



European  
Commission

# Recommendations on certifying services required to enable FAIR within EOSC

Report from the  
EOSC Executive  
Board FAIR  
Working Group  
(WG)

Independent  
Expert  
Report

EOSC Executive Board  
FAIR WG  
January 2021

Research and  
Innovation

## Recommendations on certifying services required to enable FAIR within EOSC

European Commission  
Directorate-General for Research and Innovation  
Directorate G — Research and Innovation Outreach  
Unit G.4 — Open Science  
Contact Corina Pascu  
Email [Corina.PASCU@ec.europa.eu](mailto:Corina.PASCU@ec.europa.eu)  
[RTD-EOSC@ec.europa.eu](mailto:RTD-EOSC@ec.europa.eu)  
[RTD-PUBLICATIONS@ec.europa.eu](mailto:RTD-PUBLICATIONS@ec.europa.eu)  
European Commission  
B-1049 Brussels

Manuscript completed in December 2020.

The European Commission is not liable for any consequence stemming from the reuse of this publication.  
The views expressed in this publication are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.

More information on the European Union is available on the internet (<http://europa.eu>).

---

PDF	ISBN 978-92-76-28112-2	doi: 10.2777/127253	KI-04-20-743-EN-N
-----	------------------------	---------------------	-------------------

---

Luxembourg: Publications Office of the European Union, 2021

© European Union, 2021



The reuse policy of European Commission documents is implemented based on Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Except otherwise noted, the reuse of this document is authorised under a Creative Commons Attribution 4.0 International (CC-BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

For any use or reproduction of elements that are not owned by the European Union, permission may need to be sought directly from the respective rightholders.

Cover page: © Lonely #46246900, ag visuell #16440826, Sean Gladwell #6018533, LwRedStorm #3348265, 2011; kras99 #43746830, 2012. Source: Fotolia.com.

# **Recommendations on certifying services required to enable FAIR within EOSC**

## ***Report from the EOSC Executive Board Working Group (WG) FAIR Task Force (TF)***

Edited by: Françoise Genova & Sarah Jones, Co-Chairs of the EOSC FAIR Working Group

January 2021

This report deals with certification of services enabling FAIR outputs. FAIR Metrics for EOSC are covered in the companion report (DOI: 10.2777/70791).

### Authors

Françoise Genova (Observatoire Astronomique de Strasbourg, ORCID: 0000-0002-6318-5028)

Jan Magnus Aronsen (University of Oslo, Norway, ORCID: 0000-0003-2593-1744)

Oya Beyan (Fraunhofer FIT, Germany, ORCID: 0000-0001-7611-3501)

Natalie Harrower (Digital Repository of Ireland, ORCID: 0000-0002-7487-4881)

András Holl (Hungarian Academy of Sciences, ORCID: 0000-0002-6873-3425)

Pedro Principe (University of Minho, ORCID: 0000-0002-8588-4196)

Ana Slavec (InnoRenew CoE, ORCID: 0000-0002-0171-2144)

Sarah Jones (GÉANT, ORCID: 0000-0002-5094-7126)

**Acknowledgements:** We are thankful for feedback and comments from Oscar Corcho, Esther Dzalé Yeumo, André Heughebaert, Rob Hooft, Hervé L'Hours, Wim Hugo, Rachael Kotarski, Mustapha Mokrane, Hylke Koers and Ilona von Stein.



## Contents

<b>EXECUTIVE SUMMARY .....</b>	<b>3</b>
1 INTRODUCTION .....	6
2 ACTIVITIES RELEVANT TO CERTIFICATION IN THE EOSC CONTEXT .....	7
2.1 Certification of data repositories.....	7
2.1.1 FAIRsFAIR development of FAIR certification schemes for repositories	7
2.1.2 CoreTrustSeal activities	12
2.1.3 Examples of certification and selection processes of data repositories in organisations: The World Data System, CLARIN and ELIXIR	12
2.1.4 The TRUST Principles for digital repositories: trustworthy repositories for preservation of FAIR digital objects	14
2.1.5 Characteristics derived from a Request for Public Comment from the USA Office of Science and Technology Policy	15
2.1.6 COAR Community Framework for Good Practices in Repositories	15
2.2 Certification of services enabling FAIR .....	16
2.2.1 FAIRsFAIR general principles and basic framework	16
2.2.2 Services to certify in priority	17
2.2.3 PID services	20
2.3 Registries of certified components .....	21
3 INCENTIVISATION AND SUPPORT .....	23
3.1 Incentivisation .....	23
3.2 Support .....	23
4 SUMMARY OF THE STATUS OF FAIR-ENABLING CERTIFICATION, GAPS AND POTENTIAL OPPORTUNITIES FOR EXTENSION, AND PRIORITIES FOR FUTURE WORK.....	26
4.1 Summary of the status of FAIR-enabling certification.....	26
4.2 Gaps and potential opportunities for extension.....	27
4.3 Priorities for future work .....	29
ANNEX: RELEVANT RECOMMENDATIONS FROM <i>TURNING FAIR INTO REALITY</i> .....	30

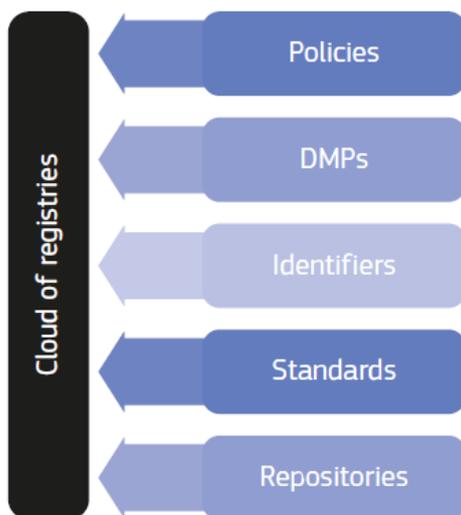
## EXECUTIVE SUMMARY

The report contains an analysis of activities relevant to certification of the services required to enable FAIR research outputs within EOSC as of November 2020. It discusses incentivisation and support, offers an analysis of gaps and potential opportunities for extension, and defines priorities for future work.

The FAIR Working Group of the EOSC Executive Board was initially tasked to define the certification approach that will be applied within EOSC for repositories that enable FAIR research outputs. We decided to expand the remit of our work to other services, because of the recognised need to define certification mechanisms for other elements of the FAIR ecosystem.

Work is on-going in FAIRsFAIR to develop FAIR certification schemes for repositories, through extension of the current “core” data repository certification with the identification of the levels of capability maturity necessary to support data which is assessed as FAIR. CoreTrustSeal is also working to extend their certification framework to additional groups of actors. Networks and projects such as the WDS and CLARIN use CoreTrustSeal in their accreditation processes, whereas ELIXIR defined its own processes to fit its requirements. Additional contributions include the TRUST principles, which provide an overarching framework for repository certification; the desirable characteristics of trustworthy repositories listed in a Request for Public Comments from the USA Office of Science and Technology Policies; and the useful framework for best practices in repositories published by the Confederation of Open Access Repositories in October 2020.

The essential components enabling the FAIR ecosystem are shown in the figure below (Figure 6 of *Turning FAIR into Reality - TFIR*). Critical building blocks that we rely on for operations should be certified. The assessment of certification of services enabling FAIR is progressing, in particular with FAIRsFAIR definition of general principles to be followed, and their on-going work on the definition of a basic framework. The services to certify in priority include PID services; the EOSC Architecture Working Group defined a set of subtopics for this certification process. Registries of certified components should be developed as shown in the figure.



*Turning FAIR into Reality* noted that activities are needed to incentivise and assist existing resources in their path towards certification. Incentivisation can be performed at different levels, through policies and projects in particular. Support can also be provided by different ways, and include support to improve data management practices.

Community concerns with metrics and certification were unequivocally expressed during the consultation on the EOSC Strategic Research and Innovation Agenda (SRIA) held

during the Summer of 2020, reflecting the need for inclusiveness and the uneven level of preparedness of services and communities. The transition period recommended by TFIR to allow existing repositories to go through the steps needed to achieve certification should be enabled. Certification like Metrics should not be a punitive method, and they should not be used for comparisons between repositories or disciplinary fields.

**At this stage, because of the need for inclusiveness and the different stages of preparedness of the communities and their services, certification status cannot be a necessary condition for a repository or other key components to be included in EOSC.**

**At some point, certification might become a prerequisite for inclusion in EOSC, in particular for data repositories but also for other key elements of the EOSC. This could be decided only after a careful assessment of the certification landscape and of the possible adverse consequences, such as excursion of valuable resources used by communities from the EOSC and putting these resources at risk.**

**We strongly recommend that repositories and services wanting to join EOSC use the certification framework criteria to check and improve their practices, with the aim to progress towards certification. Certified repositories should be clearly identified as such.**

We consider that CoreTrustSeal, which is a community-driven, international framework used by a large palette of disciplines, is the right level for research data repositories managed in the research environment with respect to DIN 31644 (nestorseal) and ISO 16363:2013.

The existing work on certification of the services required to enable FAIR should be extended under the next framework programme and ensure applicability across disciplines. One should not seek to define certification for all the types of services in the FAIR ecosystem. Priorities should be established on services to be certified.

The capability/maturity approach proposed in CoretrustSeal+FAIR should be extensively tested. All the certification frameworks proposed for other components of the FAIR ecosystem will also have to be extensively tested and feedback from a variety of stakeholders gathered.

*Turning FAIR into Reality* states that “concerted support is necessary to assist existing repositories in achieving certification.” The needs may include the construction and implementation of the necessary community standards to enable FAIR, as well as the building of the necessary skills and workforce in data management and FAIR implementation. Support for service self-assessment along certification criteria is needed to strengthen the ecosystem and ensure we can rely on the Web of FAIR data and services. Certifying a repository involves costs, depending on the repository starting status with respect to certification. Stakeholders including repository authorities and funders may understand the value for money of self-assessment and certification with respect to quality standards and trustworthiness demonstration, but in some cases the costs may be prohibitive with respect to resources. Without funding possibilities those countries, institutions or disciplines without appropriate resources might not be able to achieve certification, and the gap between the developed and less developed countries or disciplines will widen further.

More generally, support should be provided to certify services enabling FAIR once the specific certification framework is defined.

*Turning FAIR into Reality* states that “steps need to be taken to ensure that the organisations overseeing certification schemes are independent, trusted, sustainable and scalable.” Scalability is required to deal with the increase in the number of repositories seeking “core level” certification because of policy incentives. One element towards scalability may be to develop agreed community standards which could be machine-

evaluated as predefined components of the certification process. Support may have to be provided at some point to ensure scalability, but it has to take into account and preserve the necessary independence of the organisation. This support might be provided to enable certification of repositories from countries and disciplines with limited financial resources.

The governance and maintenance of this report and its recommendations should be done in close partnership with the user communities. A Working Group of the EOSC Stakeholder Forum should be established to perform the follow-up of this report, tasked to assess the evolution of the on-going activities and of the certification landscape, including the certification frameworks and the evolution of repository capacities with respect to certification.

