, a comprehensive FAIRness assessment was performed both within the atmospheric sub-domain and across the full environmental domain.

The most essential identified gaps across the atmospheric sub-domain, considered most important from a user perspective, concern the following FAIRness functions:

- Findable: globally unique identifier; indexed in searchable resource.
- Accessible: (meta)data retrievable by standardised protocol
- Interoperable: common vocabulary
- Re-usable: established license, documented provenance, (meta)data meets community standards.

3.2 The FAIR implementation plan for ACTRIS

Based on the gap analysis both within the atmospheric RIs, atmospheric sub-domain and the full environmental domain, a detailed implementation plan was developed for ACTRIS.

The tasks for implementing these functions were grouped by the following criteria:

1) Importance of the functions for the user
2) Maturity of the functions in the RI.

Furthermore, the improvement of FAIRness can be on various levels as illustrated in Figure 4. It can be within each of the ACTRIS DC units, across the ACTRIS DC, within the atmospheric domain and within the environmental domain. For joint work, the development of common products, and the harmonized user experience, the convergence of FAIRness can be very positive, in a broader context.

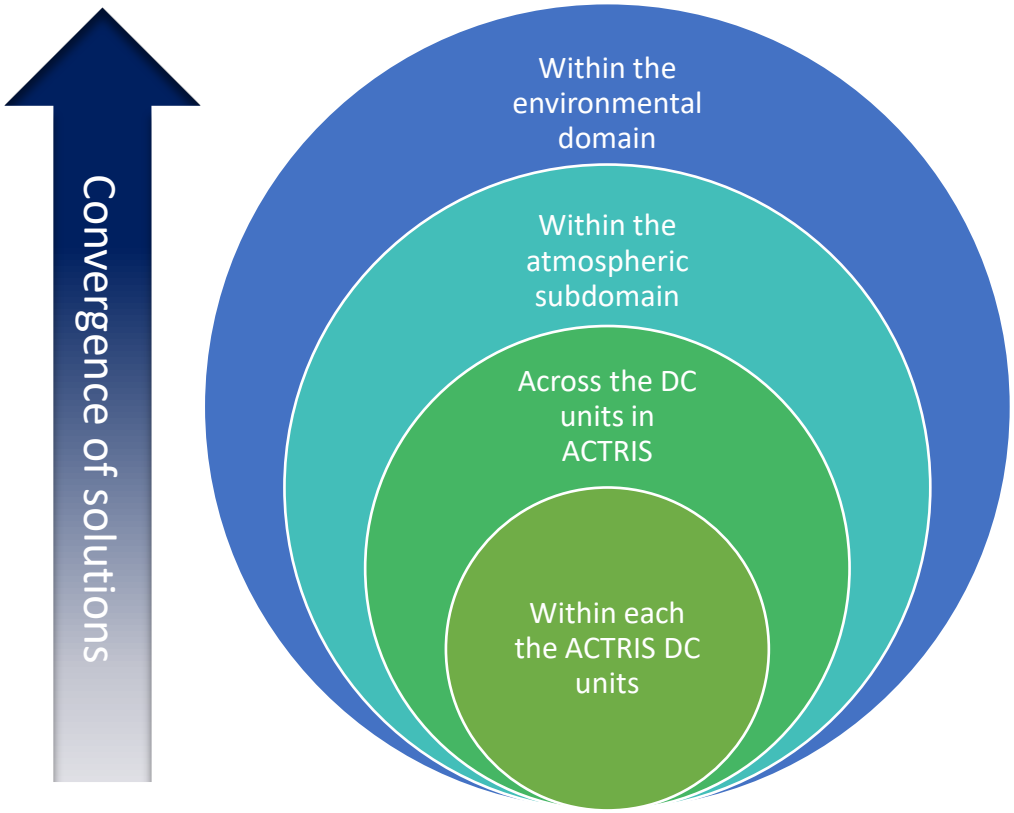


Figure 4: Hierarchy of FAIRness and illustration of the need for convergence of FAIRness from ACTRIS DC units and towards the full environmental domain

Taking these aspects and the gap analysis into account, the 1st detailed implementation plan D8.3 Atmospheric subdomain implementation plan was ready in March 2020. Later, there was a revised version, taking the requirements from the full environmental domain into account for setting up and contributing to the ENVRI-hub for ACTRIS into account. This 2nd version was finalized in June 2021 and is the most recent implementation plan for ACTRIS.

For ACTRIS, the prioritized tasks were grouped into 3 categories, where some of the tasks were targeting harmonization across the atmospheric sub-domain, not limted to ACTRIS.

Immediate implementation: the function implemented is highly important for the user, the implementation plan is consolidated.
Including the following topics:

- Consolidation of consistent use of PIDs throughout data production workflow
- Common standard interfaces for metadata and data access
- Indexing of data resources in WIS, GEOSS
- Common use of authentication schemes
- Endpoint for providing service metadata to ENVRI-hub
- Recommendations on RIs graphical user interfaces (GUI)
- Documenting provenance

Tasks for second half of project: the function implemented is considered important, but implementation can be postponed to allow focused work on the highly important FAIRness functions.

Including the following topics:

- Domain vocabulary / ontology for observed parameters, discovery and use metadata
- Consistent documentation of provenance throughout data production workflow
- Recommendations for licenses on metadata and data
- Semantic search for atmospheric ENVRI RI user interfaces Improve Graphical User Interfaces

Finally, a list of topics that were planned for the second part of the project:

- Common metadata standards and interfaces for use of metadata
- Machine-readable license and attribution metadata.
- Common strategy for structured search interfaces, including common base set of searchable items. Structured search interfaces are preferred by expert users with a high knowledge about the data they are searching.
- Traceable post-production user feedback services
- Data indexing in further data portals. (e.g. Google, EOSC, WIGOS).
- Standards for RESTful APIs for metadata and data.
- Common interfaces for data, facilitating machine readability of data, e.g. in Virtual Research Environment (VRE)s.
- PIDs for organizations and instruments to be determined in consultation with experts in ENVRI-FAIR

4 Progress on FAIRness for ACTRIS

This section includes the progress descriptions of implementation of FAIRness within ACTRIS. We have evaluated the implementation of FAIRness in relation to the implementation plan described in section 3.2. Furthermore, we have used FAIR Implementation Profiles (FIP) analysis from the ENVRI-FAIR cross domain work to assess the internal progress and achievements for the ACTRIS DC, reported in section 4.2.

4.1 The progress of FAIRness over the period 2019 – 2023 within the ACTRIS Data Centre units

A summary of the implementation tasks and status of progress within each unit is shown in this section in a set of tables, starting with the ACTRIS-DVAS. The first column lists the implementation task, then the development of the status of implementation from 2020 until May 2023.

The colours codes mean: Green: implementation ready, yellow; implementation in progress, grey: not relevant or implementation not started, red; implementation is delayed.

