Milestone 1.1: Landscape analysis of data FAIRness

Authors: Shridhar Jawak, Elisabeth Andrews, Lucia Mona, Ewan O'Connor, Giri Prakash, Ellsworth J. Welton, and Markus Fiebig

Work package no	WP1
Deliverable no.	M1.1
Lead beneficiary	NILU
Deliverable type	X R (Document, report) ☐ DEC (Websites, patent filings, videos, etc.) ☐ OTHER: please specify ....
Dissemination level	X PU (public) ☐ CO (confidential, only for members of the Consortium, incl. Commission)
Estimated delivery date	M8
Actual delivery date	30/10/2024
Version	Final
Reviewed by	Ewan O'Connor
Accepted by	Ewan O'Connor
Comments	

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

Contents

Contents 2
1. Introduction: The Importance of Data FAIRness and FAIR Convergence 3
   1.1. Data FAIRness 3
   1.2. FAIR Convergence 4
   1.3. The Ambition Level in CARGO-ACT 5
   1.4. Method 5
2. The FAIRness of CARGO-ACT Data Centres 6
3. FAIR Convergence Status Between CARGO-ACT Partner Data Centres 9
   3.1. Findability FAIR Convergence 9
   3.2. Accessibility FAIR Convergence 11
   3.3. Interoperability FAIR Convergence 13
   3.4. Reusability FAIR Convergence 15
   3.5. Overall FAIR Convergence 17
4. FAIR Enabling Resources Used by CARGO-ACT Partner Data Centres 19
5. Key Results 21

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

1. Introduction: The Importance of Data FAIRness and FAIR Convergence

Data FAIRness refers to a set of principles and services for making data **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable (FAIR). These requirements were defined by the FORCE11 group at a workshop in 2014, and published in 2016. These principles were established to improve the management and sharing of data, ensuring that it can be easily discovered, accessed, integrated, and reused by both humans and machines. FAIR Convergence involves aligning the technologies used to implement data FAIRness to ensure that different datasets and tools can work together seamlessly. This convergence is crucial for creating an ecosystem where data from diverse sources can be integrated and utilized effectively.

Problems solved by FAIRness and FAIR Convergence:

- Data silos: By making data FAIR, it breaks down barriers between datasets, allowing for more comprehensive and integrated analyses.
- Accessibility issues: Ensures that data is accessible to a wider audience, including researchers, policymakers, and the public.
- Interoperability challenges: Promotes the use of common standards and protocols, making it easier to combine and compare data from different sources.
- Reusability: Enhances the potential for data to be reused in future research, reducing duplication of effort and fostering innovation.

Benefits for data users

- Enhanced discoverability: Easier to find relevant data through improved metadata and indexing.
- Improved collaboration: Facilitates collaboration across disciplines and institutions by providing a common framework for data sharing.
- Increased efficiency: Reduces time and effort needed to prepare and clean data for analysis.
- Greater impact: Data that is FAIR can be used in more ways, increasing its value and impact on research and decision-making.

1.1. Data FAIRness

The FAIR data principles as defined by the FORCE11 group are a set of services that need to be implemented by the data repository hosting the data to be data FAIR compliant:

Findable:

F1. (meta)data are assigned a globally unique and eternally persistent identifier.
F2. data are described with rich metadata.
F3. (meta)data are registered or indexed in a searchable resource.
F4. metadata specify the data identifier.

Accessible:

A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
A1.1 the protocol is open, free, and universally implementable.

A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
A2 metadata are accessible, even when the data are no longer available.

Interoperable:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles.
I3. (meta)data include qualified references to other (meta)data.

Re-usable:

R1. (meta)data have a plurality of accurate and relevant attributes.
R1.1. (meta)data are released with a clear and accessible data usage license.
R1.2. (meta)data are associated with their provenance.
R1.3. (meta)data meet domain-relevant community standards.

The degree of data FAIRness of a data repository can be quantified by how many and which of these services have been implemented.

1.2. FAIR Convergence

Most services required to achieve data FAIRness can be implemented with various technologies, standards, or specifications. All of these technologies, standards, and specifications are, in this context, called FAIR Enabling Resources (FERs). If a community of data repositories intends to offer interoperability services between them, it is important that they use the same or compatible FERs. Otherwise, interoperability will be limited even if all participating data repositories on their own are fully compliant with the data FAIRness requirements. The process of moving towards the same or compatible FERs in a community is called FAIR Convergence.