The analysis of the status, gaps, and potential opportunities for extension, leads to define priorities for future work on the short term:

**Priority 1:** Support the current efforts to align Certification standards and assessment schemas with FAIR.

**Priority 2:** Test the proposed schemas in a variety of communities to gather feedback and update the proposed framework accordingly.

**Priority 3:** Provide support, methodologically as well as financially, to data and service providers to progress towards certification.

**Priority 4:** Monitor the progress of certification, assess the maturity of the certification landscape, and take appropriate action if fields or regions are lagging behind.

**Priority 5:** Support the establishment of core criteria and methodology to certify other key elements of the FAIR ecosystem, in particular in the first instance PID services and vocabulary repositories/metadata registries, and test them extensively.

**Priority 6:** Support the establishment and maintenance of registries of certified components of the ecosystem; if several registries are available for a given component, they should be harvestable and included in registries of registries.

**Priority 7:** Establish a Working Group under the EOSC Stakeholder Forum to ensure the implementation and further development of recommendations in this report.

### 1 INTRODUCTION

Trust is an essential component of Open Science,<sup>1</sup> and will also be essential for the success of EOSC. Certification is a path towards engendering trust. The EOSC FAIR Working Group (WG) is tasked to define the certification approach that will be applied within EOSC for repositories that enable FAIR research outputs. We decided to expand the remit of our work to services, because of the recognised need to define certification mechanisms for other elements of the FAIR ecosystem. The WG created a *Metrics and Certification Task Force* to coordinate this work alongside the companion activities on Metrics to certify FAIR digital objects.

Two of the priority recommendations of the Final Report and Action Plan from the European Expert Group on FAIR Data *Turning FAIR into Reality*<sup>2</sup> (2018) are relevant to this deliverable: one from the FAIR ecosystem chapter, **Develop assessment frameworks to certify FAIR services** (Recommendation 9), and one from the Incentives and Metrics for FAIR data and services, **Develop metrics to certify FAIR services** (Recommendation 13). The full text of the two recommendations is provided in the Annex.

Recommendation 13 is the main recommendation this report addresses. It deals with the need for certification schema to assess all components of the FAIR ecosystem, and of adapting existing frameworks, such as CoreTrustSeal,<sup>3</sup> rather than initiating new schemas (Action 13.1); the need to develop certification schemas for other components of the FAIR ecosystem (Action 13.2); the need for registries of certified components (Action 13.3); and the requirements on the organisations overseeing certification (Action 13.4). We are tasked to explore certification guidelines for repositories, but *Turning FAIR into Reality* clearly underlines the need to define certification mechanisms for other elements of the FAIR ecosystem, so we explored the additional elements of the FAIR ecosystem to be certified in line with Action 13.2 in addition to Action 13.1.

Recommendation 9 mainly deals with support and incentivisation to be provided to repositories to achieve certification, using CoreTrustSeal as a starting point (Action 9.1), and the need to allow for a transition period because certification and the preparation for certification take time (Action 9.2). The two other actions, on support programmes to be put in place when certification frameworks will have emerged for components other than repositories (Action 9.3), and mechanisms to be developed to ensure that the FAIR data ecosystem as a whole is fit for purpose (Action 9.4), are set for the longer term.

The remit of the EOSC FAIR WG is not to duplicate the relevant work done in the European and international contexts, in particular by initiatives and projects which are supported by the European Commission, but to bring this work together into a coherent direction. At the European level, the FAIRsFAIR<sup>4</sup> project - Fostering Fair Data Practices in Europe - is playing a specific role to enable a FAIR ecosystem in the EOSC. Certification is one of its main domains of activity. This report builds on their results to date. FAIRsFAIR started in March 2019, it produced its first milestones and deliverables in February 2020, and it will end in February 2022.

In this report, Section 2 deals with activities relevant to Recommendation 13, the development of certification frameworks for the components of the FAIR ecosystem. Section 3 discusses ongoing activities linked to Action 9.1, incentivisation and support to certification. Section 4 discusses the status of the certification developments, gaps and potential opportunities for extension, and priorities for future work.

---

<sup>1</sup> *Riding the Wave: How Europe can gain from the rising tide of scientific data*, Report of the EC High Level Group on Scientific Data, October 2010 [https://ec.europa.eu/eurostat/cros/content/riding-wave\\_en](https://ec.europa.eu/eurostat/cros/content/riding-wave_en)

<sup>2</sup> <https://doi.org/10.2777/1524>

<sup>3</sup> <https://www.coretrustseal.org/> CoreTrustSeal was created in 2017, on the basis of the merging of criteria and methodologies of two core certification frameworks, the Data Seal of Approval and the World Data System (WDS - <https://www.worlddatasystem.org/>).

<sup>4</sup> <https://www.fairsfair.eu/>

## 2 ACTIVITIES RELEVANT TO CERTIFICATION IN THE EOSC CONTEXT

The activities relevant to certification for repositories that enable FAIR research outputs, the initial remit of the FAIR Working Group, are discussed in Section 2.1. The more general question of certification of services enabling FAIR is discussed in Section 2.2.

The FAIRsFAIR project is one of the principal inputs for this report. The two chapters of this Section are respectively dealt with in two work packages, Work Package 4 *Certification (of repositories)*<sup>5</sup> and Work Package 2 *FAIR Practices: Semantics, Interoperability and Services*<sup>6</sup> respectively. They also work on pilots to define more rigorous requirements to support the assessment of FAIR data in trustworthy repositories as FAIR, an activity discussed in the companion *Recommendations on FAIR Metrics for EOSC*.

Section 2.1.1 discusses FAIRsFAIR activities relevant to certification guidelines for supporting FAIR research outputs. Their Work Package 4 *Certification (of repositories)* has among its objectives “to develop extensions to existing approaches to “core” data repository certification with a view to identifying the levels of capability maturity necessary to support data which is assessed as FAIR.” Section 2.1.2 describes CoreTrustSeal’s position with respect to FAIRsFAIR activities, and their own activities on the evolution of their certification framework with a possible expansion to a wider group of actors. Section 2.1.3 briefly discusses examples of accreditation processes set up by organisations to define membership in their data centre networks, the World Data System and the two ESFRI Landmarks CLARIN and ELIXIR. The development of the TRUST principles for trustworthy repositories for preservation of digital objects as an overarching framework is described in Section 2.1.4, the desirable characteristics of trustworthy repositories listed as a result of a recent Request for Public Comments from the USA Office of Science and Technology Policy is in Section 2.1.5, and the *Community Framework for Good Practices in Repositories* recently published by the Confederation of Open Access Repositories<sup>7</sup> (COAR) in Section 2.1.6.

Section 2.2.1 discusses the activities of Task 2.1 of FAIRsFAIR *FAIR Practices: Semantics, Interoperability and Services* on certification of “FAIR services.” The services to certify in priority were discussed during events organised by the FAIR Working Group: one session at the EOSC event co-located with the RDA 14<sup>th</sup> Plenary meeting in Helsinki, *The international community contributing to EOSC* (22 October 2019), and one at the EOSC Symposium in Budapest (26-28 November 2019). The findings are presented in Section 2.2.2. The discussion of certification of PID services by the EOSC FAIR and Architecture Working Groups are reported in Section 2.2.3. Registries have been identified as a key element of the FAIR ecosystem in the *Turning FAIR into Reality* report. More specifically, the need for registries of certified components is the subject of Actions 13.3 from *Turning FAIR into Reality*. This is addressed in Section 2.3.

### 2.1 Certification of data repositories

#### 2.1.1 FAIRsFAIR development of FAIR certification schemes for repositories

FAIRsFAIR WP4 Task 4.1 *Capability Maturity Model towards FAIR Certification* assesses Trusted Digital Repository requirements with respect to FAIR, taking as a starting point the CoreTrustSeal certification framework. The task aims at developing a practical and sustainable approach for repositories to self-assess their capability levels and identify target levels for “enabling” FAIR data. They align the characteristics of FAIR digital objects with the repositories that enable FAIRness, through the CoreTrustSeal Trustworthy Digital Repository requirements and the application of a capability/maturity model. CoreTrustSeal+FAIR will support the evaluation of Trustworthy Data Repositories, including their ability to offer an environment that enables and maintains FAIR data and metadata for the long term. This work is complemented by Task 4.5 work on the assessment of FAIR

---

<sup>5</sup> <https://www.fairsfair.eu/fair-certification>

<sup>6</sup> <https://www.fairsfair.eu/fair-practices-semantics-interoperability-and-services>

<sup>7</sup> <https://www.coar-repositories.org/>

data in trustworthy repositories, which is discussed in the companion report on FAIR Metrics.

The first deliverable of this task was D4.2 *Repository Certification Mechanism: A requirement on the extended requirements and procedures*<sup>8</sup> (May 2020), which has many component documents and was updated in M4.2 *Draft Maturity Model based on extensions and/or additions to CoreTrustSeal Requirements*<sup>9</sup> (August 2020). They will take into account feedback and testing, in particular by the repositories supported by FAIRsFAIR in their journey towards certification, with more extensive feedback sought.

The project has been analysing how Findable, Accessible, Interoperable and Reusable, the 15 FAIR guiding principles and the indicators developed by the RDA FAIR Data Maturity Model Working Group<sup>10</sup> map with the CoreTrustSeal requirements. A schema of the 16 CoreTrustSeal Requirements is shown in Figure 1.