Table 1: Summary of implementation of FAIRness within ACTRIS -DVAS. The colours codes mean: Green: implementation ready, yellow; implementation in progress, grey: not relevant or implementation not started, red; implementation is delayed.

Task and Milestones under implementation	Due month Implementation plan	Status June 2020	Status November 2022	Status February 2023
1.1 Use of PIDs throughout workflow				
ACTRIS: determine PID solution	3rd QTR 20			
ACTRIS: implement primary DOIs	1st QTR 22			
ACTRIS: PIDs pre-products implemented	4th QTR 22			
ACTRIS: ORCiD and org. PIDs implemented	4th QTR 22			
1.2 Standard interfaces for (meta)data access				
ACTRIS: metadata repository, 1st version	4th QTR 21			
ACTRIS: metadata M2M interface available	3rd QTR 22			
ACTRIS: THREDDS data repository and interfaces	4th QTR 21			
1.3 Data indexing in WIS and GEOSS				
ACTRIS: link to GISC in WIS implemented	4th QTR 22			
1.4 Common authentication schemes				
ACTRIS: ORCiD authentication implemented	4th QTR 22			
1.5 Service endpoint to ENVRI-hub				
ACTRIS: set up PythonAPI endpoint	3rd QTR 22			
ACTRIS: provide service record	3rd QTR 21			
1.6 GUI recommendations				
ACTRIS: new portal dev. kick-off	2nd QTR 21			
ACTRIS: new portal test release	2nd QTR 22			
ACTRIS: new portal public release	4th QTR 22			
1.7 Document provenance				
ACTRIS: first test version	4th QTR 22			
ACTRIS: final version	4th QTR 22			
Task where implementation plan is needed				
2.1 Domain vocabulary / ontology	3rd QTR 22			
2.2 Documentation of provenance	1st QTR 21			
2.3 Recommendations for licenses	3rd QTR 22			
2.4 Semantic search				

Table 2: Summary of implementation of FAIRness within ACTRIS -In-Situ. The colours codes mean: Green: implementation ready, yellow; implementation in progress, grey: not relevant or implementation not started, red; implementation is delayed.

Task and Milestones under implementation	Due month Implementation plan	Status June 2020	Status November 2022	Status February 2023
1.1 Use of PIDs throughout workflow				
ACTRIS: determine PID solution	3rd QTR 20			
ACTRIS: implement primary DOIs	4th QTR 20			
ACTRIS: PIDs pre-products implemented	2nd QTR 21			
ACTRIS: ORCiD and org. PIDs implemented	4th QTR 21			
1.2 Standard interfaces for (meta)data access				
ACTRIS: metadata repository, 1st version	4th QTR 20			
ACTRIS: metadata M2M interface available	1st QTR 21			
1.3 Data indexing in WIS and GEOSS				
ACTRIS: link to GISC in WIS implemented	3rd QTR 21			
1.4 Common authentication schemes				
ACTRIS: ORCiD authentication implemented	4th QTR 21			
Task where implementation plan is needed				
2.1 Domain vocabulary / ontology	1st QTR 21			
2.2 Documentation of provenance	1st QTR 21			
2.3 Recommendations for licenses	1st QTR 21			
2.4 Semantic search	1st QTR 21			
2.5 Graphical user interfaces	1st QTR 21			

Table 3: Summary of implementation of FAIRness within ACTRIS -ARES. The colours codes mean: Green: implementation ready, yellow; implementation in progress, grey: not relevant or implementation not started, red; implementation is delayed.

Task and Milestones under implementation	Due month Implementation plan	Status June 2020	Status November 2022	Status February 2023
1.1 Use of PIDs throughout workflow				
ACTRIS: determine PID solution	3rd QTR 20			
ACTRIS: implement primary DOIs	4th QTR 20			
ACTRIS: PIDs pre-products implemented	2nd QTR 21			
ACTRIS: ORCiD and org. PIDs implemented	4th QTR 21			
1.2 Standard interfaces for (meta)data access				
ACTRIS: metadata repository, 1st version	4th QTR 20			
ACTRIS: metadata M2M interface available	1st QTR 21			
1.3 Data indexing in WIS and GEOSS				
ACTRIS: link to GISC in WIS implemented	3rd QTR 21			
1.4 Common authentication schemes				
ACTRIS: ORCiD authentication implemented	4th QTR 21			
Task where implementation plan is needed				
2.1 Domain vocabulary / ontology	1st QTR 21			
2.2 Documentation of provenance	1st QTR 21			
2.3 Recommendations for licenses	1st QTR 21			
2.4 Semantic search	1st QTR 21			
2.5 Graphical user interfaces	1st QTR 21			

Table 4: Summary of implementation of FAIRness within ACTRIS -CLU. The colours codes mean: Green: implementation ready, yellow; implementation in progress, grey: not relevant or implementation not started, red; implementation is delayed.

Task and Milestones under implementation	Due month Implementation plan	Status June 2020	Status November 2022	Status February 2023
1.1 Use of PIDs throughout workflow				
ACTRIS: determine PID solution	3rd QTR 20			
ACTRIS: implement primary DOIs	4th QTR 20			
ACTRIS: PIDs pre-products implemented	2nd QTR 21			
ACTRIS: ORCiD and org. PIDs implemented	4th QTR 21			
1.2 Standard interfaces for (meta)data access				
ACTRIS: metadata repository, 1st version	4th QTR 20			
ACTRIS: metadata M2M interface available	1st QTR 21			
1.3 Data indexing in WIS and GEOSS				
ACTRIS: link to GISC in WIS implemented	3rd QTR 21			
1.4 Common authentication schemes				
ACTRIS: ORCiD authentication implemented	4th QTR 21			
Task where implementation plan is needed				
2.1 Domain vocabulary / ontology	1st QTR 21			
2.2 Documentation of provenance	1st QTR 21			
2.3 Recommendations for licenses	1st QTR 21			
2.4 Semantic search	1st QTR 21			
2.5 Graphical user interfaces	1st QTR 21			

Table 5: Summary of implementation of FAIRness within ACTRIS -GRES. The colours codes mean: Green: implementation ready, yellow; implementation in progress, grey: not relevant or implementation not started, red; implementation is delayed.

Task and Milestones under implementation	Due month Implementation plan	Status June 2020	Status November 2022	Status February 2023
1.1 Use of PIDs throughout workflow				
ACTRIS: determine PID solution	3rd QTR 20			
ACTRIS: implement primary DOIs	4th QTR 20			
ACTRIS: PIDs pre-products implemented	2nd QTR 21			
ACTRIS: ORCiD and org. PIDs implemented	4th QTR 21			
1.2 Standard interfaces for (meta)data access				
ACTRIS: metadata repository, 1st version	4th QTR 20			
ACTRIS: metadata M2M interface available	1st QTR 21			
1.3 Data indexing in WIS and GEOSS				
ACTRIS: link to GISC in WIS implemented	3rd QTR 21			
1.4 Common authentication schemes				
ACTRIS: ORCiD authentication implemented	4th QTR 21			
Task where implementation plan is needed				
2.1 Domain vocabulary / ontology	1st QTR 21			
2.2 Documentation of provenance	1st QTR 21			
2.3 Recommendations for licenses	1st QTR 21			
2.4 Semantic search	1st QTR 21			
2.5 Graphical user interfaces	1st QTR 21			

Table 6: Summary of implementation of FAIRness within ACTRIS -ASC. The colours codes mean: Green: implementation ready, yellow; implementation in progress, grey: not relevant or implementation not started, red; implementation is delayed.

