



# D4.4

## Report on external developments in FAIR services

Work Package	WP4
Lead partner	UKRI
Status	Final
Deliverable type	Report
Dissemination level	Public
Due date	28 February 2023
Submission date	7 July 2023

### Deliverable abstract

This deliverable provides an overview of relevant developments in the practices and technologies that support making research assets findable, accessible, interoperable, and reusable, especially where these are relevant for the ENVRI cluster of environmental research infrastructures.

A review of selected initiatives and associated technologies is provided with some context for their relevance to the ENVRI community, especially those that are represented within the ENVRI-FAIR consortium.



## DELIVERY SLIP

	Name	Partner Organisation	Date
Main Author	Helen Glaves	UKRI	20 May 2023
Contributing Authors			
Reviewer(s)	Keith Jeffery	UKRI	23 June 2023
	Barbara Magagna	GFF	29 June 2023
Approver	Andreas Petzold	FZJ	7 July 2023

## DELIVERY LOG

Issue	Date	Comment	Author
V1	15/06/2023	Initial draft	Helen Glaves
V2	23/06/2023	Improved version	Keith Jeffery
V3	25/06/2023	Final updated version	Helen Glaves

## DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the Project Manager at [manager@envri-fair.eu](mailto:manager@envri-fair.eu).

## GLOSSARY

A relevant project glossary is included in Appendix A. The latest version of the master list of the glossary is available at <http://doi.org/10.5281/zenodo.4471374>.

## 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 subdomains 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.....	4
2	FAIR in ENVRI .....	4
3	External FAIR developments.....	4
3.1	RDA .....	4
3.1.1	FAIR Maturity Model.....	4
3.1.2	FAIR Principles for Research Software (FAIR4RS Principles) .....	4
3.2	GO FAIR.....	5
3.3	FAIR’s FAIR.....	5
3.4	FAIR-IMPACT .....	6
4	FAIR Assessment Tools.....	6
4.1	Relevant FAIR Assessment Tools.....	7
4.1.1	F-UJI Automated FAIR Data Assessment Tool .....	7
4.1.2	Automated Data Management Plan (DMP) Tools .....	7
4.2	FAIR Implementation Profiles (FIPs) .....	7
5	Influence on ENVRI-FAIR Policy.....	8
6	Conclusions.....	8
7	Impact on the project.....	8
8	References.....	9
9	ANNEX A GLOSSARY.....	10

## D4.4 – Report on external developments in FAIR services

### 1 Introduction

The FAIR Principles, although formulated by Force11<sup>1</sup>, were first formally documented by Wilkinson et al (2016) and have since gained significant traction throughout the research community as a pillar of the open science paradigm. The FAIR Principles provide a set of 15 guidelines for making data Findable, Accessible, Interoperable and Reusable.

Meanwhile, many pioneering implementations and publications about FAIR have emerged. This signifies worldwide adoption of the principles in all disciplines of science and innovation, but this rapid uptake also makes the guiding principles vulnerable to misinterpretation, misapplication and inflation.

### 2 FAIR in ENVRI

The overarching goal of the ENVRI-FAIR project is that all participating RIs will improve their FAIRness as a pre-requisite for connection to the European Open Science Cloud (EOSC).

### 3 External FAIR developments

A number of ongoing initiatives are directly addressing the advancement and implementation of the FAIR Principles, which are relevant to the RIs that are participating in the ENVRI-FAIR project. These initiatives bring together relevant experts, many of whom are also connected with the ENVRI cluster of environmental research infrastructures, to further develop the implementation of the FAIR Principles.

#### 3.1 RDA

The Research Data Alliance (RDA) is an international initiative that aims to foster sharing and re-using of data. RDA operates through a series of interest and working groups that focus on specific topics and bring together experts from all disciplines that have a common interest in a specific area of interest. Within the context of RDA there are several groups that are of direct relevance to ENVRI-FAIR, and many of the beneficiaries are active members in these particular groups that are further described below.

##### 3.1.1 FAIR Maturity Model

The RDA FAIR Maturity Model working group<sup>2</sup> is maintaining deliverables and endorsed recommendations (FAIR Data Maturity Model Working Group, 2020) for a common set of core assessment criteria for FAIRness along with a generic and expandable self-assessment model for measuring the maturity level of a dataset. The outcome of this limited duration working group is to provide a framework to assess the uptake and implementation of the FAIR Principles within specific applications such as a research infrastructure.