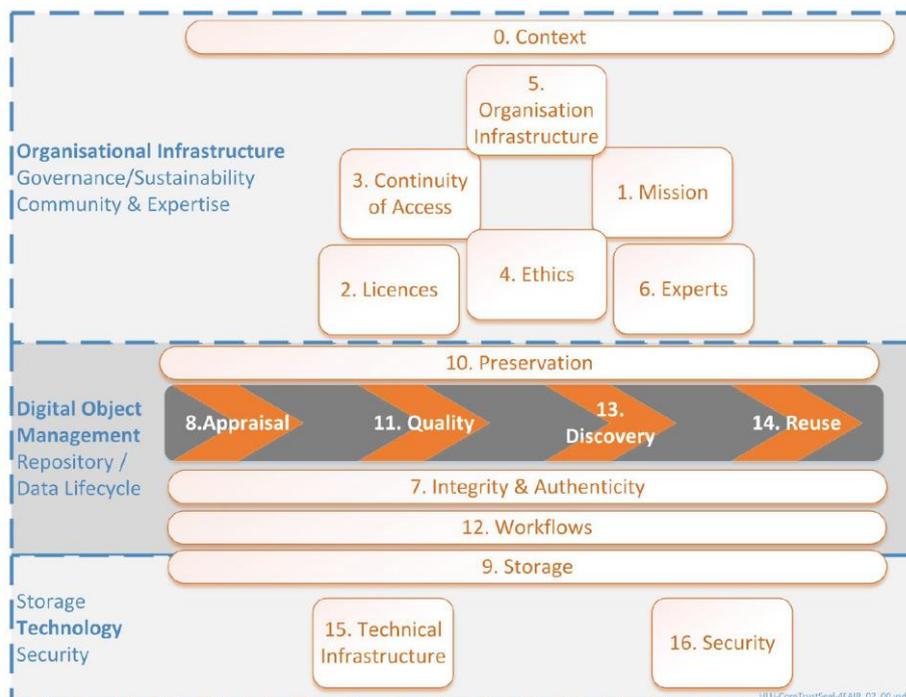


Figure 1: CoreTrustSeal Requirements in Brief (Diagram 4 of FAIRsFAIR M4.2)

They find that there are areas where the requirements can be aligned directly with repository capability, whereas in other cases a single mapping is not possible as there are multiple areas of repository activity that contribute to that aspect of FAIRness. Two successive versions of the *CoreTrustSeal+FAIR Overview*<sup>11</sup> have been published so far. The second one takes into account input from the ten repositories receiving support from FAIRsFAIR to achieve CoreTrustSeal certification and v0.90 of the FAIR indicators developed by the FAIR Data Maturity Model Working Group. The final version of *CoreTrustSeal+FAIR overview* will be integrated into a full CoreTrustSeal Requirement alignment. The FAIR to CoreTrustSeal Alignment evolved from the initial to the second version. The version available when this document is written is shown in Figure 2, in which the corresponding R numbers (e.g., R13, R15) in the left-hand column and in the bold-face text on the right refers to the associated CoreTrustSeal requirements as shown in Figure 1. The so-called “R0” criterion, which defines the context in which the repository operates, is not taken into account, but it could be used to document several aspects of the mapping, for instance by adding a question on “relevant standards.” One can note here that many of the CoreTrustSeal requirements are not directly relevant to the FAIR guiding principles,

8 <https://doi.org/10.5281/zenodo.3835698>

9 <https://doi.org/10.5281/zenodo.4003598>

10 <https://www.rd-alliance.org/group/fair-data-maturity-model-wg/outcomes/fair-data-maturity-model-specification-and-guidelines-0>

11 Initial version V01.00 <https://doi.org/10.5281/zenodo.3734897>, Second version V02.00 <https://doi.org/10.5281/zenodo.3862616>

which shows that many other important service organisation, management and operation aspects need to be covered for trustworthy services.

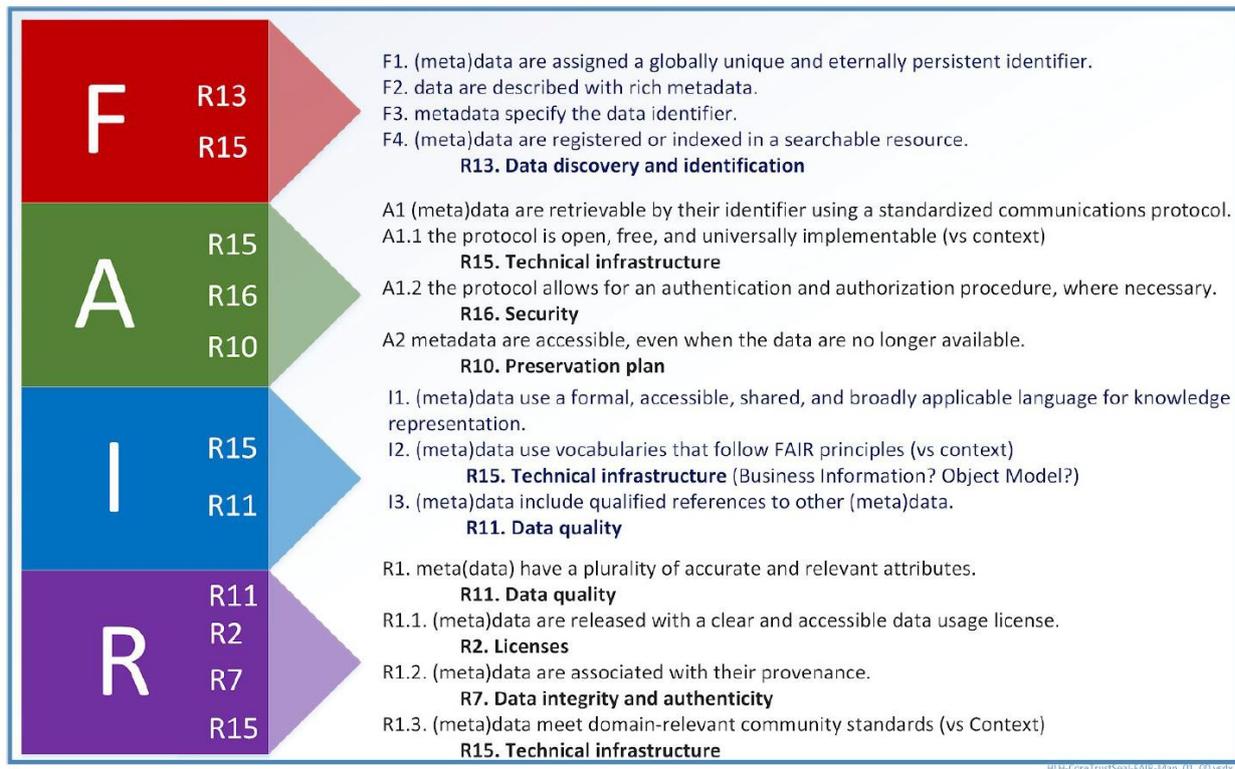


Figure 2: FAIR to CoreTrustSeal Alignment (Diagram 5 of FAIRsFAIR M4.2)

FAIRsFAIR will apply a capability and maturity approach to the CoreTrustSeal+FAIR alignment in complement to using the CoreTrustSeal compliance levels. FAIRsFAIR analysed the landscape of Capability Modelling in the CoreTrustSeal+FAIR context in a discussion paper published in May 2020.<sup>12</sup>M4.1 *Evaluation of current CoreTrustSeal guidelines and extended guidance to consider their implications for maturity modelling*<sup>13</sup> provides an overview of the capability and maturity levels and of their implications, and a discussion of the CoreTrustSeal compliance levels with respect to maturity thinking.

The CoreTrustSeal Self-Assessment Compliance Levels are the following:

- 0 - Not applicable
- 1 - The repository has not considered this yet
- 2 - The repository has a theoretical concept
- 3 - The repository is in the implementation phase
- 4 - The guideline has been fully implemented in the repository

The supporting guidance for the compliance levels states: "Compliance levels are an indicator of the applicant's self-assessed progress, but reviewers judge compliance against response statements and supporting evidence. If an applicant believes a Requirement is not applicable (0), then this must be justified in detail. Compliance Levels of 1 or 2 are not sufficient for a successful application. Certification may be granted if some Requirements are in the implementation phase (3)."

<sup>12</sup> <https://doi.org/10.5281/zenodo.3862588>

<sup>13</sup> <https://doi.org/10.5281/zenodo.3735030>

The tiered Capability Maturity Model Integration (CMMI) model 6 levels (0-5: Incomplete, Initial, Managed, Defined, Quantitatively Managed, Optimizing) are shown in Figure 3 and explained in Figure 4.

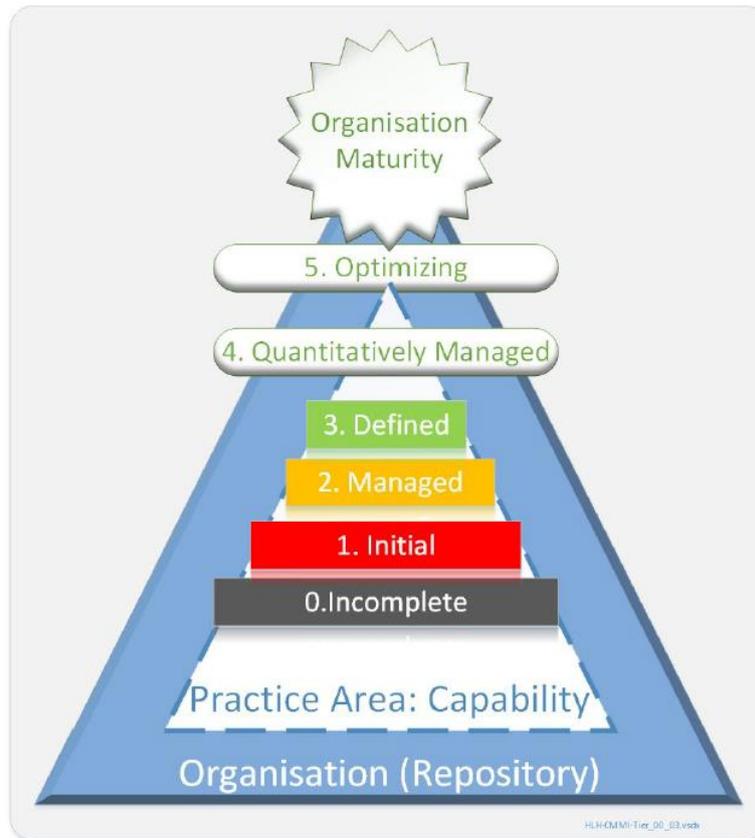


Figure 3: Tiered Capability/Maturity (Diagram 7 of FAIRsFAIR M4.2)

### Capability Levels

Capability levels apply to an organization's performance and process improvement achievements in individual practice areas. Within practice areas, the practices are organized into practice groups labelled Level 0 to Level 5 which provide an evolutionary path to performance improvement. Each level builds on the previous levels by adding new functionality or rigor resulting in increased capability.

### Maturity Levels

Maturity levels represent a staged path for an organization's performance and process improvement efforts based on predefined sets of practice areas. Within each maturity level, the predefined set of Practice Areas also provide a path to performance improvement. Each maturity level builds on the previous maturity levels by adding new functionality or rigor.

#### **Capability Level 0: Incomplete**

Incomplete approach to meeting the intent of the Practice Area.

May or may not be meeting the intent of any practice.

Inconsistent performance.

#### **Maturity Level 0: Incomplete**

Ad hoc and unknown. Work may or may not get completed.

Any self-assessments at level 0 will be reviewed and prioritised.

#### **Capability Level 1: Initial**

Initial approach to meeting the intent of the Practice Area.

Not a complete set of practices to meeting the full intent of the Practice Area.

Addresses performance issues.

#### **Maturity Level 1: Initial**

Unpredictable and reactive. Work gets completed but is often delayed and over budget.

Any self-assessments at level 1 will be reviewed and specific guidance developed.

### **Capability Level 2: Managed**

Subsumes level 1 practices.

Simple, but complete set of practices that address the full intent of the Practice Area.

Does not require the use of the organizational assets.

Identifies and monitors progress towards project performance objectives.

### **Maturity Level 2: Managed**

Managed on the project level. Projects are planned, performed, measured, and controlled.

### **Capability Level 3: Defined**

Builds on level 2 practices.

Uses organizational standards and tailoring to address project and work characteristics.

Projects use and contribute to organization assets.

Focuses on achieving both project and organizational performance objectives.