Task and Milestones under implementation	Due month Implementation plan	Status June 2020	Status November 2022	Status February 2023
1.1 Use of PIDs throughout workflow				
ACTRIS: determine PID solution	3rd QTR 20			
ACTRIS: implement primary DOIs	4th QTR 20			
ACTRIS: PIDs pre-products implemented	2nd QTR 21			
ACTRIS: ORCiD and org. PIDs implemented	4th QTR 21			
1.2 Standard interfaces for (meta)data access				
ACTRIS: metadata repository, 1st version	4th QTR 20			
ACTRIS: metadata M2M interface available	1st QTR 21			
1.3 Data indexing in WIS and GEOSS				
ACTRIS: link to GISC in WIS implemented	*			
1.4 Common authentication schemes				
ACTRIS: ORCiD authentication implemented	4th QTR 21			
Task where implementation plan is needed				
2.1 Domain vocabulary / ontology	1st QTR 21			
2.2 Documentation of provenance	1st QTR 21			
2.3 Recommendations for licenses	1st QTR 21			
2.4 Semantic search	1st QTR 21			
2.5 Graphical user interfaces	1st QTR 21			

**Note – topic 1.3 it not relevant for the ASC data centre units.*

4.2 The convergence of FAIRness within ACTRIS Data Centre and across atmospheric subdomain

In addition to the RI specific implementation of FAIRness, progress of convergence across the subdomain by implementation of common solutions has been an important goal. Harmonization across the RIs is an important principle to serve broad types of uses and users, and we work to improve access for more easily combination of data, information and facilitate interdisciplinary science. Harmonized solutions across the subdomain facilitate the development of joint products and services utilising data and data products from each RI. Moreover, adopting common solutions will significantly benefit users by reducing the barrier to access data products. It also facilitates linking services and data products to EOSC, as well as other European and global initiatives.

There were four FAIR assessments conducted over the period 2019-2023, which documented the progress of convergence within the ACTRIS DC and the sub-domain.

4.2.1 The FAIR Convergence Matrix in ENVRI-FAIR

A commonly used measure of FAIRness and FAIRness convergence is the FAIR convergence matrix. The FAIR Convergence Matrix is a platform that systematically guides any self-identified community (in this case the ENVRI community) in the decision process leading to optimal FAIR implementations and practices. The resulting collection of FAIR implementation choices composes the FAIR Implementation Profile (FIP) for that community. One important concept for this assessment is the "the single FER (fair enabling resources) of a given subprinciple". Accordingly, for each of the principles **F**indable – **A**ssessable – **I**nteroperable **R**eusable, there are principles as shown in Figure 5, and a defined list of predefined FERs across the project.

Convergence across the atmospheric subdomain and implementation of common solutions are monitored and assessed using FIPs. Within the ENVRI community these profiles are represented by 21 questions asked to the data steward of a community to assess the FAIRness of resources, see Table 1. When a question is answered with an existing resource, or a resource that is in development, it is then considered as a FAIR-enabling resource (FER) which determines if a FAIR principle is met or not. Several FERs can meet the same FAIR principle. All FERs are associated to a status at the RI levels, the shared FER status with the status values:

