Project Title	Fostering FAIR Data Practices in Europe
Project Acronym	FAIRsFAIR
Grant Agreement No	831558
Instrument	H2020-INFRAEOSC-2018-4
Topic	INFRAEOSC-05-2018-2019 Support to the EOSC Governance
Start Date of Project	1st March 2019
Duration of Project	36 months
Project Website	www.fairsfair.eu

M2.10 REPORT ON BASIC FRAMEWORK ON FAIRNESS OF SERVICES

Work Package	WP2 – FAIR practices: semantics, interoperability and services
Lead Author (Org)	Hylke Koers (SURF)
Contributing Author(s) (Org)	Patricia Herterich (DCC), Rob Hooft (DTL), Morane Gruenpeter (INRIA), Tero Aalto (CSC)
Due Date	31.10.2020 (internal milestone)
Date	30.11.2020 (published on Zenodo)
Version	1.0
DOI	10.5281/zenodo.4292599

Dissemination Level

X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Abstract

We propose a first version of an assessment framework for the FAIRness of services, together with a process to refine this model including community consultation with a view to finalizing it in August 2021. Aimed at a target audience of data service owners, the model contains concrete recommendations to improve technical aspects of services (FAIR enablement, Quality of service, Openness & Connectivity) as well as more social aspects of services (User centricity, Trustworthiness and Ethical & Legal aspects). Input was gathered from interviews with service owners, a virtual workshop held at the EOSC-hub conference earlier this year and from the recent literature on FAIR services and interoperability. The bibliography includes a series of references that have been mined for input on the suggestions in our model, and the input documents are summarized in an Annex.

Versioning and contribution history

Version	Date	Authors	Notes
0.2	3 Sept	Hylke Koers	Main structure and content description
0.21	15 Sept	Hylke Koers and all co-authors	Added section with learnings from EOSC-hub session, first draft of actual framework
0.22	20 Sept	Hylke Koers and all co-authors	Completed draft of first 3 sections of the assessment framework. Started expanding the literature section.
0.5	28 Oct	Hylke Koers and all co-authors	- Firmed up basic framework - Reviewed and improved 'findings from interviews' section - Re-factored Annex - Added abstract
0.7	30 Oct	Hylke Koers and all co-authors	Consolidated feedback, added introduction and conclusions; submitted as internal milestone
0.9	25 Nov	Hylke Koers and all co-authors	Minor edits, final review
1.0	30 Nov	Hylke Koers and all co-authors	Published on Zenodo

Disclaimer

Abbreviations and Acronyms

AAI	Authentication & Authorization Infrastructure
API	Application Programming Interface
ENVRI	Environmental Research Infrastructure
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Reusable
OLA	Operational-Level Agreement
PID	Persistent Identifier
RDA	Research Data Alliance
SLA	Service-Level Agreement

Table of contents

1. Introduction

1.1 Context and objective

To realise the full potential of a FAIR ecosystem in which research can be easily shared and optimally reused, as put forward in the *Turning FAIR into reality* report (1), it is imperative to look not only at the FAIRification of data and other digital objects, but also consider the infrastructure and services that act on those objects. In order to drive towards an optimal interplay between FAIR digital objects and services, service providers require guidance and direction on how to evolve their services. An assessment framework and/or certification scheme for data services is seen as a potentially helpful instrument to offer such guidance by offering concrete and actionable checks and recommendations (see e.g. Ref. (2) by the EOSC FAIR Working Group and Ref. (3) by this FAIRsFAIR task group, where the reader will also find a more comprehensive introduction.)

To support this endeavour, the FAIRsFAIR project has formulated the objective of developing a *FAIR assessment framework* for service owners: a framework that can be used to gauge where change is needed, together with actionable recommendations to drive incremental improvements. To be clear, the purpose of such an assessment framework is not to pass judgement on service quality or maturity; rather it proposes actionable considerations which can enhance the functioning of a service as part of a FAIR data ecosystem. This report puts forward a first version of such an assessment framework for the FAIRness of data services.

1.2 Methodology

The main result in this report is the proposed framework for FAIR assessment of services as presented in section 3. We have used three different streams of input to construct this framework:

- Literature research and analysis (described in section 2.1). We have performed a systematic analysis, extracting recommendations and requirements from sources in a structured way, to make the results as objective as possible.

- Interviews with service owners (described in section 2.2).

- Findings from audience feedback gathered during the "FAIR certification of repositories and other data services" break-out at the EOSC-hub week in May 2020 (described in section 2.3).

These different streams provided complementary views and types of input, offering a broad foundation to construct a framework on. This is important to safeguard that the framework speaks to actual needs and wishes from data services providers and offers value to all stakeholders in the wider research data community.

Using the input as described above, we have constructed an assessment framework that consists of high-level objectives together with more detailed recommendations aimed at service owners. The objectives and recommendations have been classified into a model that, at the top level, is divided into three more technically-oriented and three more socially-oriented aspects.

The framework published with this report is a first version, meant to solicit feedback and invite discussion. We also propose a refinement process, including active community consultation, to

further improve the framework between now and the completion of this FAIRsFAIR task in August 2021.

1.3 Limitations

As any work, this report has a number of limitations. We have deliberately focused on recommendations that are generally applicable across domains and across service types. This implies that we have not attempted to formulate recommendations that are either specific to an academic discipline or to a particular type of data service (and we have, thus far, shied away from any attempts to formulate a list of data service types). We would expect that, in the future, it will be useful to zoom into specific domains or service types in order to make recommendations more concrete and actionable; and also to stimulate the formation of communities of service owners who serve the same user base and who will have a strong overlap in experiences, challenges and opportunities.

Additionally, we are well aware that the literature review and analysis which we have performed is not exhaustive, if only because of the wealth of historical work (dating back to way before FAIR was formulated as such) together with rapid current developments. We have attempted to include the most pertinent current resources, but may have missed relevant work and welcome any suggestions for other work to include in the refinement process ahead of us.

2. Findings

2.1 Recent developments and insights from the literature

Since the publication of the first assessment report (3) by this task group in February 2020, there has been ample progress in specifying and operationalizing the FAIR principles for research data and other research outputs, as well as in formulating how data repositories and other services can enable FAIR data. In this section we summarize some insights and observations from the recent literature that are most relevant to the task of constructing a FAIR assessment framework for data services.

One topic of particular interest to scope of this work is the construction of FAIR metrics for digital objects. We consider such metrics as an important enabler to construct an assessment framework for FAIR-enabling services because one needs to be able to measure the FAIRness of a data object to be able to gauge the effect of a service acting on that object. (Equally, of course, a digital object cannot be made FAIR or evaluated for FAIRness in isolation from its context (4)).

Regarding FAIR metrics for objects, the RDA *FAIR Data Maturity model* Working Group[1] published its specifications and guidelines (5) in June 2020. The specification includes 41 indicators divided over three levels of priority (viz., essential / important / useful). While uptake and adoption will need to be demonstrated over time, it is foreseen that this maturity model will form a consolidated basis on which further specialized (automated) FAIR measurement tools and methods can be built – including, hopefully, tools to measure how services acting on a data object can change its level of

---

[1] https://www.rd-alliance.org/groups/fair-data-maturity-model-wg

FAIRness. Other recent work that will aid researchers in making their data more FAIR include the Danish 'how to FAIR' website[2] and the FAIR-aware tool[3] developed by the FAIRsFAIR project.