### **Maturity Level 3: Defined**

Proactive, rather than reactive. Organization-wide standards provide guidance across projects, programs, and portfolios.

### **Maturity Level 4: Quantitatively Managed**

Measured and controlled. Organization is data-driven with quantitative performance improvement objectives that are predictable and align to meet the needs of internal and external stakeholders.

### **Maturity Level 5: Optimizing**

Stable and flexible. Organization is focused on continuous improvement and is built to pivot and respond to opportunity and change. The organization's stability provides a platform for agility and innovation.

Figure 4: Capability and Maturity levels (from FAIRsFAIR M4.1)

Their assumption is that the capability levels of "Defined" (3) can deliver FAIRness, though they seek to identify any cases where level 2 "Managed" is sufficient. Maturity level 4 "Quantitatively managed" may be a dependency for sustainable complex partnership between data service providers. Achieving level 5 "Optimizing" would be desirable but resource-intensive. It is seen as valuable to support data services in defining where they need to focus resources on improvements.

They note that it seems inevitable that there will be a debate on what constitutes a level 3 maturity vs. level 5 and on what outcome is required for a given set of circumstances (e.g., 3 for low value, low cost/easy to recreate data, 5 for high value or sensitive data). They expect community expectations to evolve. They also note that they need to be sure the measurement/metric (e.g., CMMI scale) is appropriate to the object characteristics or repository features being analysed. The work takes place in parallel to efforts to test and "badge" individual digital objects as FAIR.

The approach to the tiers will be defined progressively through interactions with the FAIRsFAIR supported repositories and the wider community, including discussion with the CoreTrustSeal Board, in an iterative and evidence-based approach taking also into account the fact that metrics to evaluate FAIRness, and the tests against the metrics, are still in development.<sup>14</sup> "The overall goal is to develop a practical and sustainable approach for repositories to self-assess their current capability levels and identify target levels for enabling FAIR data. Integration of these processes into operational practice will provide a common approach to assessing and evaluating repository data services' ability to enable

---

<sup>14</sup> This is discussed in the companion report on Recommendations on FAIR Metrics for EOSC.

FAIR. The outcomes will be an overall improvement of repository practices and a pathway to certification.”

FAIRsFAIR M4.2 starts to look at standard approaches to extending the CoreTrustSeal, or elaborating additional requirements around that “Core” (this could be +FAIR or +Discipline). FAIRsFAIR will be looking at CoreTrustSeal and some alternate repository and service requirements by mapping them against a Generic Assessment & Evaluation Reference Model.<sup>15</sup>

### 2.1.2 CoreTrustSeal activities

FAIRsFAIR shares and discusses its recommendations with the CoreTrustSeal Board. D4.2 provides a statement of the CoreTrustSeal Board, stating that:

*“The mission for CoreTrustSeal endorsed by the Research Data Alliance and the wider community is to provide a single sustainable ‘core’ route for repository data service requirements and certification. The Board exists to manage and maintain that core route over time, and in response to community needs. As the FAIR Principles are clarified through indicators and evaluated through (ideally automated) tests against digital objects, CoreTrustSeal will continue to integrate ‘core’ best practices into the Requirements. We also recognise there may be more explicit FAIR requirements that may be elaborated around the foundation of the CoreTrustSeal. The CoreTrustSeal+FAIR work may be a case where we can integrate a FAIR-enabling assessment into the CoreTrustSeal process.”*

The CoreTrustSeal Board will continue to follow and engage in the work carried out by FAIRsFAIR and other FAIR-related initiatives around the world to ensure that CoreTrustSeal certification continues to address community needs for core-level certification.”

The CoreTrustSeal Board thus recognises the importance of surveying the evolution of FAIR assessment to eventually integrate more explicitly FAIR-related requirements. CoreTrustSeal requirements and processes are periodically reviewed. This would have to be included in the next periodic review of CoreTrustSeal requirements and processes, scheduled in 2022.

CoreTrustSeal is also assessing how to fulfil the demand for assessment, peer review and recognition from a wider group of actors delivering curation, storage and access services. This will be one of the aspects of the next formal review of CoreTrustSeal. The last review, which was the first since CoreTrustSeal started in 2017, was based on community comments and led to an update of the Guidance Documents in 2019, whereas the CoreTrustSeal criteria themselves remained unchanged. CoreTrustSeal recently sought community feedback<sup>16</sup> on criteria for “specialist repositories” serving a defined designated community with a clear knowledge base, which are the current main CoreTrustSeal constituents, “generalist repositories” serving a broader community, and “technical repository” service providers which could support bit-level preservation for either of the formers across the data life cycle.<sup>17</sup> They also indicate the possibility to implement immediate actions.

### 2.1.3 Examples of certification and selection processes of data repositories in organisations: The World Data System, CLARIN and ELIXIR

Three examples will be briefly described: two, the World Data System<sup>18</sup> and CLARIN,<sup>19</sup> which can be seen as an extension of CoreTrustSeal; and the specific process set up by ELIXIR<sup>20</sup> to select its Recommended Interoperability Resources<sup>21</sup> (RIR).

---

15 <https://doi.org/10.5281/zenodo.3733280>

16 <https://www.coretrustseal.org/why-certification/meeting-community-needs/>

17 <https://doi.org/10.5281/zenodo.3964071>

18 <https://www.worlddatasystem.org/>

19 <https://www.clarin.eu/>

20 <https://elixir-europe.org/>

21 <https://elixir-europe.org/platforms/interoperability/rirs>

### *The World Data System*

The World Data System is an Interdisciplinary Body of the International Science Council<sup>22</sup> (ISC; formerly ICSU) created by its 29<sup>th</sup> General Assembly in Maputo, Mozambique in 2008, to support the ISC's vision by promoting long-term stewardship of, and universal and equitable access to, quality-assured scientific data and data services, products, and information across all disciplines in the Natural and Social Sciences, and the Humanities. WDS aims to facilitate scientific research under the ISC umbrella by coordinating and supporting trustworthy scientific data services for the provision, use, and preservation of relevant datasets, while strengthening their links with the research community. The WDS has been building world-wide "communities of Excellence" for scientific data services by certifying member organisations. They started by defining their own certification framework, then joined with the Data Seal of Approval<sup>23</sup> (DSA), which was also an international "core" certification framework, in the RDA Working Group *Repository Audit and Certification on DSA-WDS partnership*.<sup>24</sup> The Working Group delivered in 2015 harmonised common procedures and requirements for the certification of data repositories at the core level, based on the DSA and WDS frameworks.<sup>25</sup> The DSA and WDS decided to align their certification process and criteria, which led to the creation of CoreTrustSeal as an independent entity in 2018. The Data Seal of Approval Board dissolved in favour of this successor standard and process.

The WDS continues to build its communities of excellence of data services. The current WDS membership process includes a dependency of achieving the CoreTrustSeal: to be accredited as a WDS Regular Member, an organisation must fulfil the membership requirements defined in the WDS Bylaws and undergo a CoreTrustSeal certification procedure.

### *CLARIN*

CLARIN is the European Research Infrastructure for Language Resources and Technology, and an ESFRI Landmark. CLARIN's distributed network is made out of centres.<sup>26</sup> There are several types of centres. The backbone of CLARIN is provided by technical centres, in particular *Service Providing Centres* or *CLARIN B-Centres*, for short. These units, often a university or an academic institute, offer the scientific community access to resources, services and knowledge on a sustainable basis. Having successfully completed CoreTrustSeal certification is a condition to become a CLARIN B-Centre.<sup>27</sup>

### *ELIXIR*

ELIXIR is an intergovernmental organisation, formally set up as a project in the European Molecular Biology Laboratory (EMBL) that brings together life science data resources from across Europe, and an ESFRI Landmark. ELIXIR established three lists of resources that it considers to be of special value to users of its infrastructure. First class are the ELIXIR Core Data Resources,<sup>28</sup> for which an original selection process was set up.<sup>29</sup> Second is the group of ELIXIR Deposition Databases:<sup>30</sup> a non-strict subset of the Core Data Resources that are suitable for direct deposition of data from research projects. The third class is formed by the ELIXIR Recommended Interoperability Resources (RIRs) which are tools and registries that have been selected by a panel of external reviewers. These resources build a network of foundational components underlying the ELIXIR Interoperability FAIR Service

---

22 <https://council.science/>

23 <https://www.coretrustseal.org/about/history/data-seal-of-approval-synopsis-2008-2018/>

24 <https://www.rd-alliance.org/groups/repository-audit-and-certification-dsa%E2%80%93wds-partnership-wg.html>

25 <https://www.rd-alliance.org/group/repository-audit-and-certification-dsa%E2%80%93wds-partnership-wg/outcomes/dsa-wds-partnership>

26 <https://www.clarin.eu/content/clarin-centres>

27 <https://www.clarin.eu/node/3577>

28 <https://elixir-europe.org/platforms/data/core-data-resources>

29 <https://f1000research.com/articles/5-2422>

30 <https://elixir-europe.org/platforms/data/elixir-deposition-databases>

Framework. The RIRs are identified as resources that facilitate the FAIR-supporting activities in scientific research, such as:

- Establish connections between data (and other) resources
- Acquire and expose metadata of data (and other) resources
- Create the infrastructure needed to build integratable data collections
- Use interoperability resources to support delivery of FAIR principles

ELIXIR maintains a selection process<sup>31</sup> and selection criteria<sup>32</sup> for RIRs. The criteria are seen as guidance for a minimum set of requirements, dealing with resource facilitation to scientific research, community, quality of resource, and legal framework, funding and governance.

ELIXIR thus built a specific framework to fulfil its requirements on selection of services to support the ELIXIR Interoperability Framework. They explain that the criteria focus on the process of FAIRisation and do not function to measure the FAIRness of the content.

*2.1.4 The TRUST Principles for digital repositories: trustworthy repositories for preservation of FAIR digital objects*

CoreTrustSeal<sup>33</sup> is only one element of the international repository certification landscape, providing a generic core-level framework for trustworthy repositories. Other generic international frameworks include ISO 16363:2013<sup>34</sup> (also known as CCSDS 652.0-M-1 - Audit and certification of trustworthy digital repositories) and DIN 31644<sup>35</sup> (nestorSeal). Discussions at the international level led to the definition of overarching principles for trustworthy digital repositories, the TRUST principles<sup>36</sup> (Lin, Crabtree, Dillo et al., 2020): **T**ransparency, **R**esponsibility, **U**ser Focus, **S**ustainability and **T**echnology (Figure 5).

Box 1 The TRUST Principles	
Principle	Guidance for repositories
Transparency	To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence.
Responsibility	To be responsible for ensuring the authenticity and integrity of data holdings and for the reliability and persistence of its service.
User Focus	To ensure that the data management norms and expectations of target user communities are met.
Sustainability	To sustain services and preserve data holdings for the long-term.
Technology	To provide infrastructure and capabilities to support secure, persistent, and reliable services.