0 - Resource not declared by community

1 - Resource in development, future use

2 - existing Resource, future use

3 - existing Resource, current use

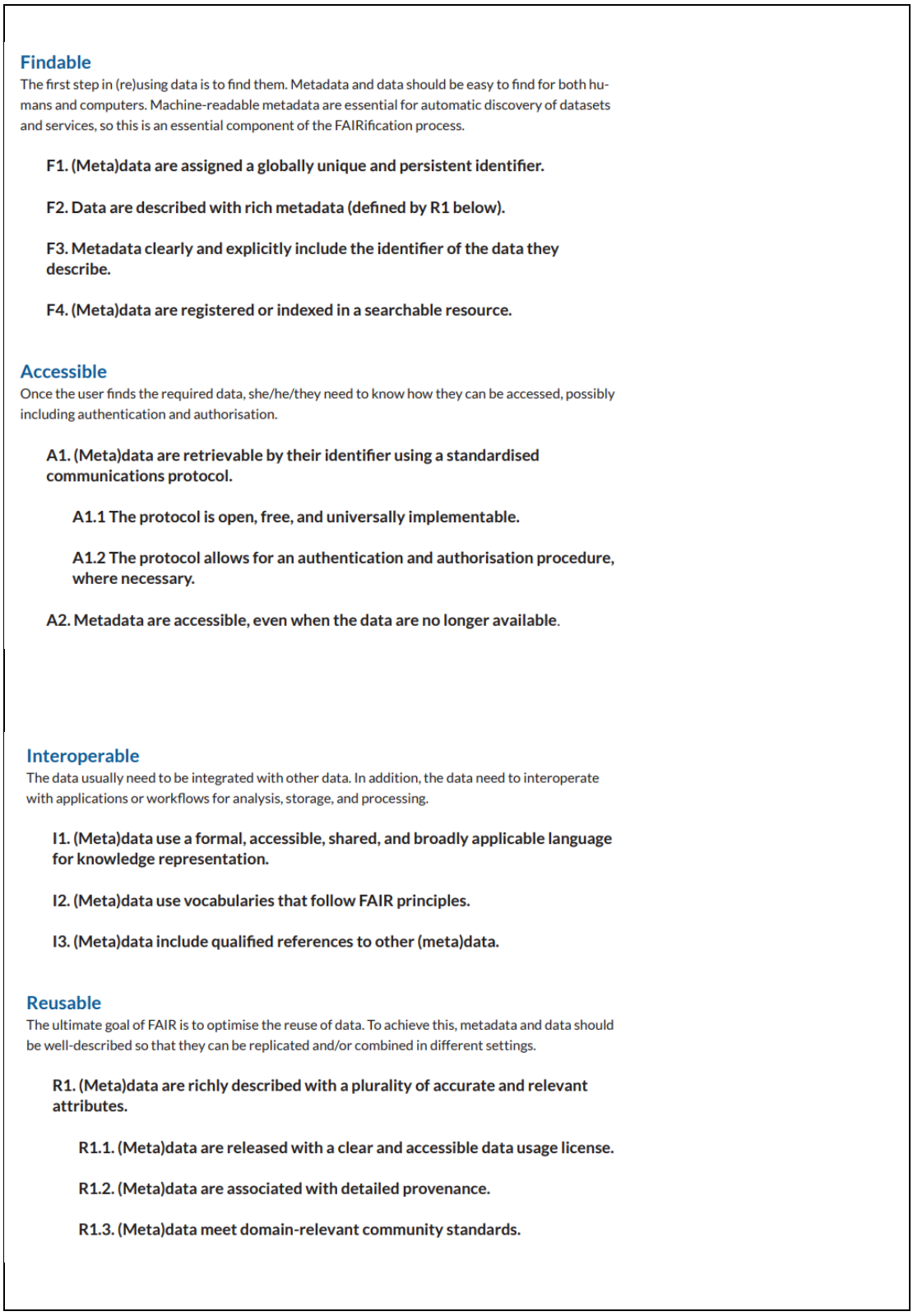


Figure 5: FAIR - findability, accessibility, interoperability, and reusability, and the associated the subprinciples are listed by e.g. GO FAIR: [FAIR Principles - GO FAIR (go-fair.org)](). For a downloadable format of the FAIR Principles, visit the [FAIR-nanopubs page on GitHub]().

Table 7: Questionnaire used to define the FAIR Implementation Profile (FIP) within ENVRI-FAIR. MD = meta data, D = Data

FAIR principle*	Question	FAIR enabling resource types
F1-MD	What globally unique, persistent, resolvable identifiers do you use for metadata records?	Identifier type
F1-D	What globally unique, persistent, resolvable identifiers do you use for datasets?	Identifier type
F2	Which metadata schemes do you use for findability?	Metadata scheme
F3	What is the technology that links the persistent identifiers of your data to the metadata description?	Metadata-Data linking mechanism
F4-MD	In which search engines are your metadata records indexed?	Search engines
F4-D	In which search engines are your datasets indexed?	Search engines
A1.1-MD	Which standardized communication protocol do you use for metadata records?	Communication protocol
A1.1-D	Which standardized communication protocol do you use for datasets?	Communication protocol
A1.2-MD	Which authentication & authorisation technique do you use for metadata records?	Authentication & authorisation technique
A1.2-D	Which authentication & authorisation technique do you use for datasets?	Authentication & authorisation technique
A2	Which metadata longevity plan do you use?	Metadata longevity plan
I1-MD	Which knowledge representation languages (allowing machine interoperation) do you use for metadata records?	Knowledge representation language
I1-D	Which knowledge representation languages (allowing machine interoperation) do you use for datasets?	Knowledge representation language
I2-MD	Which structured vocabularies do you use to annotate your metadata records?	Structured vocabularies
I2-D	Which structured vocabularies do you use to encode your datasets?	Structured vocabularies
I3-MD	Which models, schema(s) do you use for metadata records?	Metadata schema
I3-D	Which models, schema(s) do you use for datasets?	Data schema
R1.1-MD	Which usage license do you use for metadata records?	Data usage license
R1.1-D	Which usage license do you use for datasets?	Data usage license
R1.2-MD	Which metadata schemas do you use for describing the provenance of your metadata records?	Provenance model
R1.2-D	Which metadata schemas do you use for describing the provenance of your datasets?	Provenance model

** See Figure 5*

In order to provide a most complete overview of the FAIRness level, different metrics were used:

- **convergence matrix** on common technologies used in FERs among the different RIs, their state of development (0-3 as defined above)
- **FIP per RI and the total number of FERs completed each year** by each RI documents the progress (section 2"X)

This assessment using FIPs were performed four times over the period 2019 – 2023.

4.2.2 The implemented FAIR-enabling resources (FERs) in ACTRIS Data Centre and atmospheric sub-domain

This section documents the most used FAIR Enabling Resources (FERs) in the atmospheric sub-domain, and the status of the implementation. Table 8 shows the implementation profile per repository at start, and towards end of the project (April 2023). Values are the average implementation status of the single FER of a given subprinciple.

Table 8: Implementation profile per repository at start in 2019, and status April 2023. Values are the average implementation status of the single FER (fair enabling resources) of a given subprinciple (see section 4.2.1 for definitions).