Another interesting recent development is the deployment of scalable, automated assessment tools to measure the level of FAIRness of large numbers of data objects. The EOSC-Nordic project published deliverable D4.1 entitled *"An assessment of FAIR-uptake among regional digital repositories"* (6). The authors use FAIR maturity indicators adopted from the FAIR Maturity evaluator tool developed by the FAIR Data Systems group (7) to measure, for the first time, FAIRness for a reasonably sized sample of data repositories in the Nordic and Baltic regions in an automated way. Similarly, the F-UJI tool[4] developed by FAIRsFAIR provides a programmatic way to assess the FAIRness of research datasets in (at this point) five trustworthy data repositories. It is based on (at the time of this writing, November 2020) seventeen 'minimum viable' metrics that are described in more detail in Ref. (8). Both studies demonstrate how the possibility to measure FAIRness for large numbers of datasets opens the door to assigning a FAIRness level to a data repository – and, at least in principle, to other data services – based on the FAIRness of the data objects which it harbors or operates on. We will continue to collaborate with other parts of FAIRsFAIR to ensure any service relevant aspects learned from the programmatic assessment of datasets are reflected in the FAIR assessment framework for data services.

The last several months have also seen a lot of activity and progress in establishing how services can enable FAIR data and in formulating requirements and recommendations to that end. While a detailed review and analysis of the pertinent literature is deferred to Annex A: Literature review and analysis, we summarize our key findings and observations here.

First, there appears to be a broadly carried need by service providers for guidance and directions on how to make their services enable FAIR data. This need has spurred several studies including (but not limited to) work carried out by EOSC-Nordic[5], EOSC-synergy (9) and ENVRI-FAIR (10). From these and other sources detailed in the Annex, we have put together a spreadsheet with 249 desiderata for FAIR-enabling data services. These desiderata (recommendations, requirements, rules, criteria, and other forms) have been annotated, categorized and used as input for the assessment framework presented in this work. Recognizing that some level of subjectivity is hard to avoid, we have strived to make it fully transparent if and how the various desiderata were used in this work and we welcome feedback on the choices that we have made in this process.

Notwithstanding this amount of activity, we find that the topic of FAIR assessment for data services is still in an emerging state:

- We observe a broad diversity in how desiderata are formulated: Some are presented as recommendations, others as requirements, and we have also encountered rules, criteria, principles and proof points (which show adherence to a higher-level principle). This diversity makes it more difficult to compare, aggregate and harmonize across different sources. We have also observed that there is, at present, no uniform way of assigning priority across a set of recommendations or requirements.

- Several resources go rather deep in certain aspects (e.g. quality assurance of software development) but don't touch on others (e.g. FAIR enablement). This brings a risk of providing a lopsided view on the subject matter when taken in isolation.
- The process to implement recommendations or requirements, and track progress through a consistent longer-term process, is often not clear and not discussed.
- We have not come across case studies of services, other than data repositories, having implemented recommendations and reporting back benefits, challenges and pitfalls.
- Connections to earlier, more mature, work more specifically aimed at data repositories are often lacking or left implicit. Also the relation to pre-existing work in the EOSC context (e.g. EOSC rules of participation or service management onboarding requirements) are oftentimes not clear or only specified at a very high level of abstraction.

The above should not be interpreted as criticism of the excellent work that has been done; rather it provides additional motivation to build on the work that is available and work towards consolidating it into a unified FAIR assessment framework for data services.

2.2 Interviews with service owners

We carried out five semi-structured interviews with data service owners to understand how services currently support the FAIR principles, what are transferable insights and recommendations, and what are common challenges and pitfalls. The interview guide is available in Annex B. Some of the selected services provide general services and infrastructure that can be used by a wide range of disciplines, whereas others are disciplinary services working with their dedicated communities. The interviews each lasted about 60 minutes and were structured into four parts:

1. Understanding the service, its users and context;
2. Understanding the service maturity;
3. Understanding affinity and familiarity with FAIR;
4. Soliciting viewpoints on a FAIR assessment framework.

Our analysis of the interviews has focused on highlighting both social and technical aspects of how services are currently enabling FAIRness of datasets, as well as pointing out what they would require in order for them to further enable FAIR data – and how a FAIR assessment framework could support them. We have captured insights in a structured form in a spreadsheet.

Existing knowledge and expertise

There is clearly a value body of technical expertise relevant to FAIR amongst service providers as they have been providing services to support a variety of research data management aspects for several years already. Many of the interviewees feel that they *"have been working towards the FAIR principles before the term was coined"* (H. Hermjakob), for example by providing services to support findability of data or by minting persistent identifiers using interoperable standards.

Most services already make use of automated checks where possible, assessing datasets for quality (using the service's own definition), metadata richness and testing core aspects of their services for availability and common errors.

The interviewees run established services that have been in production for several years already and thus are well connected with their user communities. They have a solid understanding of their users' needs and have established processes and advisory groups to get feedback which, in turn,

provide new ideas for additional features. Publicly funded services often make their code and contributions available as open source software on a collaborative development platform (e.g. GitHub[6]), establishing the possibility for the community to engage with the service's implementation.

Most services have a good understanding of the wider landscape they operate in and are aware of other relevant services, pointing users to complementary service offerings and collaborating to improve various aspects of FAIRness.

Needs and expectations

One helpful technical solution that was highlighted by several interviewees is an automated assessment tool that will evaluate datasets for their FAIRness. How exactly such evaluations are to be carried out should be agreed by the community to ensure the results are trustworthy and meaningful.

Another important aspect that was mentioned is training, as the majority of service providers acknowledged that a FAIR-enabling service in itself does not guarantee FAIR data: It might provide the technical features to make data FAIR, but ensuring that researchers use it to the full potential is difficult and will likely require human support. While an automated assessment tool can help with regular checks of the datasets, it was highlighted that training will be required. The training will need to be aimed towards researchers that use the services so they understand how their actions impact the FAIRness of their datasets and which functionalities of the service could help them do so. Furthermore, data curators working for the services need to be trained to support end-users and any aspects of FAIRness checking that an automated tool might not capture.

> *"It's like a driving instructor. They are allowed to pass things, you know, they're allowed to grant to FAIR badges. But it's not the car company giving you FAIR. It's not the car company saying here's a car, and we'll give you a driver's licence. It's: here's a car, you can pass your driver's licence using this because it passed its MOT[7] and it's all the level that it's safe to drive on the roads. We need the FAIR data equivalent of driving instructors".* (M. Hahnel)

Training and curation are resource intensive and it is currently unclear who will fund this in the long term. Existing funding streams to develop additional features for services are often tied to specific projects and are thus only available for a limited period, making them poorly suited to support training and curation in a sustainable way.

Assessment format

Most of the interviewees were interested in evaluating their services and prefer a more formal assessment framework for services. While self-assessment can be an initial step, it was suggested that an external and community-endorsed assessment framework will engender more trust and recognition.

> *"We want a test that we are able to pass, but we nevertheless want it to be quite serious. Nothing is more horrible to an honest operator than a test that doesn't actually test.*

---

[6] https://github.com/

[7] The MOT test (Ministry of Transport, or simply MOT) is an annual test of vehicle safety, roadworthiness aspects and exhaust emissions required in the United Kingdom for most vehicles. From https://en.wikipedia.org/wiki/MOT_test

10

Self-assessments with free form questions [...] tend to select for people who know how to advertise themselves well." (G. Aben)

This assessment should be transparent and it should be seen as an instrument to help improve rather than passing a verdict of failure. It could incorporate ideas from the scientific peer review process where other service delivery experts provide feedback and highlight areas of improvement going forward. And, as going through an external assessment will consume resources, it needs to be made clear what the service will gain from being assessed.