Figure 5: The TRUST Principles (Box 1 of Lin, Crabtree, Dillo et al., 2020)

The paper concludes that “the TRUST Principles provide a mnemonic to remind data repository stakeholders of the need to develop and maintain the infrastructure to foster continuing stewardship of data and enable future use of their data holdings. The TRUST

31 [https://drive.google.com/file/d/1TDiOyNPBp-z61LkRYKjivV\\_Q0-acCcXc/view](https://drive.google.com/file/d/1TDiOyNPBp-z61LkRYKjivV_Q0-acCcXc/view)

32 <https://drive.google.com/file/d/1EDdOyJo0JKoy2eLY9Q5xcNM611aDI4Oc/view>

33 <https://www.coretrustseal.org/>, Requirements 2020-2022: [https://www.coretrustseal.org/wp-content/uploads/2019/11/Core\\_Trustworthy\\_Data\\_Repositories\\_Requirements\\_2020-2022\\_v00\\_02.pdf](https://www.coretrustseal.org/wp-content/uploads/2019/11/Core_Trustworthy_Data_Repositories_Requirements_2020-2022_v00_02.pdf)

34 The equivalent text can be obtained for free from the Consultative Committee for Space Data Systems (CCSDS) at <https://public.ccsds.org/Pubs/652x0m1.pdf>

35 DIN 31644, Ausgabe: 2012-04, Information and documentation - Criteria for trustworthy digital archives, 2014, <https://www.din.de/en/getting-involved/standards-committees/nid/standards/wdc-beuth:din21:147058907>

36 Lin, D., Crabtree, J., Dillo, I. et al. The TRUST Principles for digital repositories. *Sci Data* **7**, 144 (2020). <https://doi.org/10.1038/s41597-020-0486-7>

Principles, however, are not an end in themselves, rather a means to facilitate communication with all stakeholders, providing repositories with guidance to demonstrate transparency, responsibility, user focus, sustainability, and technology.” The TRUST principles offer the opportunity to bring different standards in a common framework.

### *2.1.5 Characteristics derived from a Request for Public Comment from the USA Office of Science and Technology Policy*

The USA Office of Science and Technology Policy (OSTP, Subcommittee on Open Science - SOS) released a request for public comment of draft desirable characteristics of repositories for managing and sharing data resulting from federally funded research<sup>37</sup> on 17 January 2020. It lists “proposed characteristics to help support discoverability, management, and sharing of research data, in a user-friendly manner, consistent with principles becoming widely adopted in the research community to make data findable, accessible, interoperable, and reusable (FAIR).” “The proposed characteristics are intended to be consistent with criteria that are increasingly used by non-Federal entities to certify data repositories, such as ISO16363 Standard for Trusted Digital Repositories and CoreTrustSeal Data Repositories Requirements, so that repositories with such certifications would generally exhibit these characteristics. SOS member agencies also anticipate that many repositories without such certifications would exhibit them as well.” The desirable characteristics for all data repositories are the following (details are provided when the meaning of the headlines might be ambiguous):

- Persistent unique identifiers
- Long term sustainability
- Metadata
- Curation and Quality Assurance
- Access
- Free & easy access to data and reuse
- Reuse (enables tracking of data reuse)
- Secure
- Privacy
- Common Format
- Provenance (maintains a detailed logfile of changes to datasets and metadata)

These are useful keywords for defining essential elements for FAIR-enabling data repositories. As explained, they are indeed well aligned with the existing certification frameworks.

### *2.1.6 COAR Community Framework for Good Practices in Repositories*

On 8 October 2020, the Confederation of Open Access Repositories (COAR) released a community framework<sup>38</sup> to assist repositories in evaluating and improving “their current operations based on a set of applicable and achievable good practices.” They aimed to create a global framework that could be adopted and used by different types of

---

<sup>37</sup> <https://www.federalregister.gov/documents/2020/01/17/2020-00689/request-for-public-comment-on-draft-desirable-characteristics-of-repositories-for-managing-and>

<sup>38</sup> <https://www.coar-repositories.org/coar-community-framework-for-good-practices-in-repositories/>

repositories, including publication and institutional repositories, as well as data repositories.

The community framework took a number of relevant assessment tools and frameworks into account.<sup>39</sup> COAR plans to enrich the framework through links to examples, guides and instructions, and will review the framework annually to ensure it reflects current practices.

The framework is organised by objectives, several of which are directly relevant to FAIR: Discovery, Access, Reuse, Integrity and authenticity, Quality assurance, Privacy of sensitive data, Preservation, and Sustainability and Governance, and notes a list of essential and desirable characteristics against each objective for a total of 45 characteristics. Overall, the COAR Community Framework for Good Practices in Repositories is designed to be inclusive, and to encompass the concerns of a wide range of repositories around the world. COAR currently has 157 members from 53 countries across all continents.

### 2.2 Certification of services enabling FAIR

#### 2.2.1 FAIRsFAIR general principles and basic framework

As explained, although our main task is to provide guidelines for repository certification, we want to explore possible certification of other elements required to enable FAIR research outputs in the FAIR ecosystem. In this ecosystem, digital objects should be FAIR at the point of consumption, which combines the initial FAIRness of the objects and the action of the services it went through on its FAIRness.

There is currently no well-defined assessment framework for FAIR enabling services. FAIRsFAIR Task 2.4 *FAIR Services* aims at closing the gap by making recommendations on how services fit in the FAIR ecosystem, providing an assessment framework with a checklist as a result. Their usage of “assessment framework” includes self-evaluation tools as well as more formal auditing and certification systems.

Their first production is M2.7 *Assessment report on 'FAIRness of services'*,<sup>40</sup> which presents a survey of existing FAIR assessment frameworks, a proposed set of guiding principles and desiderata for the FAIR assessment framework that will be constructed, and three “FAIR service assessment” case studies. Their work assesses “What does it take for a service to enable FAIR”, not “What does it mean for a service to be FAIR.”

FAIRsFAIR M2.7 provides a high-level set of requirements for frameworks, not requirements on the services themselves, which are:

- **Be comprehensive**, in that it applies to a broad range of functionalities across the data life cycle and across academic disciplines;
- **Be inclusive**, in that it addresses a wide array of service providers including commercial and public organizations;
- **Be rooted in FAIR data**, in that it clearly relates the FAIRness of a service to the FAIRness of the digital object that it acts on (thereby making an explicit connection to the original FAIR Data Principles);
- **Build upon existing work** as much as possible, for example extending concepts and criteria from frameworks such as CoreTrustSeal where possible;

---

<sup>39</sup> [Data Citation Roadmap for scholarly data repositories](https://www.nature.com/articles/s41597-019-0031-8) (<https://www.nature.com/articles/s41597-019-0031-8>), CoretrustSeal, The FAIR Data Principles, [PLOS “Criteria that Matter”](https://doi.org/10.26434/chemrxiv-2019-08-01), The TRUST Principles for Data Repositories,, COAR Next Generation Repository Technologies (<https://nrg.coar-repositories.org/>), [Plan S](https://www.coalition-s.org/addendum-to-the-coalition-s-guidance-on-the-implementation-of-plan-s/principles-and-implementation/) (<https://www.coalition-s.org/addendum-to-the-coalition-s-guidance-on-the-implementation-of-plan-s/principles-and-implementation/>).

<sup>40</sup> <https://doi.org/10.5281/zenodo.3688762>

- **Consider several dimensions of a service**, i.e., not only functional aspects (“utility” in FitSM terms) but also aspects that speak to quality, documentation, sustainability and trustworthiness (“warranty”) — where human factors including capacity building and training will be critical;
- **Be actionable** and aligned with the needs of the intended audience, in that parties developing or delivering data services can use it to, very practically, know what to put on their development roadmaps;
- **Be validated** by pilots and tests, in that the framework does not just live on paper but has been tested and practice — ideally with working exemplars; and
- **Be supported** by the community, in that it may count on informal support and formal endorsement by the broader community.

In addition, they define four modes to map services in terms of FAIR enablement, and use them in their case studies:

- **Enable**: the service actively helps to realize this particular FAIR principle - for example by adding metadata or enabling discoverability;
- **Respect**: the service does not particularly enable this particular FAIR principle, but also does not interfere with it - it can be said to respect the “FAIR-in-FAIR-out” principle;
- **Reduce**: the service actually makes data less FAIR - at least for a particular principle - for example by detaching metadata or a PID when it acts on a digital object;
- **N/A** (not clear or non-applicable): This particular element is not relevant for the service, or there was insufficient information to determine if the FAIR principle applies.

It might be useful to separate unclear cases from the ones which are not relevant for the service in the N/A category.

The team assesses the numerous recent developments of FAIR metrics, service assessment and interoperability. In M2.10 *Report on basic framework on FAIRness of services*<sup>41</sup> they propose a basic framework for the FAIR assessment of services aimed at service providers, constructed on the basis of three streams of input: a review of the literature, input from stakeholders gathered at a session of the EOSC-hub week in May 2020,<sup>42</sup> and a series of interviews with service owners. The proposed framework is organised along six lines: FAIR enablement; quality of service; open and connected, user centricity; trustworthiness; ethical and legal.

The proposed framework will be updated iteratively through public consultations and workshops. The final result, a framework for assessing FAIR enabling services (FAIRsFAIR deliverable D7.2), will be published in July-August 2021.

### 2.2.2 Services to certify in priority

FAIRsFAIR has established a pan-project Synchronisation Force<sup>43</sup> which liaises with its “European Group of FAIR Champions”<sup>44</sup>, the five ESFRI Clusters and the so-called ‘5b’ projects, the thematic and regional EOSC projects. In the framework of the Collaboration Agreement between FAIRsFAIR and EOSC Secretariat, the Synchronisation Force provides input for the EOSC Executive Board Working Groups, including the FAIR Working Group.

---

41 <https://doi.org/10.5281/zenodo.4292599>

42 <https://www.eosc-hub.eu/eosc-hub-week-2020/agenda/fair-assessment-certification-repositories>

43 <https://fairsfair.eu/advisory-board/synchronisation-force>

44 <https://www.fairsfair.eu/advisory-board/eqfc>

The second FAIRsFAIR Synchronisation Force Workshop was organised on-line as a series of eight sessions from April 29th to June 11th, 2020. Representatives from the EOSC Executive Board Working Groups, from the EOSC Clusters and '5b' projects, and the members of the European Group of FAIR Champions, were invited to attend. The EOSC FAIR Working Group participated actively in the workshop. The workshop objectives were to measure the progress towards implementing the recommendations outlined in *Turning FAIR into Reality*, and also to identify gaps in its Action Plan and propose additional actions. The workshop report<sup>45</sup> summarises the findings. TFiR Recommendation 9 *Develop assessment frameworks to certify FAIR services* was examined. It is proposed to add the following element, which could appear as Action 9.2bis in TFiR (see Annex 1): *Prepare a priority list of services that would benefit from FAIR assessment and certification*. They state that any such statement should clearly articulate the purpose and need for such assessment and propose draft criteria.