Importance of FAIR Convergence

- Enhanced interoperability: By aligning FERs, FAIR convergence ensures that data from different domains and sources can be integrated and analysed together. This is particularly important in multidisciplinary research where data from various fields need to be combined.
- Improved data quality: Convergence promotes the use of common standards and best practices, leading to higher data quality and consistency. This makes data more reliable and easier to use.
- Increased efficiency: Researchers and data users spend less time on data cleaning and preparation, allowing them to focus more on analysis and innovation. This efficiency boosts productivity and accelerates scientific discovery.
- Facilitated collaboration: FAIR convergence fosters collaboration across different research communities and institutions by providing a common framework for data sharing and integration. This leads to more comprehensive and impactful research outcomes.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

- Broader impact: When data is interoperable and reusable across different platforms and disciplines, its potential impact is significantly increased. This can lead to new insights, innovations, and solutions to complex global challenges.

In summary, FAIR convergence is essential for maximizing the value of data, enhancing collaboration, and driving innovation across various fields. By ensuring that data can be easily integrated and reused, user can unlock its full potential and make significant use in research and development.

1.3. The Ambition Level in CARGO-ACT

In CARGO-ACT, it is not the ambition to achieve full data FAIRness compliance and full FAIR convergence among the participating data repositories. Instead, CARGO-ACT WP1 intends to assess the degree of FAIRness of these participating data repositories as well as the degree of FAIR convergence, and to identify areas where improvement would be beneficial. In a second WP document, a concept for addressing or implementing these improvements will be presented. The actual implementation will be subject to partner resources, internal priorities and boundary conditions.

1.4. Method

A FAIR Implementation Profile (FIP) is a structured document that describes the specific choices and practices a community adopts to implement the services required by the FAIR principles (Findable, Accessible, Interoperable, and Reusable). Essentially, it captures how a particular community of practice applies these principles to their data and resources.

Key components of a FAIR implementation profile (FIP):

- Technology choices: Lists the tools, standards, and protocols used to achieve FAIRness.
- Community decisions: Reflects collective decisions made by the community regarding data management and sharing.
- Documentation: Provides detailed descriptions of how each FAIR principle is implemented within the community.

Importance of FIPs:

- Standardization: Helps standardize data practices within and across communities, promoting consistency and quality.
- Transparency: Increases transparency in how data is managed and shared, making it easier for others to understand and use the data.
- Reusability: Enhances the reusability of data by providing clear guidelines and documentation on how data can be accessed and integrated.
- Collaboration: Facilitates collaboration by providing a common framework that different communities can adopt and adapt.

FIPs are crucial for ensuring that data is managed in a way that maximizes its value and usability, fostering a more efficient and collaborative research environment.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

The FIP Wizard is a tool designed to capture FAIR Implementation Profiles (FIPs) through a structured questionnaire (https://catalogue.fair-impact.eu/node/212). This helps communities document their technology choices and practices in alignment with the FAIR principles. The FIP Wizard is funded by the FAIR-IMPACT project (fair-impact.eu), which aims to support the implementation of FAIR principles across various domains. The tool is maintained by the ENVRI-FAIR project, which focuses on integrating environmental research infrastructures to ensure they adhere to FAIR principles.

2. The FAIRness of CARGO-ACT Data Centres

This section summarises to what degree the CARGO-ACT data centres have implemented the services required for data FAIRness (see Table 1). In all further analysis, it needs to be noted that all data collected as part of the NOAA Federated Aerosol Network (NFAN) are archived in the Global Atmosphere Watch GAW) World Data Centre for Aerosol (WDCA), which coincides with the ACTRIS In Situ data centre unit. The data of the core NOAA stations, which are part of NFAN, are also archived by the NOAA Global Monitoring Laboratory (GML).

Table 1: Summary of the implementation status of services required to achieve data FAIRness at CARGO-ACT partner data centres. The table distinguishes between completed implementation, implementation in progress, implementation planned, and implementation not in pipeline.