## 2.3 Input from EOSC-hub session "FAIR certification of repositories and other data services"

On 19 May 2020, the FAIRsFAIR project organized a session "FAIR certification of repositories and other data services"[8] at the EOSC-hub week. In addition to disseminating earlier work generated by this task group, the main objective of the session was to gather input from the audience about their needs and wishes with regard to a FAIR assessment framework for services – with, of course, a view of feeding this into the assessment framework that will be delivered at the end of the task (August 2021). The session yielded valuable input on three different aspects of FAIR service assessment:

- **Scope**, i.e. which services are seen to be essential in enabling FAIR data and should thus be in scope for the assessment framework;

- **Desired qualities**, i.e. what are seen to be important properties for a data service to be an enabler for FAIR data;

- **Form** of the assessment framework, specifically on the continuum from *descriptive* (sharing recommendations and good practices) to *prescriptive* (formal certification)

These different types of findings are discussed in more detail below. The session was attended by more than 150 people, out of which 90 provided input to at least one question. At the start of the webinar, participants were asked to self-identify their role in an open question "How would you describe your role in relation to data repositories and services?". The answers were mapped to the role of (i) service provider; (ii) service user, or (iii) support staff (with multiple options possible). As figure 1 below illustrates, 40 participants (44% of respondents) represented the service provider community, which is the core target audience for the assessment framework that is the subject of this work.

---

[8] More information about the session can be found at
https://www.eosc-hub.eu/eosc-hub-week-2020/agenda/fair-assessment-certification-repositories
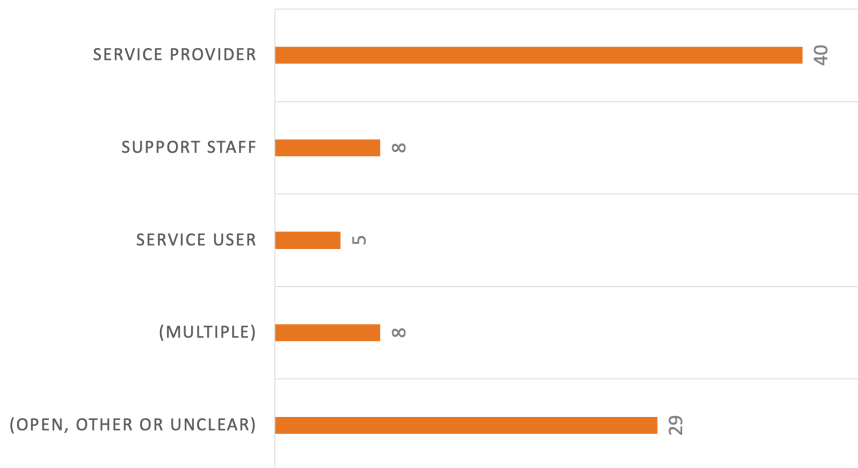
Figure 1: Break-down of the active participants (i.e. those who have provided input) over the professional roles of service provider, service user, or support staff.

Scope

The audience was asked which services they consider to be essential to enable FAIR data. The results are presented in figure 2 below as a word-cloud of the raw data (left) together with a classification into main service categories.
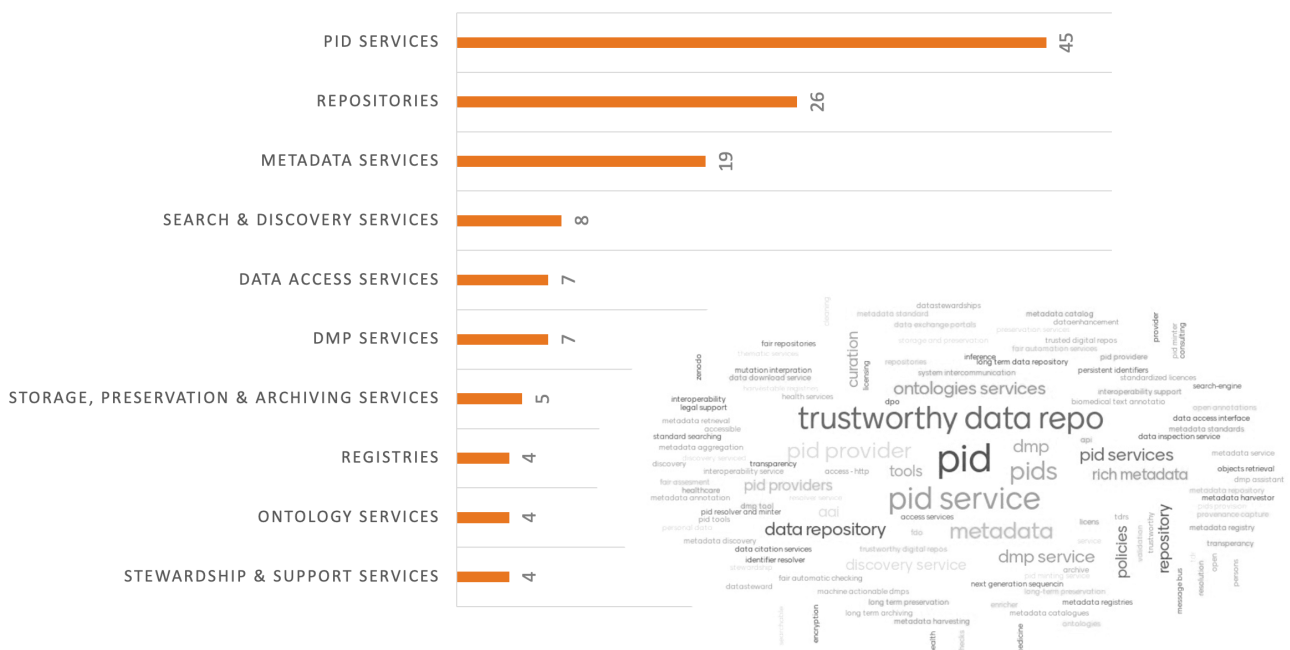
Figure 2: Audience response to the question: *"Please name three types of data services that you consider essential to enable FAIR data?"*. The word-cloud illustrates the raw audience feedback as free text; these answers are normalized and mapped onto ten service categories which are shown in the histogram.

We observe that PID services, repositories and metadata services constitute a clear top-three of services that are seen to be most essential to enable FAIR data. It is interesting to compare this to the findings reported in the "Interim recommendations on certifying the services required to enable

FAIR research outputs within EOSC" (2). The authors of that work posed a similar question "What kind of services should be certified" to a (smaller and more diverse) audience of a break-out session at the 2019 EOSC Symposium. Their audience prioritised repositories, PID systems and registries. While we notice clear alignment in the critical roles for PID systems/service and repositories, it is interesting to observe the difference in how registries are ranked in terms of importance; a possible explanation for this might lie in the difference in stakeholder representation between the two audiences.