A major contribution of the group has been the elaboration of the meaning of each of the FAIR principles as initially proposed by Force 11. In particular, this elaboration included discussion that led to conclusions about what was meant by ‘identifier’, ‘reference’, ‘qualified reference’ and so on. This detail is necessary to provide a framework for evaluation of the FAIRness of digital assets. The framework has been utilised in several other FAIR projects, many described below, and within which there is ENVRI representation.

##### 3.1.2 FAIR Principles for Research Software (FAIR4RS Principles)

In order to deliver verifiable and reproducible science, there is a need to ensure that it not only the data that are FAIR, the relevant digital research assets such as software must also be findable, accessible, interoperable, and reusable (Barker et al 2022). To this end, the RDA FAIR for Research Software

<sup>1</sup> <https://force11.org/info/the-fair-data-principles/>

<sup>2</sup> <https://www.rd-alliance.org/groups/fair-data-maturity-model-wg>

(FAIR4RS) working group adapted the FAIR Principles to create the FAIR Principles for Research Software (FAIR4RS Principles).

Increasingly, FAIR is being implemented by software services which provide access to the underlying datasets. The advantages include management of access (including authentication and authorisation, and adherence to licensing conditions which can only be achieved with software) but also as dataset sizes increase, downloading becomes expensive (in bandwidth and time) and so data management functions in the software services are invoked to select only the relevant parts of the dataset pertaining to the user request.

The community- endorsed FAIR4RS Principles have provided a further tool for ensuring transparency and reproducibility of research, which are being widely adopted throughout the research community.

Further information regarding the FAIR4RS Principles was published in a 2002 Nature paper that can be accessed at: <https://www.nature.com/articles/s41597-022-01710-x>

## 3.2 GO FAIR

The GO FAIR initiative<sup>3</sup> is a bottom-up, self-governing organisation that aims to promote and facilitate the uptake of the FAIR principles through a range of activities that include coordination, training, and capacity building. GO FAIR is also advocating for the Internet of FAIR Data and Services that relies on a minimum set of common standards and protocols that maximises accessibility for all stakeholders. This has led to an effort in the GO FAIR community to formulate an explicit and more precise set of interpretations of the original FAIR guiding principles and implementation considerations. The documented interpretation of the principles can be found on the website of the GO FAIR Foundation (GFF)<sup>4</sup>, and along with additional criteria including an open, decentralised approach, are now referred to as the GO FAIR criteria<sup>5</sup>.

Within Europe, the European Open Science Cloud (EOSC) is a pivotal component of the Internet of FAIR Data and Services that will foster access and use of data and information for all researchers.

The ENVRI-Hub being developed by the ENVRI-FAIR project has already delivered a metadata record describing service to the EOSC marketplace and will progressively deliver data and services directly to the EOSC, which will also facilitate access to the data and services provided by the participating RIs within the ENVRI cluster.

GFF has also been instrumental in developing the FAIR Implementation Profile (FIP) that is increasingly being adopted throughout the research community for the purposes of documenting the implementation of FAIR practices within a specific community of practice (CoP). See section 4.2 below for further information. GFF has set up a training programme<sup>6</sup> and a fellowship programme<sup>7</sup> to learn more how to implement FAIR. In addition, GFF is building the FAIR Connect portal<sup>8</sup>, which is becoming a knowledge base of FAIR Supporting Resources published as machine-readable nanopublications and semi-automatic generated short articles published in the FAIR Connect Journal<sup>9</sup>.

## 3.3 FAIR's FAIR

FAIRsFAIR<sup>10</sup> fosters FAIR data practices in Europe by providing practical solutions. The emphasis is on best practices. The project has been working on global standards for certifying repositories and the data within them and links this with ensuring the content of EOSC is FAIR.

---

<sup>3</sup> <https://www.go-fair.org/>

<sup>4</sup> <https://www.gofair.foundation/interpretation>

<sup>5</sup> <https://www.gofair.foundation/criteria>

<sup>6</sup> <https://osf.io/bthf8>

<sup>7</sup> <https://osf.io/x43g7>

<sup>8</sup> <https://fairconnect.pro/>

<sup>9</sup> <https://content.iospress.com/journals/fair-connect/>

<sup>10</sup> <https://www.fairsfair.eu/>

The EU-funded FAIRsFAIR project aims to develop practices and procedures that support the adoption and implementation of the FAIR Principles across data providers and repositories. This framework is backed up by training and certification processes that can be consistently adopted across those stakeholders that are responsible for some aspect of the data life cycle. The emphasis on repositories provides a higher level of abstraction than that represented with the RIs that make up the ENVRI cluster where the unit of granularity is at the level of the digital asset – e.g., dataset or software service. However, some repositories within the ENVRI-FAIR project are already certified by CoreTrustSeal. Nonetheless, the proposed best practices provide a useful basis that can be adapted as part of the continuing development of policies in ENVRI.