CARGOACT RIs
FAIRness status
- Completed
- In progress
- Not started
- New tasks, not in the implementation plan at that stage

FAIRness service	ACTRIS IN SITU / NFAN	ACTRIS ARES	ACTRIS CLU	ACTRIS GRES	ACTRIS ASC	ARM	MPLNET	NOAA GML (AERO)
Findability								
F1 Metadata: persistent identifier	Completed	In progress	Completed	Completed	Completed	Completed	New task	Completed
F1 Data: persistent identifier	Completed	Completed	In progress	In progress	Completed	Completed	New task	Completed
F2: Metadata schema	Completed	In progress	Completed	Completed	Completed	Completed	New task	Completed
F3: Schema linking data PIDs to metadata	Completed	Completed	Completed	Completed	Completed	Completed	New task	Completed
F4 Metadata: publication service	Completed	In progress	Completed	Completed	Completed	Completed	New task	Completed
F4 Data: publication service	Completed	In progress	Completed	Completed	Completed	Completed	New task	Completed
Accessibility								
A1.1 Metadata: standardized communication protocol	Completed	Completed	Completed	In progress	Completed	Completed	Completed	Completed
A1.1 Data: standardized communication protocol	Completed	Completed	Completed	In progress	In progress	Completed	Completed	Completed
A1.2 Metadata: AAI service	Completed	In progress	Completed	Completed	Completed	Completed	Completed	Completed
A1.2 Data: AAI service	Completed	In progress	Completed	Completed	Completed	Completed	Completed	Completed
A2: Metadata preservation policy	Completed	Completed	In progress	Completed	Completed	Completed	Completed	New task
Interoperability								
I1 Metadata: Knowledge representation language	Completed	Completed	Completed	Completed	Completed	Completed	Completed	Completed
I1 Data: Knowledge representation language	Completed	Completed	Completed	Completed	In progress	Completed	Completed	Completed
I2 Metadata: structured vocabulary	Completed	In progress	Completed	Completed	Completed	Completed	New task	Completed
I2 Data: structured vocabulary	Completed	Completed	Completed	In progress	In progress	Completed	Completed	Completed
I3 Metadata: semantic model	Completed	In progress	Not started	Completed	Completed	Completed	New task	Completed
I3 Data: semantic model	Completed	Completed	Completed	Completed	In progress	Completed	Completed	Completed
Reusability								
R1.1 Metadata: License	In progress	Completed	Completed	Completed	Completed	Completed	New task	Completed
R1.1 Data: License	Completed	Completed	Completed	Completed	Completed	Completed	New task	Completed
R1.2 Metadata: Provenance and provenance schema	In progress	Completed	In progress	Completed	In progress	Completed	New task	New task
R1.2 Data: Provenance and provenance schema	In progress	Completed	New task	In progress	In progress	Completed	New task	New task

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

Findability

- **F1 Metadata: Persistent Identifier**: Most data centres have *Completed* status, except ACTRIS ARES (*In progress*). Not implemented for MPLNET.
- **F1 Data: persistent identifier**: Status varies, with ACTRIS CLU and GRES (*In progress*), indicating room for improvement. Not implemented for MPLNET.
- **F2: Metadata Schema**: Consistent *Completed* status, except ACTRIS ARES (*In progress*). Not implemented for MPLNET.
- **F3: Schema linking data PIDs to metadata**: Uniform *Completed* status, except for MPLNET.
- **F4 Metadata & Data: publication service**: Similar patterns, with ACTRIS ARES (*In progress*). Not implemented for MPLNET.

Accessibility

- **A1.1 metadata & data: Standardized Communication Protocol**: Generally *Completed* across RIs, with ACTRIS ASC and GRES (*In progress*) for data.
- **A1.2 Metadata & Data: AAI Service**: Consistent *Completed* status across RIs, except ACTRIS ARES (*In progress*).
- **A2: Metadata Preservation Policy**: Mostly *Completed*, with ACTRIS CLU In progress and NOAA GML does not have any implementation plan at this stage.

Interoperability

- **I1 Metadata & Data: Knowledge Representation Language**: Status are mostly *Completed*, with ACTRIS GRES *In progress* for data.
- **I2 Metadata & Data: Structured Vocabulary**: Generally *Completed* status, with ACTRIS GRES and ASC *In progress* and MPLNET does not have any implementation plan at this stage for metadata.
- **I3 Metadata & Data: Semantic Model**: Status varies for metadata.

Reusability

- **R1.1 Metadata & Data: License**: Status is mostly *Completed*, but not implemented for MPLNET.
- **R1.2 Metadata & Data: Provenance and Provenance Schema**: Status varies, with many having implementation plans *In progress*.

Key observations:

- **Consistency**: Most RIs show consistent adherence to FAIR principles, particularly in Findability and Accessibility.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

- **Areas for improvement**: MPLNET and some ACTRIS DCs generally score low in completed status indicating areas where improvements are needed.
- **High performers**: ARM, NOAA GML and some ACTRIS DCs show high score in completed status in several categories indicating strong adherence to FAIR principles.