Desired qualities

Having established the scope of data services under consideration, the audience was asked what they consider the most important qualities for these services in order to enable FAIR data. The results are shown in figure 3 below. As expected, some concepts that are very central to FAIR – interoperability, metadata, standards, transparency – are mentioned as being the most important for FAIR-enabling services as well. In addition to these qualities, a rather wide array of technically and socially oriented qualities are also mentioned, ranging from availability and reliability to user support and openness. This supports our earlier observation that an assessment framework should be rooted in FAIR data, while also considering several dimensions of service utility and quality (see the guiding principles put forward in Ref. (3)).
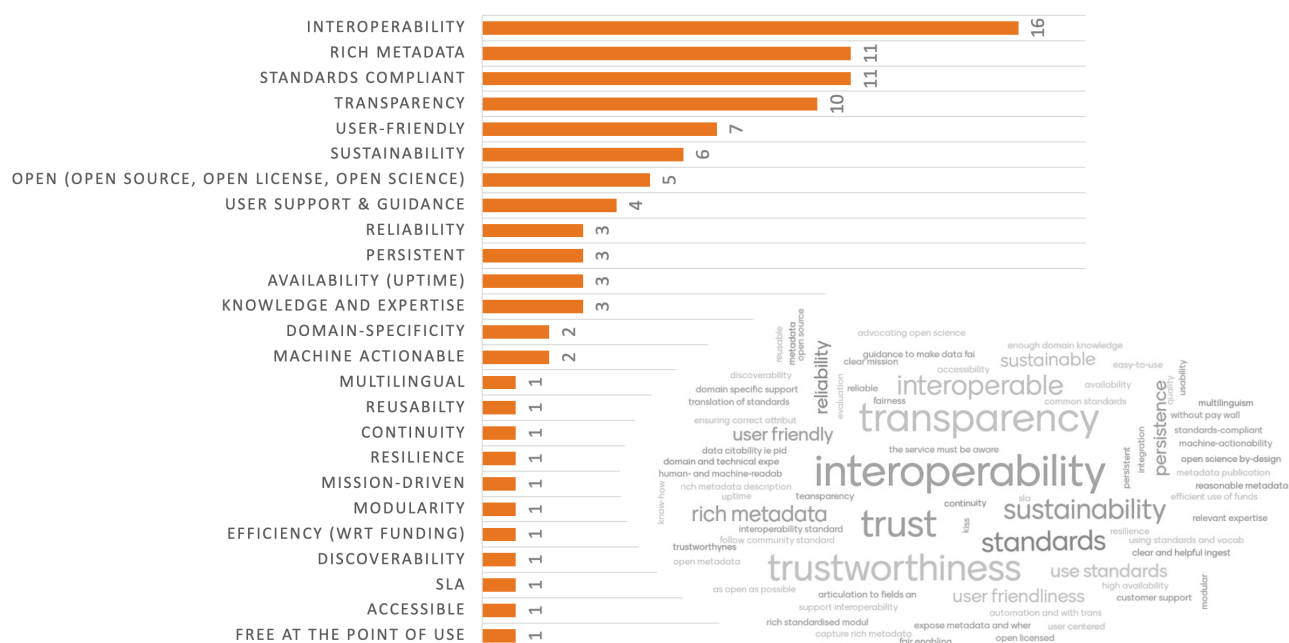

Figure 3: Audience response to the question: "*What do you consider to be the most important qualities for a data service to enable FAIR data?*". The word-cloud illustrates the raw audience feedback as free text; these answers are normalized and mapped to categories which are shown as a histogram.

Form

The last set of questions was aimed at getting input into the best approach towards making services more FAIR-enabling – with options ranging from a more descriptive to a more prescriptive framework:

1. Shared 'good practices' and recommendations for FAIR-enabling services;
2. Self-assessment tool for FAIR-enabling services;
3. Formal certification process for FAIR-enabling services.

The audience was asked to score each of these options in terms of their perceived importance on a scale from 0 to 10. The results are shown in figure 4 below. The figure summarizes responses from the whole audience (i.e. aggregated over all roles); however we have validated that restricting to the segment of 'service providers' provides similar results.
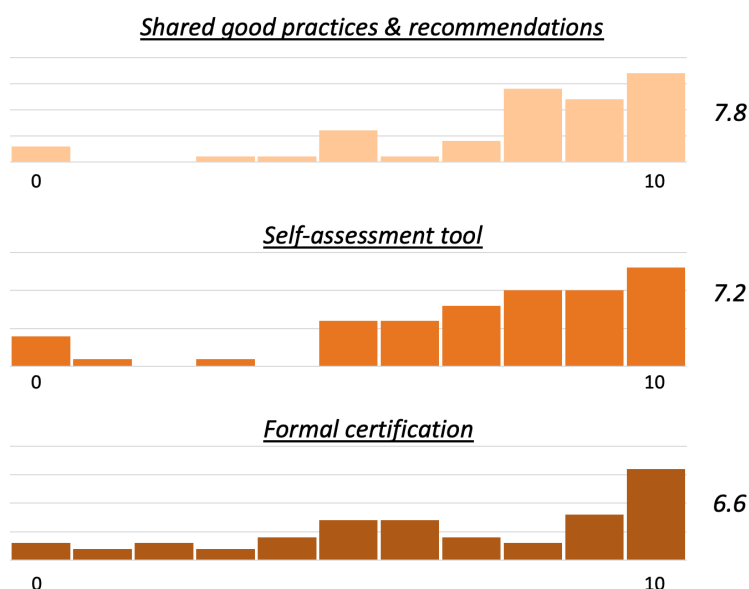
Figure 4: Distribution of audience rating on the question *"How important is it for you that (i) there are shared 'good practices' and recommendation for FAIR-enabling services; (ii) There is a self-assessment tool for FAIR-enabling services; (iii) There is a formal certification process for FAIR-enabling services".* The histograms show the distribution of the scores; the numbers on the right provide the average value.

As the figure shows, feedback from the audience indicates a preference for less formalized methods over more formalized: sharing good practices and recommendations is scored higher than self-assessment tools which, in turn, scores higher than formal certification. The 'why' question, which was asked after the scoring questions, provided some good insights into the rationale for this preferred ordering. Concerns around formal certification fall into four categories:

- Certification may be not flexible enough to do justice to the diversity of the data services landscape;
- The time may not yet be ripe for formal certification; best practices need to become established first;

- Certification may (perhaps undesired) exclude part of the community by focusing on the 'top end' of service providers;
- It is not clear whether certification will really change user behavior and, thus, incentive service providers towards desired behaviour. "In the end, users should be able to vote with their feet".

The respondents with a preference for formal certification typically indicated they see certification as the "gold standard" which the community should strive for: *"Each step represents progress toward the same goal, with certification serving as the highest standard".*

3. A proposed framework for FAIR assessment of services

3.1 Preface

In this section we propose a draft basic framework for the FAIR assessment of data services. It builds on existing frameworks and other work reported in the literature ([Recent developments and insights from the literature](#)), insights gathered from interviews with service owners ([Interviews with service owners](#)) and feedback gathered from the community ([Feedback from EOSC-hub session](#)). In addition we build on the foundational work reported in the first milestone report (3) by this task group.

The proposed framework, as laid out below, should be seen as a step towards the final deliverable of this task group: *A framework for assessing FAIR services* (due August 2021). The framework below is meant to invite feedback and initiate further discussion to iteratively develop it into an assessment framework that is solidly rooted in the needs of the community. A process to manage this refinement is discussed below.

It is important to note that the framework put forward at this stage is not meant to be prescriptive or normative. Instead, it aims to be a useful tool for service owners to develop a better understanding of points of concern for data services in general and of strong and weak points in their own service offering; to make them aware of potential gaps or points for improvement; and to offer directions for improvement through the sharing of good practices or otherwise.

Guiding principles

The first report (3) published by this task force listed a number of guiding principles and high-level requirements which were identified as important for our assessment framework. We repeat these principles here as follows:

- 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);[9]

- *Build upon existing work* as much as possible, for example extending concepts and criteria from frameworks such as CoreTrustSeal where possible;

- *Consider several dimensions of a service*, i.e. not only functional aspects ('utility' in FitSM[10] 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.

Audience and Scope

The framework is developed for **service providers**, for **all data services in all disciplines**. Challenges or recommendations that are specific to a certain domain or type of service have, by choice of scope, not been considered for inclusion in this assessment framework. While such domain-specific requirements are important for implementation, many difficulties and enablers are common across domains (11) and we decided to exclusively study these.

For all aspects of the assessment framework, it is assumed that the service is live and available to users in a production environment.

3.2 FAIR assessment of services: Basic framework

The proposed assessment framework analyses six independent aspects[11] of data services spanning both technical and social dimensions. Each of these aspects will be explored in a separate sub-section. For each aspect a high-level objective is formulated, together with a set of specific recommendations to help achieve that objective. The recommendations are formulated to specifically address the service owner and their organisation.

---

[9] The next section presents a number of case studies in which we analyze how, from a functional perspective, a service can make digital objects more, or less, FAIR.

[10] FitSM is a light-weight IT service management specification, seel https://www.fitsm.eu/

[11] While we have attempted to formulate independent aspects, some overlap will be inevitable. Feedback and suggestions for improvement will be very welcome as part of the refinement process ahead of us.