The EOSC FAIR WG organised a session at the EOSC event co-located with the RDA 14<sup>th</sup> Plenary meeting in Helsinki, *The international community contributing to EOSC*<sup>46</sup> (22 October 2019), in which one of the questions was about the elements of the FAIR ecosystem which should be certified. The questions were tackled by groups gathering the session participants by thematic field, including one for people involved in interdisciplinary work. The answers helped us to shape a session on *FAIR Service Certification* at the EOSC Symposium<sup>47</sup> (Budapest, 26-28 November 2019) including lightning talks and a facilitated discussion.

The discussion during the Helsinki EOSC event initially veered towards the potential for certification across most elements of the FAIR ecosystem, which is very ambitious, and unrealistic. A cautionary note was sounded about over-certification: the system-wide certification of most elements would impose significant overhead, both to define the adequate certification criteria for many different elements intervening at many different levels, and on providers which would have to go through heavy certification processes. In addition, these processes would have to be set up and managed.

Further discussion in Helsinki showed that for many aspects the point would be rather to have frameworks to share good practices rather than to impose formal certification. The metrics being or to be defined for different elements of the FAIR ecosystems including software, as discussed in the companion report on FAIR metrics for EOSC, will be an instrument to enable the identification of possible progress and the evaluation of progress made by making elements of good practices explicit. One can note here that CoreTrustSeal and its predecessors arose from communities seeking to define good practices and in addition give credit when it was due: community-driven definition of good practices and the evaluation of the need for formal certification is central to the process.

We built on the Helsinki meeting to poll the audience of the FAIR service session in Budapest for their input on the kind of services which should be certified, which again produced a widely diverse list. The second question on the criteria which should guide the decisions around what needs formal certification vs. sharing of good practices, also produced diverse answers. The main keywords that arose are points of risk (in which we should include components for which it is hard or impossible to use an alternative once a provider is chosen), dependencies, usage by community. Some note that maturity development is the key, and propose transparency, sharing of good practices and onboarding rather than certification, or at least to allow for different levels of certification.

In spite of this diversity, when asked for the priorities on the kinds of FAIR services which should be certified, from a predefined list of core components of the FAIR ecosystem

---

45 Second Report of the FAIRsFAIR Synchronisation Force (D5.5) <https://doi.org/10.5281/zenodo.3953979>

46 The International Research Data Community contributing to EOSC <https://www.eoscsecretariat.eu/international-research-data-community-contributing-eosc>

47 *FAIR Service Certification* (<https://www.eoscsecretariat.eu/eosc-symposium2019/FAIR-service-certification>), chaired by Françoise Genova and Pedro Principe, FAIR WG

inspired from *Turning FAIR into Reality*, repositories came first, and some cited PID systems and registries - identified as metadata registries. The attendants were also asked whether they were aware of ongoing activities to define certification schema. The outcome of the session is summarised in Table 1.

What kind of services should be certified?	What criteria should guide our decisions around what needs formal certification vs. sharing of good practices?	Are there any existing frameworks towards this?
<ul style="list-style-type: none"> <li>• Repositories</li> <li>• Template/Tools to prepare DMPs</li> <li>• PID services</li> <li>• Datasets</li> <li>• Software</li> <li>• Long term data services</li> <li>• Registries</li> <li>• Terminology services</li> <li>• APIs</li> <li>• All services supporting research data management</li> </ul>	<ul style="list-style-type: none"> <li>• Maturity</li> <li>• Sustainability</li> <li>• Persistency</li> <li>• Availability</li> <li>• Risks: data integrity, provenance, management</li> <li>• Machine readability</li> <li>• Support</li> <li>• Centrality</li> <li>• Dependencies</li> <li>• Better to share good practices.</li> </ul>	<ul style="list-style-type: none"> <li>• RDA recommendations</li> <li>• CoreTrustSeal</li> <li>• ISO16363</li> <li>• NESTOR for data repositories</li> <li>• FitSM standard</li> <li>• Proposed SHARP principles</li> <li>• ELIXIR evaluation badges and process</li> </ul>

Table 1: FAIR Service session, EOSC Symposium, Budapest - Summary of participant feedback

Figure 6 (Figure 6 from *Turning FAIR into Reality*) shows the essential components of the FAIR ecosystem. An EOSC service that a user chooses once and then uses sustainably (like a repository) requires trust. A service that can be tested and rejected when better competition arises (like a search engine) does not require certification. Only critical building blocks that we rely on for core operation and long-term preservation should be certified. The figure shows that registries are critical components of the system, and as such are candidates for the definition of a specification framework. The relevant reports of the EOSC Executive Board Architecture and FAIR Working Groups point at PID services and vocabulary repositories/metadata registries as priorities for defining a certification framework.



Figure 6: The components of the FAIR ecosystem (Figure 6 of Turning FAIR into Reality)

### 2.2.3 PID services

The FAIR and Architecture Working Groups of the EOSC Executive Board worked together on a *Persistent Identifier (PID) Policy for the European Open Science Cloud*.<sup>48</sup> The document discussed requirements on PID services and PID providers, which explicitly mention trustworthiness: "A set of trusted registration PID Authorities and PID Service Providers is needed that are regularly certified based on agreed rule sets. Certification should cover both resolvability of PIDs to information from PID Service Providers and their management processes for maintenance of PIDs. It should clarify who is responsible for keeping the Kernel Information up-to-date, if necessary, by enabling third parties to modify it."

The PID policy implementation will be guided through recommendations on the *PID Technical Architecture for EOSC*<sup>49</sup> provided by the EOSC Executive Board Architecture Working Group. The document has a section on certification. They state that PID services need a special level of trustworthiness, and that certification is a possibility to raise the level of trust in these services; that PID registration authorities and PID service providers have to be certified by independent agencies - this would be not only a control of technical processes, but also operational and governance aspects are evaluated. They insist on the fact that the persistence expectation for PIDs raises a special requirement that goes beyond institutional borders, because an institutional failure in providing the service needs a fallback by other institutions, for short institutional failures as well as for a permanent shutdown of the service for whatever reason. They plead that this kind of trust can only be reached by special contracts made between service providers that have themselves already a certain level of trustworthiness and institutional persistence promise - this is certainly a more general point because of the wide range of co-dependencies to be expected across EOSC.

They list subtopics for the certification process, such as:

- Quality assurance (stability and performance)
  - public service level agreements,
  - organisational measures taken to ensure the persistency of the service,
  - PID stability and consistency checks including contractor notifications in case of errors,
  - redundancy of PIDs to guarantee services according to the EOSC PID Policy,
  - measurements to prevent information loss in case of crashes (backup, mirroring),
  - a guarded PID deletion strategy, with a process to create a tombstone note in case that a Digital Object was deleted.
- Long term persistence
  - With which temporal time frame (long-term funding statements and business model)?
  - an exit strategy in case of ceasing a service (with cross institutional contracts about service continuation)
- Security
  - ensure that no unauthorised changes are possible
- Support

---

<sup>48</sup> [https://ec.europa.eu/info/publications/persistent-identifier-pid-policy-european-open-science-cloud\\_en](https://ec.europa.eu/info/publications/persistent-identifier-pid-policy-european-open-science-cloud_en) DOI: 10.2777/926037

<sup>49</sup> Draft for consultation: <https://docs.google.com/document/d/1T-bpNsmuxQewsLq48XTyUJJoe0lsV7poaXohpgDo9W34/edit>

- What if there are complaints against resolution or generation of PIDs?
- Certification process
  - What if there are complaints against the result of evaluation (like not certified)?

Since persistence is a promise with PIDs this aspect must have a special priority in the certification process. It can be supported by technical processes, but as said is more dependent on social contracts that have been agreed upon which is difficult to assess. In addition, compliance to the described PID architecture can also be a topic in a certification process.

### 2.3 Registries of certified components

*Turning FAIR into Reality* identifies registries as a key component of the FAIR ecosystem. Recommendation 3 states that there need to be registries cataloguing each essential component of the ecosystem, and automated workflows between them. Figure 6 above summarises the components of the FAIR ecosystem with the cloud of registries cataloguing them.

The FAIRsFAIR WP4 Discussion Document *FAIR Principles Baseline Comments*<sup>50</sup> lists challenges about the FAIR guiding principles and the areas of clarification that are necessary beyond the FAIR principles to define indicators and metrics, particularly as they apply to the context of Trusted Digital Repositories as the enabling environment for FAIR data. From the analysis, the following registries are required for the implementation of the different guiding principles - some of them are explicitly cited in the text, others can be deduced from the comments:

- Persistent identifier registry
- Registry of resource discovery systems
- Registry of communication protocols
- Registry of standards (for data and metadata)

In fact, in most cases there will likely not be a single registry for a given type of component. When there are multiple registries, “registries of registries” will have to be established, and registries should have harvesting capabilities.

Re3data<sup>51</sup> and FAIRsharing<sup>52</sup> are both cited in *Turning FAIR into Reality* Action 13.3 as examples of registries gathering information on data repositories. They are both in liaison with FAIRsFAIR.

Re3data is a DataCite<sup>53</sup> service which provides a global registry of research data repositories for permanent storage and access to datasets from a diverse range of academic disciplines. It gives an overview of the data repository landscape, helps researchers to find appropriate repositories for the storage and access of data sets, and provides information about the repositories, in particular whether they are certified and if yes, in which certification framework. FAIRsFAIR Task 4.4 *Tools to identify relevant trustworthy certified repositories* works to improve the re3data registry to enable researchers to identify FAIR-enabling repositories, and to find relevant datasets or deposit their research data. The task takes input from multiple efforts within FAIRsFAIR to identify relevant re3data metadata or improve them and develop functionality to provide discoverability and access.

---

50 <https://doi.org/10.5281/zenodo.3728131>

51 <http://re3data.org/>

52 <https://fairsharing.org/>

53 <https://datacite.org/>

FAIRsharing, an RDA Recommendation, is a resource based at the University of Oxford, UK with an international user base. It provides manually curated metadata records on data and metadata standards, databases, repositories, knowledge bases and journal and funder data policies, as well as the relationships between them. FAIRsharing and FAIRsFAIR set up an agreement in July 2020 to collaborate on areas of common interest and interconnect their activities.<sup>54</sup> FAIRsharing indicates that they “will display certifications and interoperability features on the relevant records.”

---

<sup>54</sup> <https://www.fairsfair.eu/news/fairsharing-fairsfair-join-forces-support-repositories-all-around-europe-their-effort-towards>

### 3 INCENTIVISATION AND SUPPORT

*Turning FAIR into Reality* Action 9.1 states that “a programme of activity is required to incentivise and assist existing domain repositories, institutional services and other valued community resources to achieve certification, in particular through CoreTrustSeal.” Incentivisation and support can come through different paths. Incentivisation often comes through policies at different levels -- national, funders, publishers, projects. Support can also be provided by different pathways, such as direct support to repositories undergoing certification (e.g., within projects or at the national level), as well as guidance and best practices.

#### 3.1 Incentivisation

Incentivisation for repositories to undergo certification comes mainly through national, funders’ and publishers’ policies, but can also come from organisations and projects, as shown by the examples discussed in Section 2.2.3.