Recommendations

1. **Enhance documentation**: RIs with lower scores in completed status should enhance their metadata and documentation practices to improve findability and reusability.
2. **Adopt standard protocols**: Implementing standardized protocols for data and metadata can help improve accessibility and interoperability.
3. **Increase collaboration**: Sharing best practices among services with high FAIRness scores can help elevate the overall quality of data management.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

3. FAIR Convergence Status Between CARGO-ACT Partner Data Centres

3.1. Findability FAIR Convergence

Findability

	ACTRIS IN SITU /	ACTRIS ARES	ACTRIS CLU	ACTRIS GRES	ACTRIS ASC	ARM	MPLNET	NOAA GML (AERO)
ACTRIS IN SITU / NFAN		7	1	3	3	1	0	4
ACTRIS ARES	7		6	6	6	4	0	3
ACTRIS CLU	1	6		1	1	2	0	2
ACTRIS GRES	3	6	1		8	2	0	2
ACTRIS ASC	3	6	1	8		2	0	2
ARM	1	4	2	2	2		0	1
MPLNET	0	0	0	0	0	0		0
NOAA GML (AERO)	3	2	1	1	1	0	0	

Figure 1: Findability FAIR convergence matrix between CARGO-ACT partner data centres, i.e. number of Findability FAIR Enabling Resources (FER) shared between each pair of data centres.

This report summarizes the findability scores for different services, highlighting key observations and areas for improvement.

High scores

- **ACTRIS IN SITU**: Achieved a high score of 7 for ACTRIS ARES, indicating strong findability practices.
- **ACTRIS ARES**: Consistently scored 6 for ACTRIS CLU, GRES, and ASC, demonstrating good findability.
- **ACTRIS GRES**: Notably scored 8 for ACTRIS ASC, the highest score in the dataset, reflecting excellent findability.

Moderate scores

- **ACTRIS CLU**: Scored 6 for ACTRIS ARES, showing moderate findability.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

- **ARM**: Scored 4 for ACTRIS ARES and 2 for several other services, indicating moderate findability.

Low scores

- **MPLNET**: Scored 0 across all services, indicating a lack of findability.
- **NOAA GML**: Mostly scored 1s and 2s, suggesting low findability.

Key observations

- **ACTRIS IN SITU and ACTRIS ARES**: Generally, have higher findability scores, suggesting better practices in making their data findable.
- **MPLNET**: Consistently scores 0, indicating significant room for improvement in findability.
- **ACTRIS GRES and ACTRIS ASC**: Show strong mutual findability with scores of 8, indicating excellent practices in making their data findable to each other.

Recommendations

1. **Improve documentation**: Services with low scores should enhance their metadata and documentation practices to improve findability.
2. **Adopt standard protocols**: Implementing standardized protocols for data and metadata can help improve findability across all services.
3. **Increase collaboration**: Sharing best practices among services with high findability scores can help elevate the overall quality of data management.

This summary provides a clear overview of the current state of findability across various services and offers actionable recommendations for improvement.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

3.2. Accessibility FAIR Convergence

Accessibility

	ACTRIS IN SITU / NFAN	ACTRIS ARES	ACTRIS CLU	ACTRIS GRES	ACTRIS ASC	ARM	MPLNET	NOAA GML (AERO)
ACTRIS IN SITU / NFAN		3	4	5	5	2	1	2
ACTRIS ARES	3		4	7	5	4	1	2
ACTRIS CLU	4	4		5	6	2	1	2
ACTRIS GRES	5	7	5		6	3	1	2
ACTRIS ASC	5	5	6	6		3	1	2
ARM	2	4	2	3	3		1	2
MPLNET	1	1	1	1	1	1		1
NOAA GML (AERO)	2	2	2	2	2	2	1	

Figure 2: Accessibility FAIR convergence matrix between CARGO-ACT partner data centres, i.e. number of Accessibility FAIR Enabling Resources (FER) shared between each pair of data centres.

This report summarizes the accessibility scores for different RIs, highlighting key observations and areas for improvement.

High scores

- **ACTRIS GRES**: Achieved the highest score of 7 for ACTRIS ARES, indicating excellent accessibility practices.
- **ACTRIS ASC**: Consistently scored 6 for ACTRIS CLU and GRES, showing strong accessibility.
- **ACTRIS CLU**: Scored 6 for ACTRIS ASC, reflecting good accessibility.