Technically-oriented aspects

FAIR enablement

Objective:

The service enables FAIR data by elevating the FAIRness of digital objects and/or supporting the FAIRification process. FAIR enablement is actively driven through the implementation of community-supported standards and interoperability frameworks.

Recommendations:

- Perform a self-assessment on how the function(s) of the service *enable, respect* or *reduce* each of the FAIR principles for the data that it operates on.[12] Make the results of the self-assessment publicly available, together with an outlook on the desired state for the service (including a cost/benefit analysis).[13]

- Use automated tests that show how the service increments FAIRness of digital objects in a verifiable, measurable, repeatable and scalable way. Root such tests in community-supported methodologies that measure the FAIRness of digital objects in an objective way.

- In consultation with the target community (or communities), identify which metadata schemas and other standards (e.g. technical and semantic aspects of data encoding) should be adopted. Consider in particular domain-specific standards and practices. Strive to include accessibility conditions in metadata. Where applicable, generate and capture metadata automatically.

- Engage with both the user community and other service providers to improve interoperability between services. Of particular attention here are authentication & authorization infrastructure (AAI), PIDs, and data and metadata encoding specifications. Seek alignment with existing or emerging data type registries and interoperability frameworks, e.g. the EOSC interoperability framework.

- Consider both human and machine access to the service, specifically with a view towards supporting automated pipelines for the FAIRification of digital objects.

Quality of service

Objective

The service is delivered in a reliable, secure, high-quality way, consistent with its specifications.

Recommendations:

- Codify the service's availability and other non-functional aspects in a public Service Level Agreement (SLA) which is easy to understand by users from different communities.

---

[12] The case studies presented in Ref. (3) offer a suggested format for this self-assessment. Of course other formats are acceptable as well, however we do recommend to include all of the aspects listed in the case studies (i.e.: Summary; Users; Purpose; Adoption; Services; Target Digital Objects; Examples; FAIR enablement mapping).

[13] Note that a service does not need to address all aspects of FAIR, and integration with other FAIR-enabling services (e.g. PID minting) is often preferable over developing your own solutions.

17

- Deploy the service on appropriate and well-supported hardware or virtual (cloud) infrastructure. Define operational-level agreements (OLA) with 3rd-party infrastructure services that enable service delivery.

- Take reasonable technical and non-technical measures to prevent, detect, and respond to cyber or physical security threats; securing the service and protecting sensitive information resources (e.g. using secure HTTP connections only). Organize security audits and pen-tests at regular intervals, ideally at least every two years.

- Assess whether the service deals with sensitive data (e.g. patient records) and, if so, take additional measures in line with both applicable legislation and expectations from the user community.

- Implement service management processes to bolster a reliable and predictable service delivery (including but not limited to capacity planning).

- Implement service management processes to govern changes in a controlled way. Make release notes and documentation publicly available. Announce maintenance breaks well ahead of time. Maintain backward compatibility when possible.

- Implement (ideally automated) testing procedures for every change to the service or a service (component) that it integrates with. Testing should ideally include not only functional testing, but also performance and stress testing.

- Implement service management processes to deal with incidents or vulnerabilities in an effective and transparent way. Implement and test disaster recovery procedures. In case of service interruptions, aim to restore service as soon as possible even if that requires work-arounds or other temporary measures.

- Implement a service monitoring system that generates alerts in case of unexpected behavior, including functional, performance and security-related issues.

- Implement and make available a set of metrics as indicators for the performance, stability and adoption of the service.

Open & Connected

Objective:

The service is operated in a transparent, low-barrier and inclusive way; seeking integrations and connections with other services; and championing principles of openness consistent with Open Science and Open Research.

Recommendations:

- Publish clear, inclusive and non-discriminatory licences and/or terms of use. Enable wide access to the service.

- Where possible, offer the service free-at-the-point-of-use for researchers (which does not preclude other monetization or cost-recovery models)

- Adopt well-documented and community-supported open standards and specifications, in particular for API's and other interfaces.

- Make the service and all documentation available online through URLs that are fully qualified domain names and assigning PIDs where applicable.

- Seek integrations with other services rather than replicating functionality, especially for common reusable infrastructure components. Adopt EOSC architectural components and standards as enablers for deep interoperability with other services in the EOSC portfolio.[14]

- Use community-supported PIDs to integrate with other services; keep data, metadata and PID's tightly connected. Consider implementing the FAIR Digital Object model to enable interoperability with other data services.

- Where possible, make any source code that is used to run the service available under a common open-source licence.[15]

- Seek inclusion in relevant service catalogs, ideally obtaining and using a PID for the service.

Socially-oriented aspects

User centricity

Objective

The service is managed such that it serves the (possibly evolving) goals of the user community, and maximises usability while minimizing burden.

Recommendations:

- Drive for continual improvements to the user experience. Actively work with the community to understand and improve usability, for example through user tests or design studios.

- Ensure that there is an ongoing, consistent dialogue between the service and its user community, such that users can optimally make use of the service and influence its development.

- Understand how the service fits within the data management norms and expectations of the target user community.

- Engage the user community in establishing and prioritizing the service's backlog and roadmap.

- Ensure that sufficient documentation is available for users and organize a process to regularly review and update (at least with every change to the service). Documentation should cover functional aspects, a description of the various service components and their relationship, and explain which phases of the data life cycle and data management processes are supported by the service. Ideally documentation should be version-controlled, have a PID and an (open) licence.

---

[14] Part of the EOSC interoperability framework, the EOSC Profiles (https://data.d4science.net/13af) specify common data models for EOSC entities (Providers, Resources, etc) which helps drive interoperability of resources within EOSC.
[15] See e.g. https://spdx.org/licenses/ for a list of relevant software licences.

- Include multi-lingual support and accessibility features[16], both for the service and its documentation, to the extent relevant for the service's (potential) user base. Key information must be available in English if the service is intended to be included within EOSC.

- Invest in user training and outreach activities to help users understand the service's value proposition and how to effectively use it.

- Ensure the service provider organization has adequate support staff available to assist users where needed.

Trustworthiness

Objective:
The service is perceived by the user community as reliable and trustworthy, both in terms of its utility and its warranties[17], now and in the future.

Recommendations:

- Clearly communicate the service's core value proposition and any pertinent (technical or non-technical) features, as well as its limitations.

- Be open and transparent about organisational mission, business model, legal status and target user communities. Be transparent and accountable about costs, profits and cost-recovery models.

- Take reasonable measures to ensure a sustainable long-term operation — including both financial and organisational aspects. Aim to reduce long-term operational dependencies on short-lived project funding. If available, provide clear information to indicate how long the service will minimally be available and maintained.

- For services that are meant to preserve research objects over a longer period of time (such as data repositories), have a clear minimum preservation timeframe.

- Implement technical measures to safeguard the continuity of the service, and the longevity and integrity of any (meta)data that is stored as part of the service. This includes keeping backups on independent systems, implementing fail-over mechanisms and exercising proper life-cycle service management.

- Implement an appropriate and transparent governance structure that includes representation of the service's target user community.

- Seek to attain certification where relevant community-endorsed certification mechanisms exist.

- Ensure the service provider organization has sufficient staff with knowledge to operate the service, now and in the future.

---

[16] For accessibility on the web, we specifically recommend the Web Content Accessibility Guidelines (WCAG) overview: https://www.w3.org/WAI/standards-guidelines/wcag/

[17] The concepts of 'utility' and 'warranty' are adopted from FitSM to speak to a service's key purpose and additional factors that impact the customer's quality perception, respectively. See https://www.fitsm.eu/

Ethical & Legal

Objective:

The service complies with all applicable legal and ethical guidelines, in a transparent and auditable way.

Recommendations:

- Define, publish and adhere to a code of conduct that is in accordance with commonly agreed principles regarding the conduct of research in the service's user community.

- Take reasonable measures to ensure data is handled in compliance with disciplinary and ethical norms, and that data licences are clearly defined and respected.