At the national level, for example, the French National Plan for Open Science has among its measures “implement a certification process for data infrastructures.”<sup>55</sup> The Dutch National Plan for Open Science notes that long term storage and reuse of research data relies on defining “the characteristics of an approved data repository, using all the certifications and initiatives for this purpose already in place at the international level.”<sup>56</sup> More recently, Ireland’s National Framework on the Transition to an Open Research Environment notes “The quality and reliability of the infrastructure will be ensured, including through the use of internationally recognised certification mechanisms, specifications, and standards...”<sup>57</sup> Similarly, Austria’s national policy explicitly recommends the use of certified Trusted Digital Repositories.<sup>58</sup>

Depositing data in a certified repository is also seen as a good point in Data Management Plans and in some publishers’ policies. Science Europe states, in their *Practical Guide to the International Alignment of Research Data Management*<sup>59</sup> that “some repositories have been certified as trustworthy repositories by one of several acknowledged certification bodies. In order to facilitate the recognition of trustworthy repositories for researchers, it is strongly recommended that repositories that have not yet been certified seek certification by such a body.”

As discussed in Section 2.2.3, organisations and projects can require that their members undergo certification, CoreTrustSeal in the case of the WDS and CLARIN, their own framework for ELIXIR. The CESSDA ERIC conditions of membership are defined in the “Annex 2 Obligations” which include a reference to certification against a trustworthy digital repository standard, which has been specified to mean CoreTrustSeal. At the international level, for instance, the South African Ilifu states on their website<sup>60</sup> that “Repository instances across the consortium are expected to have CTS certification.”

#### 3.2 Support

Projects can play a role in providing support to certification. FAIRsFAIR has two relevant tasks in its Work Package 4: Task 4.2 *European Network of trustworthy repositories enabling FAIR data* and Task 4.3 *Support and guidance for certification of data repositories*. Calls offering support for certification and interoperability allowed the project to identify repositories at different maturity levels. The certification of the 10 selected repositories will provide feedback on certification process and FAIR alignment. Support and guidance are also provided for the repositories which applied but were not selected.

---

55 [http://cache.media.enseignementsup-recherche.gouv.fr/file/Recherche/50/1/ SO\\_A4\\_2018\\_EN\\_01\\_leger\\_982501.pdf](http://cache.media.enseignementsup-recherche.gouv.fr/file/Recherche/50/1/ SO_A4_2018_EN_01_leger_982501.pdf)

56 [https://www.openscience.nl/files/openscience/2019-02/nationalplanopenscience\\_en.pdf](https://www.openscience.nl/files/openscience/2019-02/nationalplanopenscience_en.pdf)

57 [http://norf-ireland.net/wp-content/uploads/2019/07/NORF\\_Framework\\_10\\_July\\_2019-2.pdf](http://norf-ireland.net/wp-content/uploads/2019/07/NORF_Framework_10_July_2019-2.pdf)

58 <https://www.fwf.ac.at/en/research-funding/open-access-policy/open-access-to-research-data/>

59 [https://www.scienceurope.org/media/jezkhnoo/se\\_rdm\\_practical\\_guide\\_final.pdf](https://www.scienceurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf)

60 <http://www.ilifu.ac.za/il/reports-and-publications>

Support can be provided at the research community level, with the advantage of knowing the community culture and practices. The SSHOC Cluster provides an overview<sup>61</sup> of certification approaches in CESSDA ERIC, CLARIN ERIC, DARIAH ERIC and E-RIHS communities. They opened a call for application<sup>62</sup> from Social Science and Humanities repositories interested in receiving support in achieving CoreTrustSeal certification.

Support to improving repository practices, with the aim to bring at least some of them to certification, can also be implemented at the national level. The Research Data Alliance National Nodes set up by the RDA Europe 4.0 project in March 2018 can play a role in incentivisation and support, as well as the Clusters and the regional/thematic "EOSC 5b" project (one can cite for instance EOSC-Nordic). Certification is seen as important by a number of RDA Nodes, some of those, such as the Hungarian node, having their first emphasis on making repositories aware of the FAIR aspects, and helping them to comply. RDA Europe 4.0 created a group of nodes interested in developing certification through CoreTrustSeal in their countries, which allowed the sharing of good practices on repository on-boarding and support.

The RDA National Nodes from France and the Netherlands have repository certification as a priority, and provide supporting activities. In France moreover, a common working Group on Certification of data repositories was set up in July 2020 by the national Open Science Committee and RDA France to define and perform supporting activities, with funding from the National Open Science Fund. In Canada, the non-profit organization Portage<sup>63</sup> is planning a repository certification support project to begin in early 2021 and aimed at supporting a cohort of repositories in obtaining or progressing toward CoreTrustSeal certification. The project will aim to provide training, facilitate information exchange among participants, and provide direct support to enable the repositories' self-assessments.

Support can be provided more generally to repositories for improving their data management practices and progress in the implementation of the FAIR principles. Cooperative comparisons and exchange between repositories with similar contexts (similar type, similar discipline, similar organisational structure, etc) can aid mutual learning and best practice development.

The companion report on FAIR Metrics for EOSC underlines the risks of using automated tools to measure compliance before they are fully tested in a variety of contexts and their biases fully understood. When such tools will be deemed acceptable, they would provide the repositories with elements supporting their FAIR enabling practices and their certification self-assessment and facilitate the work of the certification bodies.

At the technical level, most repositories are built on widely used software. It would be very helpful to make explicit which requirements of the certification these software fulfil by construction, so the repositories providers can focus on the non-technical aspects of the certification. Being able to rely on these capacities is a way to spread FAIR culture through the usage of tools enabling FAIR. FAIR-enabling should now be kept in mind in the development of software platforms by community effort.

The TRUST principles can be used for communication in particular towards repositories, to raise awareness and as a first step for analysing repository practices. The OSTP criteria can be used as a second step. The CoreTrustSeal criteria can also be used as a guideline. At the international level, in the *Enabling FAIR Data* project,<sup>64</sup> a repository cohort in Earth, Space and Environmental Sciences (ESES) has been established in partnership with CoreTrustSeal and the World Data System and supported by the Council of Data Facilities to advance the implementation of FAIR principles in ESES repositories. Similarly, the

---

61 Kleemola, Mari, Alaterä, Tuomas J., Koski, Niko, Ala-Lahti, Henri, Jerlehag, Birger, L'Hours, Hervé, ... Van Horik, René. (2020). SSHOC D8.2 Certification plan for SSHOC repositories (Version v1.0). Zenodo. <https://doi.org/10.5281/zenodo.3725867>

62 <https://sshopencloud.eu/news/call-applications-sshoc-repository-certification-support-0>

63 <https://portagenetwork.ca/>

64 <https://eos.org/agu-news/advancing-fair-data-in-earth-space-and-environmental-science>

*Community Framework for Good Practices in Repositories* released by COAR provides a strong pathway to a range of repositories, and is deeply focused on member/community support. Many of the criteria in the framework align with FAIR.

## 4 SUMMARY OF THE STATUS OF FAIR-ENABLING CERTIFICATION, GAPS AND POTENTIAL OPPORTUNITIES FOR EXTENSION, AND PRIORITIES FOR FUTURE WORK

### 4.1 Summary of the status of FAIR-enabling certification

As stated in the “Turning FAIR into Reality” action plan, there is a need for certification schemas to assess all components of the FAIR ecosystem. Significant work has been devoted to certification of data repositories, with an international landscape which includes in particular CoreTrustSeal, which provides a generic core framework for trustworthy repositories, DIN 31644 (nestorSeal) and ISO 16363:2013.

In the context of FAIR, work is on-going, in particular in the FAIRsFAIR project, on FAIR alignment of repository certification schemas. This is complementary to the evaluation of the FAIRness of the data itself. More generally, the certification of FAIR-enabling services is also studied - a service can enable, respect or reduce the FAIRness of its holdings.

The EOSC will be a federation of existing resources. It will of course give access to new data, but it will primarily be a federation of existing thematic data repositories and services, interfaced with existing data sharing frameworks. Inclusiveness is thus a critical keyword for success: repositories and services should be incentivised and supported to improve their data management practices and in their journey towards FAIR, to provide and get the maximum benefit from EOSC.

Another critical keyword for EOSC success is trust. Certification is a way to display trustworthiness, evaluated by external evaluators with respect to a set of criteria. The criteria also provide a framework for the improvement of practices, and are thus an important tool for increasing quality. They can be used by repositories to self-evaluate and improve their practices and processes, even if they do not candidate for formal certification. The definition of compliance levels, as in CoreTrustSeal, enables to evaluate the repository status with respect to trustworthiness and to measure progress. The FAIR guiding principles are already present in the certification frameworks, in general implicitly, and to better include them is a way to enable repositories in their journey towards FAIR.<sup>65</sup>

One of the recommendations of the companion report on FAIR Metrics for EOSC, namely “do not reinvent the wheel” is replicated here. *Turning FAIR into Reality* states that “certification schemas are needed to assess all components of the ecosystem as FAIR services.” It recommends that “when existing frameworks exist to certify data services, these should be reviewed and adjusted to align with FAIR,” and cites CoreTrustSeal as an existing framework to be adapted.

CoreTrustSeal is a community-based initiative launched in 2017 which currently gathers a growing international community of repositories which cover a wide palette of disciplinary fields.<sup>66</sup> It is adopted by communities and projects as an appropriate mechanism to certify trustworthiness. FAIRsFAIR is working on a CoreTrustSeal+FAIR framework, with the aim of applying a capability and maturity approach to the CoreTrustSeal+FAIR alignment in complement to using the CoreTrustSeal compliance levels.

In parallel, ELIXIR maintains a specific framework to fulfil its requirements on selection of services to support the ELIXIR Interoperability Framework.

Certification frameworks should be regularly assessed to verify that they remain fit for purpose, gather community feedback and update them if needed. CoreTrustSeal has a regular 3-year assessment and revision framework, which rely on feedback and suggestions from its extended community and can also take advantage of the RDA/WDS

---

<sup>65</sup> See for example: [https://ipres2019.org/static/pdf/iPres2019\\_paper\\_74.pdf](https://ipres2019.org/static/pdf/iPres2019_paper_74.pdf)

<sup>66</sup> <https://www.coretrustseal.org/why-certification/certified-repositories/>

Certification of Digital Repositories Interest Group<sup>67</sup> to reach the wider community interested in certification of trustworthy repositories. The first assessment was completed in 2019; community feedback led to an update of the guidance, whereas the criteria were not changed. CoreTrustSeal began its next assessment cycle by polling the community to assess the demand for extending its framework to better include actors delivering curation, storage and access services. They also follow and engage in the work performed by FAIRsFAIR on alignment of CoreTrustSeal with FAIR.

Research artefacts are made FAIR by the services in which they are created, discovered and reused: FAIR data maturity depends on the capabilities and trustworthiness of services such as repositories, but not only. The different kinds of services making up the FAIR ecosystem are at very different stages of definition of a certification framework. FAIRsFAIR is working on the definition of certification of services enabling FAIR.