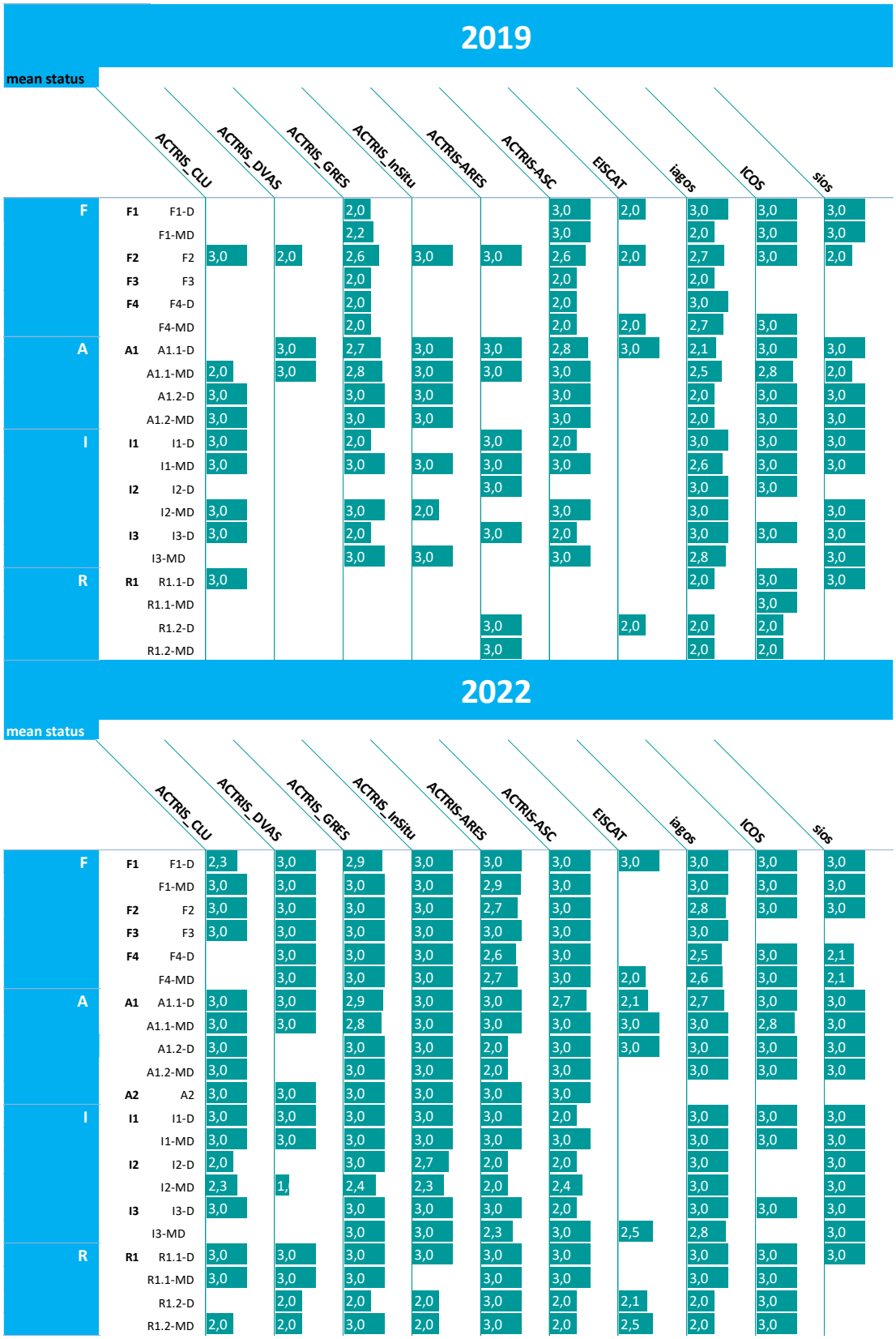


Figure 6 shows the number of FERs any pair of RI data centre has in common. E.g, one vocabulary and one particular CC-license would sum up to a shared FER count = 2. Values are aggregated per principle, year etc. by totalling the number of shared FERs, not taking into account the implementation status. Plots for all years and all FERs are shown. Figure 6a shows the development for the sub principles categorized under **F**indable and **A**ccessible, and Figure 6b shows the development for the sub principles under "**I**nteroperable and **R**eusable.

*Figure 6a: The count of FERs any pair of RIs has in common for **F**indable – upper panel, **A**ccessible in lower panel.*

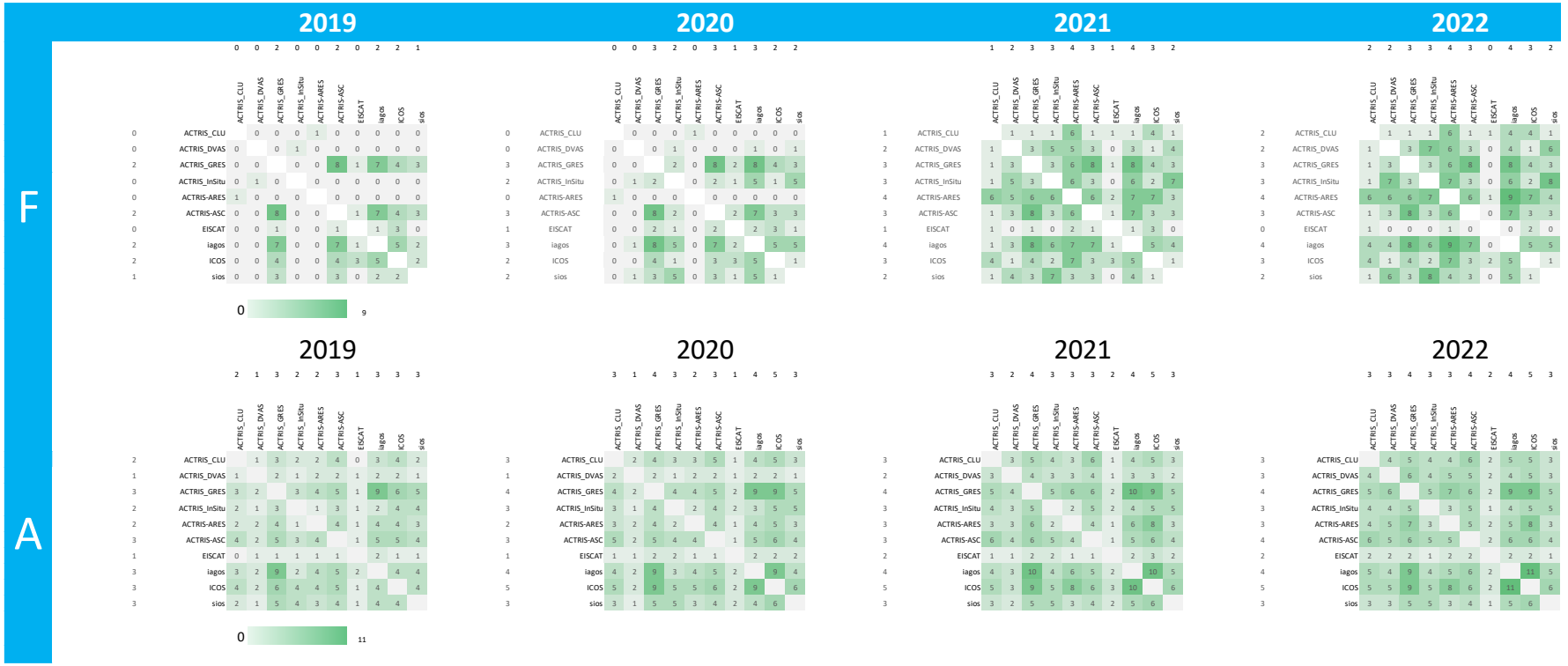


*Figure 6b: The count of FERs any pair of RIs has in common for **I**nteroperable in upper panel and **R**e-usable in the lower panel.*