- Take reasonable measures to manage the intellectual property rights of data producers.

- Maintain a publicly available privacy policy.

- Clearly communicate a dedicated contact address for security issues including hacks, vulnerabilities and privacy breaches. Ensure the address is actively monitored by multiple staff members.

- Implement auditable measures to ensure that the service respects all applicable legislation and regulations around user privacy and sensitive data (including but not limited to GDPR in Europe). In particular, when processing personal data, roles and responsibilities must always be well-defined and data subjects must be provided with the name and contact details of the data controller and of the Data Protection Officer.

3.3 Next steps: Refinement process

The basic assessment framework we have presented above marks an important milestone in our journey towards a FAIR assessment framework for data services. To reiterate, we have constructed the proposed framework on the basis of three streams of input: a methodological review of the current literature; input from stakeholders representing our target audience at the EOSC-hub week; and a series of interviews with service owners.

This broad approach is meant to ensure that the framework is solidly rooted in the needs and wishes from service owners and the broader community. Nevertheless further work is needed to validate this alignment and further improve the framework. Specific points of attention where we expect the framework may be improved include:

- Recommendations may be sharpened and be made more actionable;

- Recommendations may be assigned a PID for unambiguous and stable reference;

- Recommendations could be prioritized, possibly segmented by different types of data services (e.g. PID services and data repositories may have different priority recommendations);

- Per recommendation, additional detail and working examples of successful adoption or implementation could be collected and made available;

● Specifying which recommendations, when implemented, could be validated automatically and where human interpretation will remain needed.

To achieve these improvements, we propose an extended process of refinement through community consultation starting at the publication of the present report and drawing to a close with this task's final deliverable due in August 2021. Concretely, we propose the following schedule:

Dec – Jan 2020	Public consultation on assessment framework as proposed in this work – including webinar to kick off the public consultation (perhaps with RDA FAIR maturity WG)
Feb 2021	Workshop with different stakeholder groups to validate completeness
March 2021	Dissemination of workshop findings, including next iteration of the assessment framework
April 2021	Second workshop for service owners to validate utility and actionability
May 2021	Draw up final framework for assessing FAIR services
June 2021	"Last chance" external expert review
June 2021	Internal review within FAIRsFAIR
July – Aug 2021	Finalization and publication of framework for assessing FAIR services (deliverable D2.7)

In parallel to this cycle of consultations and workshops, the task group will continue to actively engage with service owners and other representatives from the community, as well as pioneers of FAIR and EOSC services.

4. Conclusions

With this report we put forward a first version of an assessment framework for the FAIRness of data services. While we expect the framework will be of value to multiple stakeholders in the FAIR data ecosystem, the framework is specifically targeted at service owners (i.e. individuals or organisations developing and operating services for research data). It offers a structured list of high-level objectives together with specific recommendations towards realizing those objectives. By design, the framework is generalistic in the sense that it covers all domains and all types of services.

The framework was constructed on the basis of a methodological literature analysis, interviews with service owners and audience feedback gathered at the *"FAIR certification of repositories and other data services"* break-out session at the EOSC-hub week earlier this year. This report marks the beginning of a process of refinement, during which we will actively engage with service owners and the wider FAIR community to improve and sharpen the framework — leading up to the delivery of a "framework for assessing FAIR services" in August 2021.

5. Acknowledgements

6. Bibliography

1. European Commission Expert Group on FAIR Data. Turning FAIR into reality: final report and action plan from the European Commission expert group on FAIR data. Brussels: European Commission. Directorate-General for Research and Innovation. 2018. Available from: https://doi.org/10.2777/1524

2. Genova, F., Aronsen, J., Beyan, O., Harrower, N., Holl. A., McIntosh, L. et al. Interim recommendations on certifying the services required to enable FAIR research outputs within EOSC. 2020. Available from: https://repository.eoscsecretariat.eu/index.php/s/zCnHTcytBHaLjRp

3. Koers, H., Gruenpeter, M., Herterich, P., Hooft, R., Jones, S., Parland-von Essen, J., Staiger, C. Assessment report on 'FAIRness of services'. Zenodo 2020. Available from: https://doi.org/10.5281/zenodo.3688762

4. L'Hours, H., Von Stein, I., Huigen, F., Devaraju, A., Mokrane, M., Davidson, J. et al. D4.2 Repository Certification Mechanism: a Recommendation on the Extended Requirements and Procedures. Zenodo 2020. Available from: https://doi.org/10.5281/zenodo.3835698

5. Research Data Alliance FAIR Data Maturity Model Working Group. FAIR Data Maturity Model: specification and guidelines. 2020. Available from: https://doi.org/10.15497/RDA00050

6. Jaunsen, A., Kleemola, M., Alaterä, T., Lehvaslaiho, H., Hasan, A., Nordling, J., Assinen, P. D4.1 An assessment of FAIR-uptake among regional digital repositories. Zenodo 2020. Available from: https://doi.org/10.5281/zenodo.4045402

7. Wilkinson, M., Dumontier, M., Sansone S., Bonino da Silva Santos, L., Prieto, M., Batista, D. et al. Evaluating FAIR maturity through a scalable, automated, community-governed framework. Sci Data. 2019 Dec; 6(1):174. Available from: https://doi.org/10.1038/s41597-019-0184-5

8. Devaraju, A., Huber, R., Mokrane, M., Cepinskas, L., Davidson, J., Herterich, P. et al. FAIRsFAIR Data Object Assessment Metrics. Zenodo 2020. Available from: https://doi.org/10.5281/zenodo.3934401

9. Orviz, P., David, M., Gomes, J., Pina, J., Bernardo, S., Campos, I. et al. EOSC-synergy: A Set of Common Service Quality Assurance Baseline Criteria for Research Projects. 2020. Available from: https://dx.doi.org/10.20350/digitalCSIC/12533

10. Papale, D. ENVRI-FAIR D11.1 Biodiversity and Ecosystem subdomain implementation short term plan. Zenodo 2020. Available from: https://doi.org/10.5281/zenodo.3885361

11. Deniz Beyan, O., Chue Hong. N., Cozzini, S., Hoffman-Sommer, M., Hooft, R., Lembinen, L. et al. Seven Recommendations for Implementation of FAIR Practice. Zenodo 2020. Available from: https://doi.org/10.5281/zenodo.3931993

12. Mokrane, M., Recker, J. CoreTrustSeal-certified repositories: Enabling Findable, Accessible, Interoperable, and Reusable (FAIR). In: 16th International Conference on Digital Preservation. Amsterdam, The Netherlands: iPRES; 2019. Available from: https://doi.org/10.17605/OSF.IO/9DA2X

13. Lin, D., Crabtree, J., Dillo, I., Downs, R., Edmunds. R., Giaretta, D. et al. The TRUST Principles for digital repositories. Sci Data. 2020 May 14;7(1):144. Available from: https://doi.org/10.1038/s41597-020-0486-7

14. Koers, H., Bangert, D., Hermans, E., Van Horik, R., De Jong, M., Mokrane, M. Recommendations for Services in a FAIR Data Ecosystem. Patterns 2020. Available from: https://doi.org/10.1016/j.patter.2020.100058

15. Riungu-Kalliosaari, L., Hooft, R., Kuijpers, S., Parland-von Essen, J., Tana, J. D2.4 2nd Report on FAIR requirements for persistence and interoperability. Zenodo 2020. Available from: https://doi.org/10.5281/zenodo.4001631

16. Corcho, O., Eriksson, M., Kurowski, K., Ojsteršek, M., Choirat, C., Van de Sanden, M. et al. EOSC Interoperability Framework : DRAFT. 2020. Available from: https://www.eoscsecretariat.eu/sites/default/files/eosc-interoperability-framework-v1.0.pdf