Moderate scores

- **ACTRIS IN SITU**: Scores range from 2 to 5, indicating moderate accessibility.
- **ARM**: Scores range from 2 to 4, showing moderate accessibility.
- **NOAA GML**: Consistently scores 2, indicating moderate but consistent accessibility.

Low scores

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

- **MPLNET**: Scores 1 across all services, indicating very low accessibility.
- **ACTRIS ARES**: Scores 1 for MPLNET, suggesting low accessibility in this specific interaction.

Key observations

- **ACTRIS GRES and ACTRIS ASC**: Generally, have higher accessibility scores, suggesting better practices in making their data accessible.
- **MPLNET**: Consistently scores 1, indicating significant room for improvement in accessibility.
- **ACTRIS IN SITU and ARM**: Show moderate accessibility, with room for improvement in certain areas.

Recommendations

1. **Enhance protocols**: Services with low scores should adopt standardized communication protocols to improve accessibility.
2. **Improve documentation**: Better documentation and user guides can help increase accessibility.
3. **Increase collaboration**: Sharing best practices among services with high accessibility scores can help elevate the overall quality of data management.

This summary provides an overview of the current state of accessibility across various services and offers actionable recommendations for improvement.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

3.3. Interoperability FAIR Convergence

Interoperability

	ACTRIS IN SITU / NFAN	ACTRIS ARES	ACTRIS CLU	ACTRIS GRES	ACTRIS ASC	ARM	MPLNET	NOAA GML (AERO)
ACTRIS IN SITU / NFAN		4	5	4	4	4	3	6
ACTRIS ARES	4		4	3	3	4	3	4
ACTRIS CLU	5	4		4	4	4	4	4
ACTRIS GRES	4	3	4		7	4	4	4
ACTRIS ASC	4	3	4	7		4	4	4
ARM	4	4	4	4	4		3	3
MPLNET	3	3	4	4	4	3		3
NOAA GML (AERO)	6	4	4	4	4	3	3	

Figure 3: Interoperability FAIR convergence matrix between CARGO-ACT partner data centres, i.e. number of Interoperability FAIR Enabling Resources (FER) shared between each pair of data centres.

This report summarizes the interoperability scores for different RIs, highlighting key observations and areas for improvement.

High scores

- **ACTRIS GRES**: Achieved the highest score of 7 for ACTRIS ASC, indicating excellent interoperability practices.
- **NOAA GML**: Scored 6 for ACTRIS IN SITU, reflecting strong interoperability.

Moderate scores

- **ACTRIS IN SITU**: Scores range from 3 to 6, indicating moderate to strong interoperability.
- **ACTRIS CLU**: Consistently scores 4 and 5, showing good interoperability.
- **ARM**: Scores range from 3 to 4, indicating moderate interoperability.
- **ACTRIS ARES**: Scores mostly 3 and 4, showing moderate interoperability.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

Low scores

- **MPLNET**: Scores mostly 3 and 4, indicating moderate but lower interoperability compared to other services.

Key observations

- **ACTRIS GRES and ACTRIS ASC**: Generally, have higher interoperability scores, suggesting better practices in ensuring data can be integrated and used across different systems.
- **NOAA GML**: Shows strong interoperability with ACTRIS IN SITU, indicating effective practices in data integration.
- **MPLNET**: Consistently scores lower, indicating room for improvement in interoperability.

Recommendations

1. **Adopt common standards**: Services with lower scores should adopt common standards and protocols to improve interoperability.
2. **Enhance data integration practices**: Implementing better data integration practices can help improve interoperability across all services.
3. **Increase collaboration**: Sharing best practices among services with high interoperability scores can help elevate the overall quality of data management.

3.4. Reusability FAIR Convergence

Reusability

	ACTRIS IN SITU / NFAN	ACTRIS ARES	ACTRIS CLU	ACTRIS GRES	ACTRIS ASC	ARM	MPLNET	NOAA GML (AERO)
ACTRIS IN SITU / NFAN		0	2	3	3	1	0	0
ACTRIS ARES	0		1	0	0	2	0	0
ACTRIS CLU	2	1		2	2	1	0	0
ACTRIS GRES	3	0	2		4	1	0	1
ACTRIS ASC	3	0	2	4		1	0	1
ARM	1	2	1	1	1		0	0
MPLNET	0	0	0	0	0	0		0
NOAA GML (AERO)	0	0	0	1	1	0	0	

Figure 4: Reusability FAIR convergence matrix between CARGO-ACT partner data centres, i.e. number of Reusability FAIR Enabling Resources (FER) shared between each pair of data centres.