These efforts to coordinate and align the implementation of the FAIR principles across different stakeholders that include the research infrastructures is critical for the successful realisation of the European Open Science Cloud (EOSC). As such, the resulting Rules of Participation and associated best practices should be implemented as part of the ENVRI-Hub and adopted by the participating RIs within the ENVRI-FAIR project consortium.

### 3.4 FAIR-IMPACT

The FAIR-IMPACT project<sup>11</sup> has commenced more recently. It is one of several initiatives funded by the EU to support the realisation of the EOSC. The project aims to deliver practical solutions for FAIR implementation with a focus on the use of persistent identifiers (PIDs), metadata, ontologies, metrics, certification standards and interoperability. solutions that are demonstrated through four integrated use cases.

The overall aim of the FAIR-IMPACT project is to develop a web of transversal FAIR data related services that cut across discipline and builds upon FAIRsFAIR. The project has engaged with ENVRI requesting a presentation on the ENVRI approach to FAIRness and, further, the architecture of RIs with associated portals and catalogues leading to the ENVRI-Hub portal with its catalogue acting as a unifying node. The ENVRI community will maintain an interaction with this project, which provides real practical guidelines, best practice and tools that will be achieved through co-development with selected stakeholders and research communities.

## 4 FAIR Assessment Tools

The implementation of the FAIR Principles has also driven development of associated technologies including assessment tools that determine the compliance of individual repositories or digital assets in those repositories with the FAIR principles. The overarching objective of these tools is to assess the FAIRness of a research asset and determine its availability for reuse. The RDA FAIR Maturity Model working group conducted an evaluation of the current FAIR assessment tools that are available published in 2019 (Bahim et al 2019). Depending on the interpretation of the FAIR Principles the results of these FAIR Assessment Evaluator can vary a lot (Wilkinson et al., 2022) and thus it is necessary critical to make those interpretations explicit.

The landscape of tools has changed since then across different initiatives, disciplines and domains and several recent research papers, e.g. (Krans et al., 2022), have attempted to make an evaluation of the effectiveness of the various FAIR assessment tools.

There has been much discussion among the various FAIR projects about FAIR assessment. At one end of the spectrum, some wish to have an agreed international quantitative method and set of metrics for assessing FAIRness, and on the other end of the spectrum qualitative assessments like the FIP approach were promoted to draw the landscape of used FAIR Enabling Resources (FERs) and how this might converge between communities like ENVRI. Some consider this as unnecessary heavyweight processes and form-filling to assess FAIRness, as experienced metadata analysts can actually look at the schema for a digital asset and assess its FAIRness quickly. Opinions within ENVRI vary between these extremes. At present, there has been an effort to assess FAIRness of ENVRI digital assets using FIPs, which has informed RIs about the amount of effort required to do this evaluation. It might though be useful to apply

---

<sup>11</sup> <https://fair-impact.eu/>

a mixed approach – describe the implemented FERs and assess their FAIRness level, to actually be aware which roadmap might best help to implement FAIR across ENVRI.

At present there is no universally accepted FAIR assessment tool that can be used for datasets, data products, software services, software source (including scripts), workflows and so on. ENVRI will continue to monitor the situation – especially through contacts in the current FAIR-IMPACT project and RDA.

The RDA FAIR Maturity Model working group conducted an evaluation of the current FAIR assessment tools that are available as a published report (Bahim et al 2019). However, this report is not entirely comprehensive due to more recent developments.

## 4.1 Relevant FAIR Assessment Tools

### 4.1.1 F-UJI Automated FAIR Data Assessment Tool

The FAIRsFAIR project has developed the F-UJI Automated FAIR Data Assessment Tool as a pilot for assessing the FAIRness of datasets within selected repositories. A test instance is also available via the FAIRsFAIR website at: <https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>

### 4.1.2 Automated Data Management Plan (DMP) Tools

Data Management Plans (DMPs) are in narrative form and describe intentions of the researchers; many funders require them as part of the grant application and its execution. The RDA DMP Common Standard WG has developed the *maDMP Common Model*, which has already been adopted to varying degrees by several DMP platforms (e.g., DSW<sup>12</sup>, Argos, DAMAP<sup>13</sup>). Furthermore, recognising the diversity of data and RDM practices across scientific disciplines and domains, Science Europe has proposed the concept of [Domain Data Protocols](#) (DDPs)<sup>14</sup> to serve as domain-specific DMP templates.