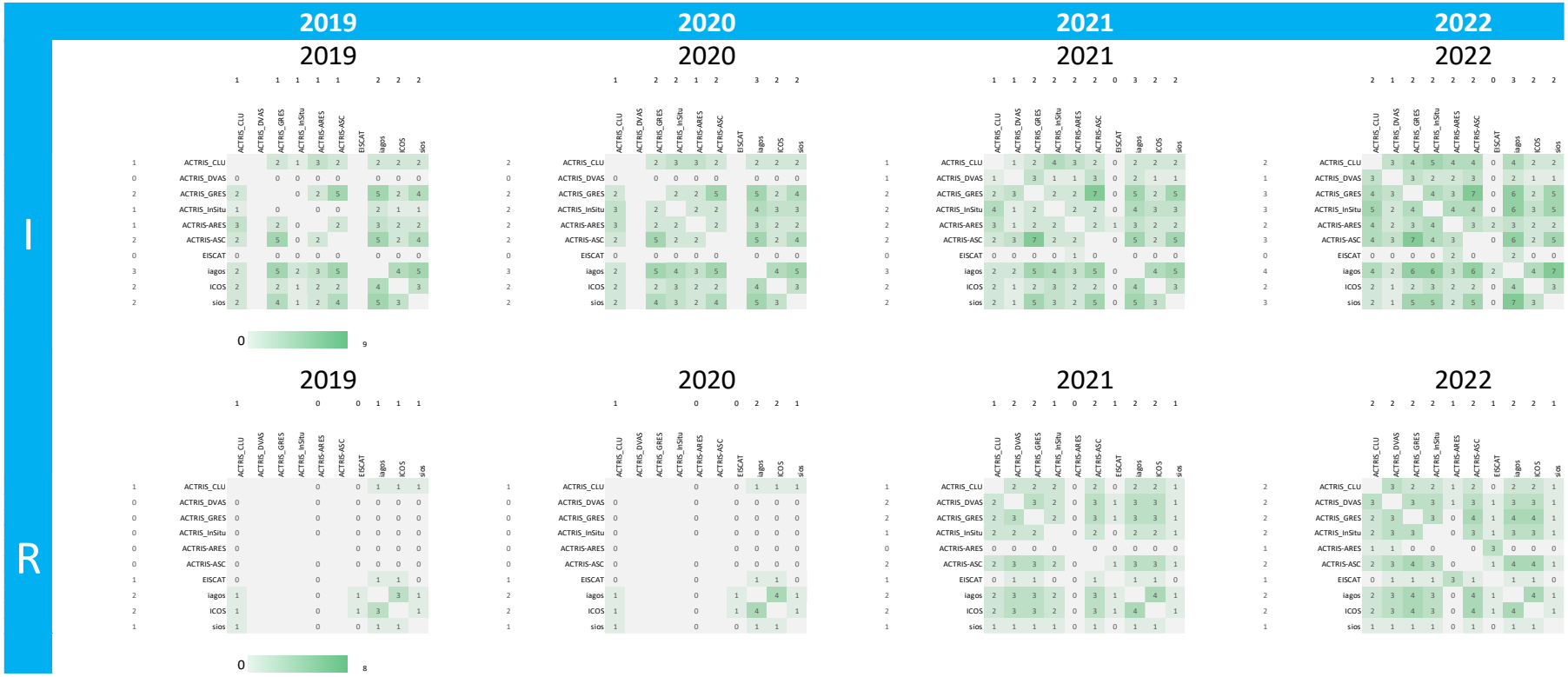
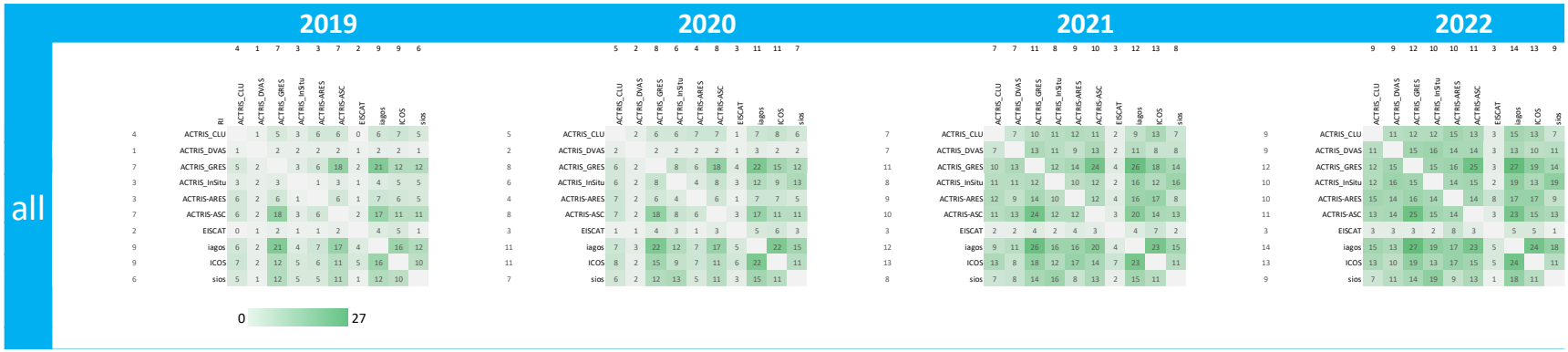


Figure 7: The total count of FERs for any pair of RIs has in common accumulated for all FAIR sub-principles



The strong progress in FAIRness across the sub-domain is evident, in particular for the categories “Findable” and Re-usable”. For these principles, the RIs more or less started from scratch with respect to the FERs defined for the full environmental domain (se Figure 5 for description of principles).

Analysing these plots in relation to the tables in section 4.1, it is possible to check each RI’s implementation tasks over the period, and the largest development, and which implementation steps that have been implemented on the repository level to improve the atmospheric FAIRness and convergence with the other RIS.

4.2.3 Convergence of FAIRness within the atmospheric subdomain

The documentation of the improved convergence across the atmospheric subdomain is included in Figure 8. The given values are average implementation status of the list of predefined FERs used to monitor the implementation of FAIRness. These plots are also called “heat maps”.

The shared FER status with the status values 0 – 3 (0=Resource not declared by community. 1 = Resource in development, future use, 2 = existing Resource, future use. 3 - existing Resource, current use). Figure 4 shows the FAIR convergence from 2019 – 2022.

*Figure 8a: The convergence matrix for the atmospheric subdomain showing the development of joint solutions divided in "**F**indable – upper panel, **A**ccessible in lower panel.*