This report summarizes the reusability scores for different RIs, highlighting key observations and areas for improvement.

High scores

- **ACTRIS GRES**: Achieved the highest score of 4 for ACTRIS ASC, indicating strong reusability practices.
- **ACTRIS ASC**: Scored 4 for ACTRIS GRES, reflecting good reusability.

Moderate scores

- **ACTRIS IN SITU**: Scores range from 0 to 3, indicating moderate reusability.
- **ACTRIS CLU**: Scores range from 1 to 2, showing moderate reusability.
- **ARM**: Scores range from 1 to 2, indicating moderate reusability.

Low scores

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

- **MPLNET**: Scores 0 across all services, indicating no reusability.
- **NOAA GML**: Scores mostly 0, with a single score of 1 for ACTRIS GRES and ASC, indicating low reusability.
- **ACTRIS ARES**: Scores mostly 0, with a single score of 2 for ARM, suggesting low reusability.

Key observations

- **ACTRIS GRES and ACTRIS ASC**: Generally, have higher reusability scores, suggesting better practices in ensuring data can be reused effectively.
- **MPLNET**: Consistently scores 0, indicating significant room for improvement in reusability.
- **ACTRIS IN SITU and ACTRIS CLU**: Show moderate reusability, with room for improvement in certain areas.

Recommendations

1. **Enhance documentation**: Services with low scores should improve their documentation and metadata to enhance reusability.
2. **Adopt standard licenses**: Implementing standardized licenses can help improve the reusability of data.
3. **Increase collaboration**: Sharing best practices among services with high reusability scores can help elevate the overall quality of data management.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

3.5. Overall FAIR Convergence

Overall

	ACTRIS IN SITU / NFAN	ACTRIS ARES	ACTRIS CLU	ACTRIS GRES	ACTRIS ASC	ARM	MPLNET	NOAA GML (AERO)
ACTRIS IN SITU / NFAN		14	12	15	15	8	4	12
ACTRIS ARES	14		15	16	14	14	4	9
ACTRIS CLU	12	15		12	13	9	5	8
ACTRIS GRES	15	16	12		25	10	5	9
ACTRIS ASC	15	14	13	25		10	5	9
ARM	8	14	9	10	10		4	6
MPLNET	4	4	5	5	5	4		4
NOAA GML (AERO)	11	8	7	8	8	5	4	

Figure 5: Overall FAIR convergence matrix between CARGO-ACT partner data centres, i.e. overall number of FAIR Enabling Resources (FER) shared between each pair of data centres.

This report provides an overall analysis of the FAIRness scores for different RIs, highlighting key observations and areas for improvement.

High scores

- **ACTRIS GRES**: Achieved the highest overall score of 25 for both ACTRIS ASC and itself, indicating excellent adherence to FAIR principles.
- **ACTRIS ASC**: Also scored 25 for ACTRIS GRES, reflecting strong FAIR practices.

Moderate scores

- **ACTRIS IN SITU**: Scores range from 8 to 15, indicating moderate to strong adherence to FAIR principles.
- **ACTRIS ARES**: Scores range from 4 to 16, showing moderate to strong adherence.
- **ACTRIS CLU**: Scores range from 5 to 15, indicating moderate adherence.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

- **ARM**: Scores range from 4 to 14, showing moderate adherence.
- **NOAA GML**: Scores range from 4 to 11, indicating moderate adherence.

Low scores

- **MPLNET**: Consistently scores 4 to 5, indicating low adherence to FAIR principles.

Key observations

- **ACTRIS GRES and ACTRIS ASC**: Generally, have the highest overall scores, suggesting excellent practices in implementing FAIR principles.
- **ACTRIS IN SITU and ACTRIS ARES**: Show moderate to strong adherence, with room for improvement in certain areas.
- **MPLNET**: Consistently scores low, indicating significant room for improvement in all aspects of FAIRness.