17. Wittenburg P. From Persistent Identifiers to Digital Objects to Make Data Science More Efficient. Data Intelligence 2019 1:1, 6-21. Available from: https://doi.org/10.1162/dint_a_00004

18. Landi, A., Thompson, M., Giannuzzi, V., Bonifazi, F., Labastida, I., Bonino da Silva Santos, L., Roos, M. The "A" of FAIR – As Open as Possible, as Closed as Necessary. Data Intelligence 2020 2:1-2, 47-55. Available from: https://doi.org/10.1162/dint_a_00027

19. McQuilton, P., Batista, D., Beyan, O., Granell, R., Coles, S., Izzo, M. et al. Helping the Consumers and Producers of Standards, Repositories and Policies to Enable FAIR Data. Data Intelligence 2020 2:1-2, 151-157. Available from: https://doi.org/10.1162/dint_a_00037

20. Juty, N., Wimalaratne, S., Soiland-Reyes, S., Kunze, J., Goble, C., Clark, T. Unique, Persistent, Resolvable: Identifiers as the Foundation of FAIR. Data Intelligence 2020 2:1-2, 151-157. Available from: https://doi.org/10.1162/dint_a_00025

21. Dekker, R. Social Data: CESSDA Best Practices. Data Intelligence 2020 2:1-2, 220-229. Available from: https://doi.org/10.1162/dint_a_00044

.

Annex A: Literature review and analysis

In this section we discuss and analyze selected recent work which has presented recommendations, requirements and other desiderata for FAIR-enabling data services. These desiderata are collected in a spreadsheet which we have used as input for the assessment framework put forward in Section 3 of this report. The spreadsheet also collects requirements and recommendations from a number of publications and resources that were already discussed in Ref. (3), such as the SHARP framework [18] and CoreTrustSeal (12). Alignment between the FAIR principles and CoreTrustSeal ("CoreTrustSeal + FAIR") – with the ultimate objective of "maximizing the quantity of FAIR data under trustworthy curation" – has recently been discussed by FAIRsFAIR WP4 in Ref. (4).

Collectively, the analyses presented here and in the associated spreadsheet also show the emerging state of this field: There are ample suggestions, recommendations and requirements – but what is still lacking is an overall cohesive framework that includes all relevant aspects and provides practical suggestions and guidance to service owners, while doing justice to the intrinsic differences between academic disciplines and communities as well as to the different levels of maturity within the current services landscape. We hope that the framework presented in this work may serve as a good step in that direction.

EOSC Portal and Service Portfolio onboarding requirements

The EOSC portal website[19] lists a number of requirements for service providers in order for their service(s) to be included in the EOSC Portal:

- The service is accessible by users outside its original community.
- The service is described through a common template focused on value proposition and functional capabilities.
- At least one service instance is running in a production environment available to the user community.
- Publish Research data is Findable, Accessible, Interoperable and Reusable [reference to FAIR].
- Release notes and sufficient documentation are available.
- Helpdesk channels are available for support, bug reporting and requirements gathering.

Additionally, there are a number of criteria which are checked as part of the service onboarding process. At the time of writing, work is being undertaken to improve and streamline this process; however the working document with a set of "Criteria for possible inclusion in the EOSC Service Portfolio" is available through EOSC-hub.[20] These criteria, and the above requirements, have been captured and used as input in the construction of the assessment framework presented here.

---

[18] See https://repository.eoscsecretariat.eu/index.php/s/XLXpPsJ5HZGqq3M#pdfviewer
[19] EOSC portal website: https://www.eosc-portal.eu/for-providers
[20] See https://wiki.eosc-hub.eu/display/EOSC/Criteria+for+possible+inclusion+in+the+EOSC+Service+Portfolio

EOSC Rules of Participation

The EOSC Rules of Participation Working Group – one of the working groups that have been set up following the establishment of the EOSC Executive Board – is tasked with "defining the rights, obligations governing EOSC transactions between EOSC users, providers and operators"[21]. This includes a set of rules for EOSC services which we consider relevant input for the assessment framework presented here.

At the moment of writing, the formulation of these Rules of Participation is still a work-in-progress and the most recent draft document is v0.2 stemming from 29 January 2020.[22] Section C of this document applies specifically to services.

EOSC-Nordic service maturity model assessment tool

EOSC-Nordic developed a maturity model assessment tool[23] which aims to help service providers in the Nordic and Baltic regions to improve service quality, in particular with a view to aligning services with requirements from the EOSC. The tool is made available in the form of a spreadsheet with a total of 29 requirements – mapped onto minimum / intermediate / high maturity levels – grouped into five topical areas:

- Service management
- Data management and quality
- Accessibility and legal requirements
- Sustainability and Financial
- EOSC architecture compatibility

The work is directly relevant to this report because it presents a set of requirements for data services that may serve as input and inspiration. Here we note that, even though the document is explicitly targeted at "services in the Nordic and Baltic region", the model itself appears to be readily applicable to other geographies. After review, selection and consolidation, the 29 stated requirements have been used as input for the assessment framework put forward here.

In terms of content and scoping, we note that it relies heavily on general service management requirements as specified by the FitSM framework. Beyond these requirements, the framework also makes reference to FAIR and compatibility with EOSC, but it does not provide specific requirements or recommendations that would give guidance or direction to service owners.

EOSC-Synergy: A Set of Common Service Quality Assurance Baseline Criteria for Research Projects

The EOSC-synergy project published a document "Set of Common Service Quality Assurance Baseline Criteria for Research Projects" (9), which proposes 39 high-level criteria (some of which having underlying more detailed criteria). Targeting web services, web applications and platform or service composition, the criteria are meant to establish Service Quality Assurance criteria for the development deployment, operation and integration of services into a production research infrastructure such as the EOSC – with the ultimate view to building trust; ensuring functional suitability; improving usability; and promoting automated validation.

The proposed criteria are associated with requirement levels "MUST", "SHOULD" and "MAY" according to RFC 2119[24] and are organized into the following 11 categories:

- API Testing
- Integration Testing
- Functional tests
- Performance tests
- Documentation
- Security
- Policies
- Support
- Automated Deployment
- Monitoring
- Metrics

The listed criteria are heavily rooted in IT service management and DevOps software development practices including automated testing and API's[25]. It provides detailed and specific recommendations on topics such as testing, service integration, monitoring and metrics; on the other hand, it does not address any specific requirements stemming from the FAIR principles.

We have included the criteria as input for the assessment framework presented here, though many of the criteria are of a greater level of specificity compared to the intended scope of this work.

TRUST principles

The TRUST principles, presented in Ref. (13), propose a set of principles that complement the FAIR principes by demonstrating *trustworthiness* in data repositories. A mnemonic that stands for Transparency, Responsibility, User Focus, Sustainability, Technology, the TRUST principles call attention to long-term data stewardship including topics such as governance, resources and security. While primarily formulated for data repository owners, the principles are also expected to

---

[24] See https://tools.ietf.org/html/rfc2119

[25] In particular, the document embraces specifications from the OpenAPI initiative (see https://www.openapis.org/)

be beneficial for a broader audience including funders, researchers as well as organisations or individuals managing other data services.

In some more detail, the TRUST principles advocate:

- **Transparency** — by being open and transparent about mission statement and scope, terms of use, preservation timeframe and any other pertinent additional features or services.

- **Responsibility** — by adhering to the designated community's metadata and curation standards; providing stewardship of the data holdings; providing data services and appropriate interfaces; managing intellectual property rights; protecting sensitive information resources; and ensuring overall security of the system and its content.

- **User Focus** — by implementing relevant data metrics; providing or contributing to community catalogues; and monitoring and acting upon changing needs within the designated user community.

- **Sustainability** — by planning for risk mitigation, business continuity, disaster recovery, and succession; secure funding; and providing adequate governance

- **Technology** — by implementing relevant and appropriate standards, tools, and technologies; and by having measures in place to protect against (cyber or physical) security threats.