Current research activity in several organisations is underway to transform DMPs from simple plans into live, machine-actionable guidance to enable FAIR-by-design and to be used and revised at any point in the research life cycle, to track and to assess the level of FAIRness reached by the relevant digital objects (DOs). Specifically, this implies a cultural shift in research by cultivating an interoperable maDMP ecosystem that (a) includes various types of DOs, especially software since this is becoming a first-class citizen, (b) interconnects them in a coherent way to form an RDM story, (c) is fully connected to and utilises information from EOSC infrastructures (e.g., PIDs, SKGs, funding, ontologies) via common APIs, (d) considers specificities of scientific domains (e.g., what SKGs to link to, and which FAIR assessments are better suited to a domain) by extending the *RDA maDMP Common Model* with DDPs and by creating a commons to lower the adoption barrier, (e) automates assessment via seamless connections to FAIR Assessment tools (via common APIs) providing the much needed semi-automated guidance, and (f) explores the concept of *Reproducibility Management Plans* (RMPs) by extending maDMPs with additional necessary elements and/or metrics.

ENVRI-FAIR participants have been involved in the relevant RDA groups and in discussions with the current providers of maDMPs. This has also been discussed across the five science clusters (led by ENVRI, specifically UKRI). This approach opens an evolutionary route to ensure FAIRness by design rather than FAIRness by evaluation and improvement.

## 4.2 FAIR Implementation Profiles (FIPs)

The FAIR Implementation Profile (FIP) has been promulgated by the GFF as a system for a community of practice to document their decision choices for implementation of each of the FAIR Principles<sup>15</sup> using the FIP Wizard<sup>16</sup>. These FIPs can be shared with other CoPs as a resource for fostering and supporting

<sup>12</sup> <https://ds-wizard.org/>

<sup>13</sup> <https://damap.org/>

<sup>14</sup> [https://www.snf.ch/media/en/Sz1vykx3F3IOx6Eht/SE\\_Guidance\\_Document\\_RDMPs.pdf](https://www.snf.ch/media/en/Sz1vykx3F3IOx6Eht/SE_Guidance_Document_RDMPs.pdf)

<sup>15</sup> <https://osf.io/5ygzx>

<sup>16</sup> <https://fip-wizard.ds-wizard.org/>

wider adoption of the FAIR guiding principles that support open and accessible research assets (data, software, workflows etc.)

To support the generation of FIPs and encourage convergence of practice, the GFF has developed the FAIR Convergence Matrix, which allows communities to optimise their FAIR implementation practices. More details on the concept and implementation of FAIR Implementation Profiles can be found in Schultes et al. (2020). The FIP Wizard team has developed a service to allow mappings between the FIP and the maDMP translating the FER choices into community-specific directives that data stewards can subsequently implement (Magagna et. al, 2022).

## 5 Influence on ENVRI-FAIR Policy

FAIR has been a large influence on ENVRI-FAIR Policy. As indicated in the associated deliverables, FAIR (along with EOSC, GDPR and other lesser drivers) has formed the basis of the requirements leading to policy statements and hence to the policy documents at each RI (which in turn have led to guidelines (desk instructions) and to IT implementation to follow the policies based on the guidelines – such as open access (subject to authentication and authorisation), personal data protection, security (to ensure service availability), metadata standards to be used and more.

However, policy has also been heavily influenced by legal requirements such as The General Data Protection Regulation (GDPR), and attention to personal data privacy, terms and conditions and cookies (with declarations of organisational legitimate business use and a user consent pop-up) and this has taken priority over FAIRness. Having said that, with a rich metadata model for the ENVRI-Hub (based on that of EPOS) a high level of FAIRness has been achieved.

## 6 Conclusions

While the concepts of findable, accessible, interoperable, and reusable are well document in the FAIR Principles, and widely recognised throughout the ENVRI community there are differing levels of adoption and implementation across the ENVRI cluster of RIs.