Recommendations

1. **Enhance documentation and metadata**: Services with lower scores should improve their documentation and metadata practices to enhance overall FAIRness.
2. **Adopt standard protocols and licenses**: Implementing standardized protocols and licenses can help improve adherence to FAIR principles.
3. **Increase collaboration and sharing of best practices**: Services with high scores should share their best practices with those scoring lower to elevate the overall quality of data management.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

4. FAIR Enabling Resources Used by CARGO-ACT Partner Data Centres

Table 2: Summary of FERs used at the CARGO-ACT partner data centres to implement the FAIR data services. Topics where FAIR convergence could be improved most efficiently are highlighted in green.

CARGOACT RIs
FERs used

FAIRness service	ACTRIS IN SITU / NFAN	ACTRIS ARES	ACTRIS CLU	ACTRIS GRES	ACTRIS ASC	ARM	MPLNET	NOAA GML (AERO)
Findability								
F1 Metadata: persistent identifier	DOI	Handle system, DOI, ORCID	Handle system, ROR, ORCID	UUID, DOI	UUID, DOI	DOE Data ID Service (DOI)	No data*	DOI
F1 Data: persistent identifier	DOI	Handle system, DOI	Handle system, DOI	DOI, Epic	DOI	DOE Data ID Service (DOI)	No data*	DOI
F2: Metadata schema	WMO	WMO Core Profile, NetCDF 1.7, DataCite	DataCite	DataCite, ISO 19115-GI-Metadata	DataCite, ISO 19115-GI-Metadata	ISO 19115-GI-Metadata, Schema.org, NetCDF 1.7,	No data*	WMO Core Profile
F3: Schema linking data PIDs to metadata	DataCite	DataCite, Handle System	Handle system	DataCite	DataCite	DataCite	No data*	DOI
F4 Metadata: publication service	ACTRIS data portal, WMO Search, GEOSS Portal, SIOS metadata search	ACTRIS data portal, DataCite, WMO Search, re3data	ACTRIS data portal, DataCite	DataCite	DataCite	DataCite, DataONE, re3data, ORCID, Google Data Search	No data*	EBAS database
F4 Data: publication service	EBAS database, ACTRIS data portal, DataCite	DataCite, WMO Search, re3data	ACTRIS data portal, DataCite	DataCite	DataCite	DataONE, Google dataset search	No data*	EBAS database
Accessibility								
A1.1 Metadata: standardized communication protocol	OAI-PMH	HTTPS, REST, OPenDAP	HTTPS, REST	HTTPS, CSW, OAI-PMH, REST	HTTPS REST	HTTPS, FTP, OPeNDAP, REST	HTTPS	HTTPS
A1.1 Data: standardized communication protocol	HTTPS, OPenDAP	HTTPS, FTP, OPenDAP	HTTPS	HTTPS, REST, OPenDAP	HTTPS, OPenDAP	HTTPS, FTP, OPeNDAP	HTTPS	HTTPS, FTP
A1.2 Metadata: AAI service	Open Data	Oauth, SAML1.1, ORCID	Open Data	OpenData	OpenData	HTTPS	HTTPS	OpenData
A1.2 Data: AAI service	Open Data	Oauth, SAML1.1, ORCID	Open Data	Oauth, ORCID, eduGAIN	OpenData	LDAP	HTTPS	OpenData
A2: Metadata preservation policy	ACTRIS Data management plan	ACTRIS Data management plan	ACTRIS Data management plan	ACTRIS Data management plan	ACTRIS Data management plan	DOE Data policy	No data*	NCEI
Interoperability								
I1 Metadata: Knowledge representation language	XMLS	JSON Scheme, SKOS	JSON	JSON	JSON	XML, JSON	CSV, JSON	XMLS
I1 Data: Knowledge representation language	NetCDF	NetCDF	NetCDF	NetCDF	NetCDF	JSON-LD, NetCDF, CSV	NetCDF	NetCDF, Text
I2 Metadata: structured vocabulary	ACTRIS, CF-SN	ACTRIS, NetCDF 1.7, WIGOS	ACTRIS, CF SN	ACTRIS, GCMD	ACTRIS, GCMD	CF SN, GCMD, Schema.org, NetCDF 1.7	No data*	ACTRIS, CF-SN
I2 Data: structured vocabulary	ACTRIS, CF-SN	ACTRIS, NetCDF 1.7	ACTRIS	ACTRIS, CF SN	ACTRIS, CF SN	NetCDF 1.7	CF SN	ACTRIS, CF-SN
I3 Metadata: semantic model	WMO Core Profile, DataCite, I-ADOPT (planned)	DataCite (planned), I-ADOPT framework	DataCite, I-ADOPT (planned)	ISO 19115-GI-Metadata	ISO 19115-GI-Metadata	DataCite, NetCDF 1.7	No data*	WMO Core Profile
I3 Data: semantic model	NetCDF CF-1.7	DataCite, I-ADOPT framework	NetCDF 1.7	NetCDF CF 1.7, GEOMS	NetCDF 1.7, JCAMP-DX, EUROCHAMP data format	NetCDF 1.7	NetCDF 1.7	NetCDF 1.7
Reusability								
R1.1 Metadata: License	CC BY 4	CC BY 4	CC BY 4	CC0 1.0	CC0 1.0	CC BY 4	No data*	CC0 1.0
R1.1 Data: License	CC BY 4	CC BY 4	CC BY 4	CC BY 4	CC BY 4	CC BY 4	No data*	CC0 1.0
R1.2 Metadata: Provenance and provenance schema	PROV-O	PROV-O	PROV-O	PROV-O	PROV-O	ISO 19115-GI-Metadata, Schema.org, NetCDF 1.7	No data*	No data*
R1.2 Data: Provenance and provenance schema	PROV-O	PROV-O	PROV-O	PROV-O	PROV-O	NetCDF 1.7	No data*	No data*