PIDs are key enablers of FAIR; the Architecture Working Group of the EOSC Executive Board underlines that PID services need a special level of trustworthiness and lists subtopics to take into account in the certification process.

### 4.2 Gaps and potential opportunities for extension

Alignment of repository certification schemas with FAIR and the definition of a framework for certification of services enabling FAIR are underway but need to be further developed and tested. FAIRsFAIR second Synchronisation Workshop pinpoints the need to add the action to “prepare a priority list of services that would benefit from FAIR assessment and certification” to Recommendation 9 of *Turning FAIR into Reality*. The assessment of PID service certification initiated by the EOSC Architecture Working Group should be developed and tested.

*Turning FAIR into Reality* states that “a transition period is needed to allow existing repositories without certification to go through the steps needed to achieve trustworthy digital repository status.” The community unequivocally expressed its concerns with metrics and certification during the consultation on the EOSC SRIA held during the Summer of 2020: Metrics and Certification are given a low priority in the survey of relevance of action areas, ranking only second-to-last with 39% of votes compared to 78% for the highest ranked priority, metadata and ontologies, followed by identifiers at 72%. These concerns are reflecting the need for inclusiveness and the uneven level of preparedness of the services and communities. They have to be taken into account, and the transition period advocated by TFIR has to be enabled.

Certification like Metrics should not be a punitive method but be implemented progressively and incrementally, to enable gradual and ultimately wide by-in. They should not be used for comparisons between repositories or disciplinary fields. It is worth noting also that there is a dependency between the achievement of the identified priorities and the capacity to enable data FAIRness. Effective metadata/ontology and identifier provision are critical for repository trustworthiness.

**At this stage, because of the need for inclusiveness and the different stages of preparedness of the communities and their services, certification status cannot be a necessary condition for a repository or other key components to be included in EOSC.**

**At some point, certification might become a prerequisite for inclusion in EOSC, in particular for data repositories but also for other key elements of the EOSC. This could be decided only after a careful assessment of the certification landscape and of the possible adverse consequences, such as exclusion of valuable**

---

67 <https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html>

**resources used by communities from the EOSC and putting these resources at risk.**

**We strongly recommend that repositories and services wanting to join EOSC use the certification framework criteria to check and improve their practices, with the aim to progress towards certification. Certified repositories should be clearly identified as such.**

We consider that CoreTrustSeal, which is a community-driven, international framework used by a large palette of disciplines, is the right level for research data repositories managed in the research environment with respect to DIN 31644 (nestorseal) and ISO 16363:2013. CoreTrustSeal is regularly reviewed, a key asset to improve its fitness for purpose through community feedback, comments and suggestions for updates. FAIRsFAIR begins to assess standard approaches to extending the CoreTrustSeal or elaborating cooperatively additional requirements around the "Core", which could lead to CoreTrustSeal+FAIR or +Discipline to avoid duplication of efforts.

The existing work on certification of the services required to enable FAIR should be extended under the next framework programme and ensure applicability across disciplines. Dependencies between the critical elements of the ecosystem will also have to be assessed and taken into account. Certification comes with significant work to define the certification framework, and with a significant overhead for the services which apply. One should not seek to define certification for all the types of services in the FAIR ecosystem. Priorities should be established on which services require certification at earlier stages of EOSC, which include service registries. We point at PID services and vocabulary repositories/metadata registries as an immediate priority following the recommendations of the relevant documents of the EOSC Executive Board Architecture and FAIR Working Groups.

The capability/maturity approach proposed in CoreTrustSeal+FAIR should be extensively tested, since as explained in FAIRsFAIR D4.2 neither CoreTrustSeal nor FAIR are designed with this approach in mind. All the certification frameworks proposed for other components of the FAIR ecosystem will also have to be extensively tested and feedback from a variety of stakeholders gathered.

As stated in *Turning FAIR into Reality* Action 9.2, "concerted support is necessary to assist existing repositories in achieving certification." The needs may include the construction and implementation of the necessary community standards to enable FAIR, discussed in the companion report on FAIR Metrics for EOSC, as well as the building of the necessary skills and workforce in data management and FAIR implementation. Support for services to self-assess is needed to strengthen the ecosystem and ensure that we can rely on the Web of FAIR data and services. Certifying a repository involves costs (mainly in-house human resources, possibly also involvement of contractors), the level of costs depending on the initial repository status with respect to the criteria. In some cases, the stakeholders, including repository authorities and funders, understand the value for money of self-assessment and certification with respect to quality standards and trustworthiness demonstration. But for countries, institutions or communities not well endowed, those costs might be prohibitive with respect to resources. To get those repositories certified, support will be needed, not only guidance, but funding. Without funding possibilities those countries, institutions or communities without appropriate resources might not be able to achieve certification. If repositories remain uncertified, or researchers in certain countries or disciplines could only use central/generic repository services, the gap between the developed and less developed countries and disciplines will widen further. This must be avoided and funding schemes directed at levelling the playing field.

More generally, support should be provided to certify services enabling FAIR once the specific certification framework is defined.

*Turning FAIR into Reality* Action 13.4 states that “steps need to be taken to ensure that the organisations overseeing certification schemes are independent, trusted, sustainable and scalable.” Scalability is required to deal with the increase in the number of repositories seeking “core level” certification because of policy incentives. One element towards scalability may be to develop agreed community standards which could be machine-evaluated as predefined components of the certification process, minimising the need for human evaluation. Support may have to be provided at some point to ensure scalability, but it has to take into account and preserve the necessary independence of the organisation. This support might be provided to enable certification of repositories from countries with limited financial resources.

### 4.3 *Priorities for future work*

These priorities are for work in the immediate future. Work is on-going on several of them, in particular in FAIRsFAIR.

**Priority 1:** Support the current efforts to align Certification standards and assessment schemas with FAIR.

**Priority 2:** Test the proposed schemas in a variety of communities to gather feedback and update the proposed framework accordingly.

**Priority 3:** Provide support, methodologically as well as financially, to data and service providers to progress towards certification.

**Priority 4:** Monitor the progress of certification, assess the maturity of the certification landscape, and take appropriate action if fields or regions are lagging behind.

**Priority 5:** Support the establishment of core criteria and methodology to certify other key elements of the FAIR ecosystem, in particular in the first instance PID services and vocabulary repositories/metadata registries, and test them extensively.

**Priority 6:** Support the establishment and maintenance of registries of certified components of the ecosystem; if several registries are available for a given component, they should be harvestable and included in registries of registries.

**Priority 7:** Establish a Working Group under the EOSC Stakeholder Forum to ensure the implementation and further development of recommendations in this report.

## **ANNEX: RELEVANT RECOMMENDATIONS FROM *TURNING FAIR INTO REALITY***

### **Rec. 9: Develop assessment frameworks to certify FAIR services**

Data services must be encouraged and supported to obtain certification, as frameworks to assess FAIR services emerge. Existing community-endorsed methods to assess data services, in particular CoreTrustSeal (CTS) for trustworthy digital repositories, should be used as a starting point to develop assessment frameworks for FAIR services. Repositories that steward data for a substantial period of time should be encouraged and supported to achieve CTS certification.

Action 9.1: A programme of activity is required to incentivise and assist existing domain repositories, institutional services and other valued community resources to achieve certification, in particular through CTS.

*Stakeholders: Funders; Data service providers; Standards bodies.*

Action 9.2: A transition period is needed to allow existing repositories without certifications to go through the steps needed to achieve trustworthy digital repository status. Concerted support is necessary to assist existing repositories in achieving certification. Repositories may need to adapt their services to enable and facilitate machine processing and to expose their holdings via standardised protocols.

*Stakeholders: Data service providers; Institutions; Data stewards.*

Action 9.3: As certification frameworks emerge for components of the FAIR data ecosystem other than repositories, similar support programmes should be put in place to incentivise accreditation and ensure data service providers can meet the required service standards.

*Stakeholders: Funders; Data service providers; Standards bodies.*

Action 9.4: Mechanisms need to be developed to ensure that the FAIR data ecosystem as a whole is fit for purpose, not just assessed on a per service basis.

*Stakeholders: Coordination fora; Research communities; Standards bodies.*

### **Rec. 13: Develop metrics to certify FAIR services**

Certification schemes are needed to assess all components of the ecosystem as FAIR services. Existing frameworks like CoreTrustSeal (CTS) for repository certification should be used and adapted rather than initiating new schemes based solely on FAIR, which is articulated for data rather than services.

Action 13.1: Where existing frameworks exist to certify data services, these should be reviewed and adjusted to align with FAIR. The language of the CTS requirements should be adapted to reference the FAIR data principles more explicitly (e.g., in sections on levels of curation, discoverability, accessibility, standards and reuse).

*Stakeholders: Coordination fora; Data service providers; Institutions; Research communities.*

Action 13.2: New certification schemes should be developed and refined by the community where needed to assess and certify core components in the FAIR data ecosystem such as identifier services, standards and vocabularies.

*Stakeholders: Global coordination fora; Data service providers; Standards bodies.*

Action 13.3: Formal registries of certified components are needed. These must be maintained primarily by the certifying organisation but should also be communicated in community discovery registries such as Re3data and FAIRsharing.

*Stakeholders: Data service providers; Funders.*

Action 13.4: Steps need to be taken to ensure that the organisations overseeing certification schemes are independent, trusted, sustainable and scalable.

*Stakeholders: Funders; Research communities.*



## **Getting in touch with the EU**

### **IN PERSON**

All over the European Union there are hundreds of Europe Direct information centres.

You can find the address of the centre nearest you at: [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)

### **ON THE PHONE OR BY EMAIL**

Europe Direct is a service that answers your questions about the European Union.

You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by email via: [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)

## **Finding information about the EU**

### **ONLINE**

Information about the European Union in all the official languages of the EU is available on the Europa website at: [https://europa.eu/european-union/index\\_en](https://europa.eu/european-union/index_en)

### **EU PUBLICATIONS**

You can download or order free and priced EU publications from:

<https://op.europa.eu/en/publications>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en))

### **EU LAW AND RELATED DOCUMENTS**

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: <http://eur-lex.europa.eu>

### **OPEN DATA FROM THE EU**

The EU Open Data Portal (<http://data.europa.eu/euodp/en>) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

The Recommendations on certifying services required to enable FAIR within EOSC document contains an analysis of activities relevant to certification of the services required to enable FAIR research outputs within EOSC as of November 2020. It discusses incentivisation and support, offers an analysis of gaps and potential opportunities for extension, and defines priorities for future work.

The FAIR Working Group of the EOSC Executive Board was initially tasked to define the certification approach that will be applied within EOSC for repositories that enable FAIR research outputs, but decided to expand the remit of their work to other services, because of the recognised need to define certification mechanisms for other elements of the FAIR ecosystem.

The analysis of the status, gaps, and potential opportunities for extension, leads to define seven priorities for future work in the short term.

*Research and Innovation policy*



Publications Office  
of the European Union