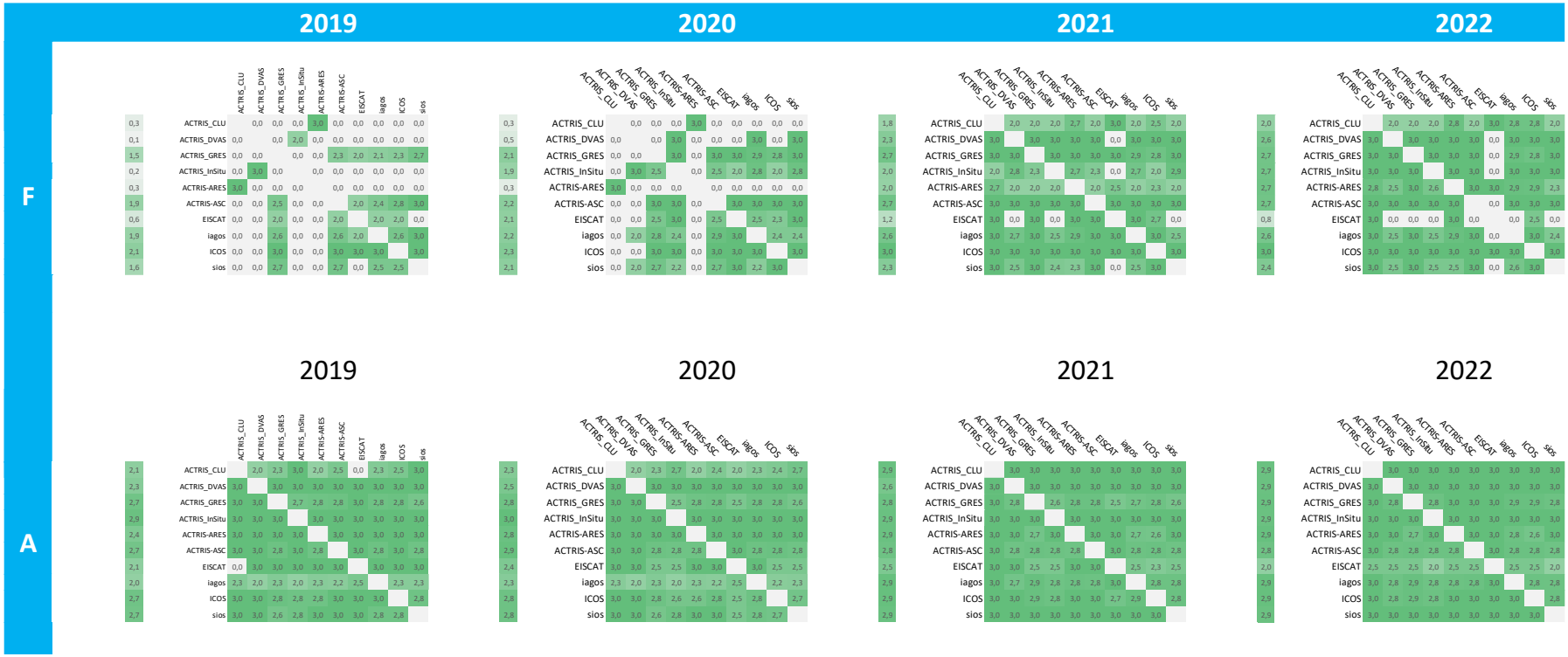


*Figure 8b: The convergence matrix for the atmospheric subdomain showing the development of joint solutions dividend in **I**nteroperable in panel 1 and **R**e-usable in the lower panel.*

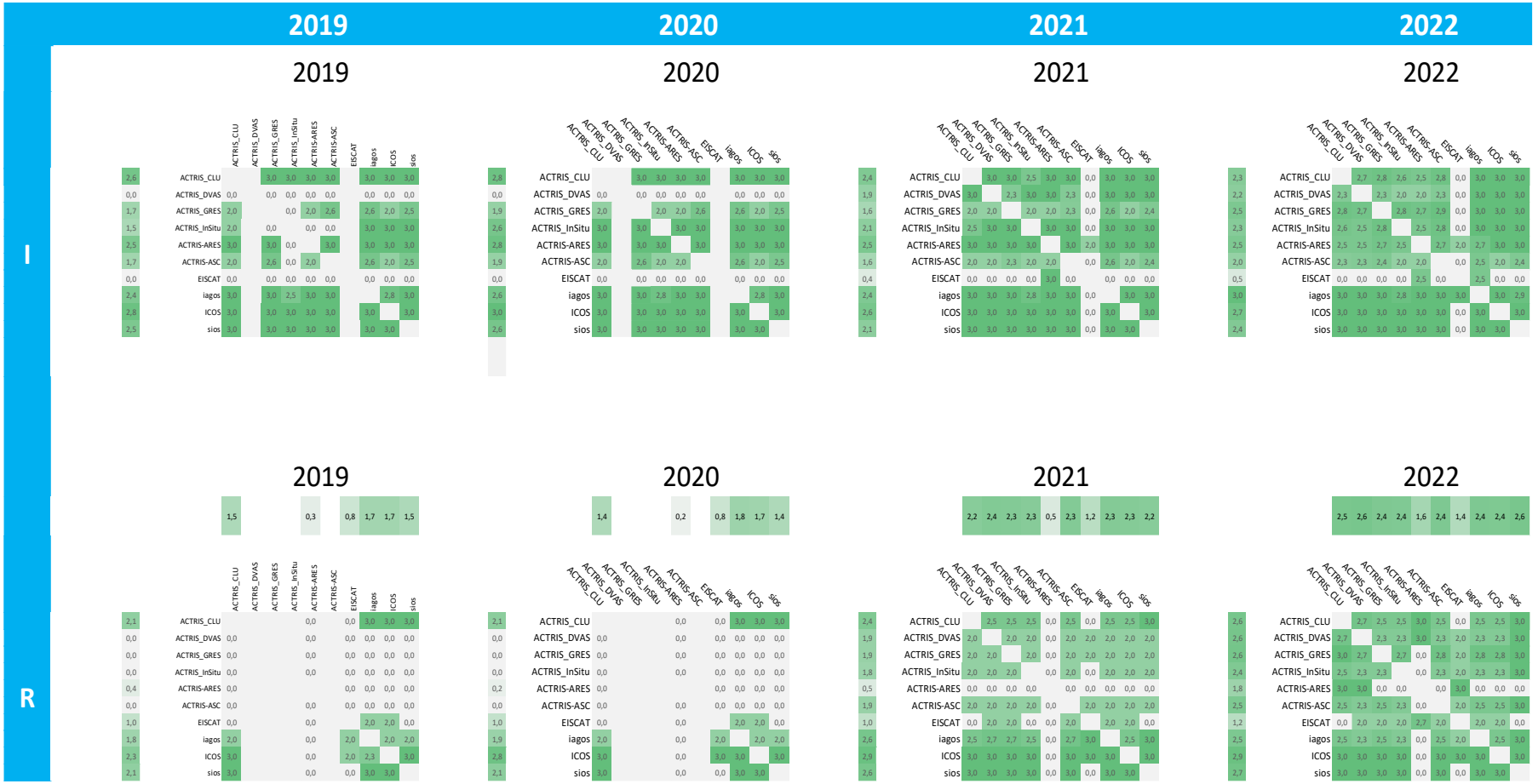


Figure 9: The convergence matrix for the atmospheric subdomain showing the accumulated status for all FAIR sub-principles.

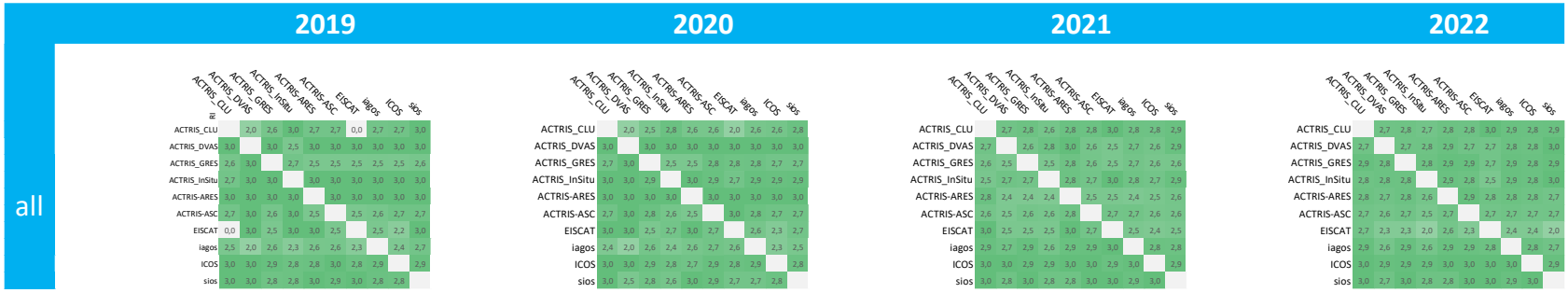


Figure 8a and Figure 8b documents the evolution and the strong progress of convergence within the atmospheric sub-domain over the period. The sub-domain has successfully implemented many harmonised solutions, and as seen in Table 7 on page 21. Table 7 are also finally implemented (also seen in the heat maps).