This section provides an analysis of the FAIRness data for different data centres, focusing on the four FAIR principles: Findability, Accessibility, Interoperability, and Reusability.

Findability

- **Persistent identifiers**: Most data centres use DOIs for both metadata and data, ensuring unique and persistent identification. Exceptions include MPLNET, which lacks data and metadata identifiers, and ACTRIS CLU and GRES, which are implementing data identifiers.
- **Metadata schema**: Commonly used schemas include WMO, DataCite, and ISO 19115-GI-Metadata. ACTRIS ARES and ARM use a broader range of schemas, enhancing findability.
- **Schema linking data PIDs to metadata**: DataCite is widely used, ensuring that data and metadata are linked effectively.
- **Publication services**: Services like DataCite, WMO Search, and EBAS database are commonly used for publishing metadata and data, facilitating discoverability.

Accessibility

- **Standardized communication protocols**: HTTPS and OPeNDAP are commonly used, ensuring secure and standardized access to metadata and data.
- **AAI services**: Most services use open data access, with some employing OAuth, SAML, and ORCID for authentication and authorization.
- **Metadata preservation policy**: ACTRIS services follow a consistent data management plan, while ARM adheres to the DOE Data policy. NOAA GML lacks a documented preservation policy.

Interoperability

- **Knowledge representation languages**: JSON and NetCDF are widely used, ensuring that data can be understood and processed by different systems.
- **Structured vocabularies**: ACTRIS and CF-SN vocabularies are commonly used, promoting consistency in data description.
- **Semantic models**: DataCite and ISO 19115-GI-Metadata are frequently used, enhancing the semantic understanding of data.

Key Observations

- **Strengths**: Most services demonstrate strong adherence to FAIR principles, particularly in findability and interoperability.
- **Areas for improvement**: MPLNET consistently lacks data identifiers, and structured vocabularies for metadata, indicating significant room for improvement. NOAA GML also shows gaps in data identifiers and preservation policies.

Recommendations

1. **Enhance documentation and metadata**: Ris with lower scores should improve their documentation and metadata practices to enhance overall FAIRness.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093

2. **Adopt standard protocols and licenses**: Implementing standardized protocols and licenses can help improve adherence to FAIR principles.
3. **Increase collaboration and sharing of best practices**: RIs with high scores should share their best practices with those scoring lower to elevate the overall quality of data management.

This summary provides an overview of the current state of FAIRness across various services and Ris and offers actionable recommendations for improvement.

5. Key Results

- The granularity of data DOIs varies between data centres. This can at least be inconvenient for the user, and could be improved.
- CARGO-ACT data centres don't use any common data search and discovery service.
- There is diversity in the technologies for metadata endpoints, schemas, and profiles. Convergence here would make the data much easier to use, e.g. in virtual research environments.
- Data centres use different vocabulary for metadata annotation. Convergence here would improve Findability and Re-usability.
- Data centres don't have a common concept, if any, for documenting data provenance.

CARGO-ACT (www.cargo-act.eu) is supported by the European Commission under the Horizon Europe (HORIZON) – Strengthening the international dimension of ESFRI and/or ERIC research infrastructures, HORIZON-INFRA-2023-DEV-01-07, Grant Agreement #: 101132093