ENVRI-FAIR has driven the uptake of the FAIR Principles throughout the ENVRI community both through the development of relevant policies and best practices as part of the project, and by active participation in other relevant initiatives such as RDA and GO FAIR. The ENVRI-FAIR project has capitalised on the relevant outcomes of other European and international initiatives that are seeking to promote FAIR including those developing the technical solutions necessary for assessment of compliance with the FAIR Principles. In addition, ENVRI-FAIR has developed the ENVRI-Hub that supports the delivery of FAIR compliant services to the European Open Science Cloud (EOSC), which has been a major driver for wider adoption of the FAIR Principles within the ENVRI cluster.

## 7 Impact on the project

The ongoing review of developments related to the FAIR Principles relevant to the project has provided wider insights that ensure that the data and services delivered by the participating Research Infrastructures are aligned with current best practice. This is particularly important to ensure that the data delivered by the different RIs can be reused as widely as possible including interoperability with data from other disciplines and domains. This extended interoperability is becoming an increasing priority for addressing large scale global challenges such as climate change and food security.

In addition, the differing maturity of the participating RIs implies that there are different levels of awareness and adoption of the FAIR Principles. The ENVRI-FAIR project can ensure that relevant advances in both best practices and technologies are disseminated and implemented throughout the ENVRI community including through adoption within the ENVRI-Hub. Ensuring that the RIs and ENVRI-Hub are fully aligned with current best practices and technological innovations is critical for the sustainability of both the participating RIs and the ENVRI-Hub.



## 8 References

Bahim, Christophe, Dekkers, Makx, & Wyns, Brecht. (2019). Results of an Analysis of Existing FAIR Assessment Tools. <https://doi.org/10.15497/rda00035>

Barker, M., Chue Hong, N.P., Katz, D.S. et al. *Introducing the FAIR Principles for research software*. Sci Data 9, 622 (2022). <https://doi.org/10.1038/s41597-022-01710-x>

FAIR Data Maturity Model Working Group. (2020). FAIR Data Maturity Model. Specification and Guidelines (1.0). <https://doi.org/10.15497/rda00050>

N.A. Krans, A. Ammar, P. Nymark, E.L. Willighagen, M.I. Bakker, J.T.K. Quik, (2022) *FAIR assessment tools: evaluating use and performance*; NanoImpact, Volume 27, 2022,100402 ISSN 2452-0748, <https://doi.org/10.1016/j.impact.2022.100402>

Magagna B, Schultes EA, Suchánek M, Kuhn T (2022) FIPs and Practice. Research Ideas and Outcomes 8: e94451. <https://doi.org/10.3897/rio.8.e94451>

Schultes, E., Magagna, B., Hettne, K.M., Pergl, R., Suchánek, M., Kuhn, T. (2020). *Reusable FAIR Implementation Profiles as Accelerators of FAIR Convergence*. In: Grossmann, G., Ram, S. (eds) *Advances in Conceptual Modeling. ER 2020. Lecture Notes in Computer Science*, vol 12584. Springer, Cham. [https://doi.org/10.1007/978-3-030-65847-2\\_13](https://doi.org/10.1007/978-3-030-65847-2_13)

Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJ, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PA, Hoofstede R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone SA, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. *The FAIR Guiding Principles for scientific data management and stewardship*. Sci Data. 2016 Mar 15;3:160018. doi: 10.1038/sdata.2016.18. Erratum in: Sci Data. 2019 Mar 19;6(1):6. PMID: 26978244; PMCID: PMC4792175.

Wilkinson, M.D, Sansone, S.-A., Grootveld, M., Nordling, J., Dennis, R., & Hecker, D. (2022). FAIR Assessment Tools: Towards an "Apples to Apples" Comparisons. Zenodo. <https://doi.org/10.5281/zenodo.7463421>

## 9 ANNEX A GLOSSARY

<b>ACRONYM</b>	<b>MEANING</b>
DDP	Data Domain Protocols
DMP	Data Management Plan
DO	Digital Object
EOSC	European Open Science Cloud
EPOS	European Plate Observing System
ESFRI	European Strategy Forum on Research Infrastructures
FAIR	Findable, Accessible, Interoperable, Reusable
FIP	FAIR Implementation Plan
GDPR	General Data Protection Regulation
GEANT	The European network connecting National Research and Education Networks
IT	Information Technology
maDMP	machine actionable Data Management Plan
PID	Persistent Identifier
RDA	Research Data Alliance
RI	Research Infrastructure
SKG	Science Knowledge Graph