Figure 8b shows the convergence matrix for the atmospheric subdomain based on averaged numbers and the accumulated status for all FAIR sub-principles, is not very illustrative, as the atmospheric subdomain was mature in some aspects and principles from the start and many RIs had worked on implementation of FAIR tasks before the first FER assessments. Accordingly, the pattern was hidden behind the solutions ready from early of the project. However, looking at Figure 6a and b and Figure 7, the strong evolution and progress of FAIRness are evident and well documented over all years, and all sub-principles.

Describing this in more detail, from the matrix, with the “Accessible” panel 2 in Figure 8a it is evident that all the data centres involved offered relatively good and proper access to data from the beginning of ENVIR-FAIR in 2019. This has been a priority for the RIs at an early stage and the RIs were relatively mature in this regard. However, looking at panel 1 in Figure 6b showing the development of the FERs for “Findable data” there has been a strong development in all RIs, and also Figure 8a panel 1 documents the high convergence across the RI for this principle. This is a very important objective and achievement for the subdomain as for the users' commons solutions in the way of identifying and finding the data is important.

For “Re-Useable” FERs the development and achievements over the period 2019-2022 has been extremely strong. One example of important achievements is the establishment of a community standard for the use of licence on data; in practical terms all are using CCBY 4. This is an important joint solution, facilitating the possibility of developing new products and services across the subdomain.

5 Remaining tasks

At the end of the ENVRI-FAIR project, the ACTRIS DC units have the most essential data FAIRness functionalities in place, even if there is not full convergence. The most essential data FAIRness functionalities comprise:

- identification of data products and pre-products
- machine interfaces to data and metadata repositories, in place.

Remaining gaps in these functionalities as listed in the tables above should be addressed as soon as possible.

6 Main achievements and improvement of FAIRness for ACTRIS over the period 2019 – June 2023

In the course of the ENVRI-FAIR project, the ACTRIS DC has increased its data FAIRness in all parts of the concept:

- **Findability:**
  Most ACTRIS data are already findable and accessible through a common web-portal for human interaction. The next generation portal has been specified during the project and is currently under implementation. A common concept for data identification has been specified which is based on a unified granularity, allowing for consistent qualification of data use. The next generation data portal will be based on standardised, FAIR interoperability functions.

- **Accessibility**:
  All ACTRIS DC units have implemented standardised machine interfaces for metadata and data, or are in the process of finishing this implementation. These interfaces are the basis of the common ACTRIS metadata interface offering machine-access to ACTRIS data for everyone interested, as well as the basis for the next generation data portal for human interaction. The implementation has made ACTRIS data and metadata visible and accessible in e.g. the ENVRI-hub, Data indexing in WIS and GEOSS portal, and visible also in the EOSC marketplace.

- **Interoperability:**
  In an effort involving the whole research infrastructure, ACTRIS has established a common, FAIR compliant vocabulary for describing its data products, along with routines for maintaining the vocabulary. For describing variable names, the vocabulary complies with the RDA recommended I-ADOPT ontology, making it suitable for cross-domain integration.

- **Reusability**:
  ACTRIS has agreed on a concept of licences for both data products, pre-products, and software.

All of the above services, along with implementation plans for future ones, have been documented in a public data management plan and the FAIRness implementation plan is still an important document for further advancement of ACTRIS.

6.1 What can we do now which we could not before advancing of the FAIRness?

This section includes a list of functions, topics and aspects we can do now, compared to January 2019. The work enables user services such as:

- Direct citation of data (e.g. minting DOI on all quality assured data, with harmonised granularity across the data centre)
- Automatic data access into user applications.
- Access to ACTRIS data resources in VREs such as Jupyter notebooks.
- Reliable re-use of data due to license.

Some more details on new achievements are listed here:

- Machine access to the full ACTRIS data repository, in addition to metadata, with granularity corresponding to identification service.
- Improvement of the quality of our (meta)dataparticularly in terms of traceability and provenance across the whole DC even if we still need to complete the implementation (provenance, instrument PIDs are in progress)
- Identify datasets in the repository, both as full time series and as individual data deliveries, connected by documented provenance.
- Controlled vocabulary extending beyond our RI that has enabled better integration with other RIs and sub-domains, and internally within ACTRIS, not only for the data centre but crucial also for the interactions with the topic centres. This includes definition and use of FAIR and I-ADOPT compliant vocabulary.
- Detailed DMP for infrastructure, including all FAIRness aspects down to the level of data production workflows.
- Improve satellite data extraction and colocation through machine-to-machine interoperable services
- Established license for metadata, data and software

6.2 What would we like to do / achieve within the next 5 years?

- Implement new data portal for humans that makes use of all FAIRness achievements so far (vocabulary, easier machine data access, data streaming)
- Improving and enhancing the link between ACTRIS and EOSC, GEO, Copernicus.
- Integrate more ACTRIS resources (services and datasets) in the EOSC marketplace
- Utilize the recent cross RI FAIRness development and continue to develop cross RI services
- Cross-RI searching capability on RI portals (for example possibility on ACTRIS portal to query ICOS data for variables of interest in the same period /close by stations...)
- Implement service for identification of data collections.
- Fully implement documentation of data provenance, and make the information accessible and usable for the data user.
    - improvements on data provenance and traceability by providing for all new datasets (experiments) as a mandatory fields: a control experiment, the auxiliary mechanism (see DMP), the ORCID of the data provider, a flagging system providing information on the quality level of the data, assignment of PID to instruments
- Further improvements on data access: provide L2 data NetCDF format (in addition to the edf format used actually)
- CoreTrustSeal certification of data centre, including functional test of all FAIRness features.
- Make metadata available in RDF compatible format.

- Implement EOSC-compatible single-sign-on mechanism for services requiring authentication.
- Implement quantification of data use feature based on OpenAire API.
- Empowering search synergistic capability on ACTRIS data portal
- Improve the colocation tool and support more satellite data, for atmospheric science but also extended to solid earth, marine and surface observation
- Work more on machine actionability, provide user friendly API and develop on-demand processing services (through OGC WPS for instance)
  => dashboards or VREs (to cover either generic or specific needs), and maybe allow some users to build their own

NILU is an independent, nonprofit research institute established in 1969.

NILU is a highly qualified research institute with a leading competence profile within atmospheric composition, climate change, air quality, environmental contaminants, health effects, sustainable systems, circular economy, and digitalisation.

NILU shall, through research, services, innovation and enabling technologies, contribute to sustainable solutions on Grand

Challenges. NILU shall have leading research infrastructure, including/such as laboratories, observatories and open data centres.

nilu