While the TRUST principles are primarily targeting the data repository community, many of these elements are readily generalized to other data services. We therefore consider these principles, and the more detailed recommendations included in the article, as input for the assessment framework presented here.

Recommendations for Services in a FAIR Data Ecosystem

The article "Recommendations for Services in a FAIR Data Ecosystem" (14), co-authored by representatives of FAIRsFAIR together with RDA Europe, OpenAIRE, EOSC-hub and FREYA, presents 20 recommendations for services to support FAIR data. These recommendations were collected through a series of workshops including different stakeholders within the broader FAIR community. The recommendations are grouped into seven categories as follows:

- Certification
- Essential Infrastructure Components
- Stewardship
- Costs
- Rewards
- Collaboration and Support
- Data Management

The recommendations were prioritized, and priority recommendations were mapped to actions and associated with action owners. These priority recommendations and actions are (see Table 1. of Ref. (14)):

Recommendation	Action
Consider FAIR alignment and data sharing as part of research assessment, among other criteria.	Infrastructures should be evaluated and rewarded to be FAIR-aligned
	Reward researchers who apply the FAIR principles to their research, e.g., through incentives such as increased visibility for their work
Domain-specific ontologies, as domain-specific requirements have to be taken into account.	Identify disciplines which do not have ontologies and create awareness for registries of ontologies and enrich them
	Make repositories support FAIR by developing tools, such as APIs, and share best practices and user stories
Establish data stewardship programs providing simple and intuitive training for researchers, and enable data stewards and researchers who support applications of FAIR.	Identify and present the cost of developing supporting infrastructure, including human resources

This work is of particular interest because it followed a bottom-up approach, collecting and prioritizing recommendations through a series of workshops involving multiple communities and stakeholders. The 20 recommendations have all been considered as input for the assessment framework presented in this work, though it should be noted not all of them are applicable for service owners (e.g. there are also recommendations targeting funders or researchers).

ENVRI-FAIR D11.1: Biodiversity and ecosystem subdomain implementation short term plan

ENVRI-FAIR WP12 published its first deliverable *"D11.1 Biodiversity and ecosystem subdomain implementation short term plan"* (10) at the end of February. The report analyzes the status, in terms of FAIRness, of seven Research Infrastructures (RI's) within the Biodiversity and Ecosystem sub-domain. For each of the seven RI's, the report presents strengths and weaknesses in FAIRness, together with a 2-year plan for improvement. Because the improvement plans are very specific to the individual RI's, we did not consider them as input into the assessment framework presented here. However, the report also outlines a number of ways to improve interoperability and accessibility (see pp 21-29 in Ref. (10)) – which we here summarize as follows:

- Clearly identify and communicate core competencies as well as shareable tools
- Define and use common parameters
- Seek alignment on data and technical metadata formats.

In addition to providing these reusable recommendations, the paper is also of value because it describes a 'good practice' example of collaboration across RI's to work together to advance FAIRness.

FAIRsFAIR D2.4 2nd Report on FAIR requirements for persistence and interoperability

At the end of August, the *"2nd Report on FAIR requirements for persistence and interoperability"* was published on Zenodo (15). It is the second iteration (of three reports) sketching the state of FAIR within Europe written by the FAIRsFAIR project. The report reviews practices and presents recommendations to enable FAIR through PID's, semantic interoperability and metadata. While the majority of the discussion and recommendations target researchers and data stewards, section 2.1.3.2 specifically addresses how service providers may support FAIR by assigning and integrating with PID's. Table 1 below, copied from the Ref. (15), summarizes this. The publication also offers a helpful analysis of the role of PID's in general, the relation between PID's and metadata, and an overview of PID systems for different entities in the scholarly ecosystem.

	Reduce	Enable	Support
Infrastructure level	The service does not accept persistent identifiers as values paired with natural language values when creating metadata or exposing metadata.	The service offers the option to add external persistent identifiers (DOI, ORCID) and can create new PIDs for the digital objects it hosts.	The service also offers integrated common reference metadata and presents (meta)data both for humans and machines.
Project level	The service does not enable linking versions and requires manual (free text) creation of descriptive metadata.	The service offers the possibility to create structural metadata (internal and external PIDs to versions and other relevant DOs)	The service also creates PREMIS or other types of controlled event metadata and links workflows and provenance metadata automatically

Table 1: "How services can support FAIR through using PIDs", copied from Ref. (15).

EOSC Interoperability Framework

Early May, the Interoperability Task Force of the EOSC FAIR Working Group, with participation from the Architecture WG, published a version 1.0 of the EOSC Interoperability Framework (16). At the time of writing, the report is still open for community review. Building on pre-existing work from the European Interoperability Framework[26], the document sets out to identify general principles and recommendations for "efficient machine-enabled exchange of digital objects within EOSC and

---

[26] https://ec.europa.eu/isa2/publications/european-interoperability-framework-eif_en

between EOSC and the outside world." (copied from the text). Targeting an audience of "service providers of EOSC-related services", the document lists technical, semantic, organisational problems, needs and requirements with respect to interoperability. It also provides a number of recommendations, some of which will be immediately relevant for service providers (see the associated spreadsheet for our analysis on the applicability of these recommendations for service owners). It is also of interest to note that the document proposes the FAIR Digital Object model (see e.g. Ref. (17) and reference therein) as a cornerstone for the EOSC interoperability framework.

Figure 5: Four levels of interoperability — legal, organisational, semantic and technical — as discussed in Ref. (16). Figure copied from that work.

The FAIR Principles: First Generation Implementation Choices and Challenges

This Special Issue[27] of the journal Data Intelligence, published under the editorship of Barend Mons, Erik Schultes, Fenghong Liu and Annika Jacobsen, collects 28 articles which report experiences, practices and challenges encountered by individuals and organizations spearheading the implementation of the FAIR principles. For the purpose of the framework proposed in this work, we have analyzed Refs. (18), (19), (20) and (21) in detail and extracted requirements and recommendations for service owners in the main spreadsheet feeding into the assessment framework proposed in this work.

---

[27] See http://www.data-intelligence-journal.org/p/67/

Annex B: Service owner interview script

This script was used to guide the conversation, but not followed in a very rigid way.

Part 1: Understanding the service, its users and context

Goal: Getting to understand what the service aims to do, and how it fulfills that function. Also get a sense of what the service owner believes are "good properties" of the service.

Specific questions:
- Name of the service
- Which organization is responsible for running it?
- What does it do, i.e. what is the value proposition and what is its functionality?
  - What are competing / alternative services?
  - What is the "unique selling point" of your service?
- Usage
  - Who are the primary users?
  - How many users?
  - Usage metrics? (E.g. % users vs. addressable market)
  - What are Alternatives for your service? Competition? Different approaches?

Part 2: Understanding service maturity

Goal: Getting to understand the level of maturity of the service (and the organization responsible for running the service) and if this is an area of interest. Also, do they feel they could benefit from guidance on these aspects?

Specific questions:
- Quality & warranty
  - Do you have an SLA or other performance warranties?
  - Do you have a process for quality control?
- Do you have disaster recovery plans in case of a failure?
- How do you deal with incidents and/or changes? (Are there incident and/or change management processes in place?)
- How are user requests (usage help, bugs, feature requests) handled?
- How is your service funded? Is there a plan to ensure its sustainability?

Part 3: Understanding affinity and familiarity with FAIR

Goal: Getting to understand the familiarity with FAIR and (active or passive) support. Get to know if they see their service as 'FAIR enabling'? If so, get to know about good practices to share; if not, find out if they could benefit from guidance on this.

Specific questions:
- Familiarity with FAIR
  - What can you tell me about FAIR?
  - Would you say that you support it?

- ○ Do you see this as a 'cost' or an 'investment'? What do you feel you could gain?
- ● Sharing expertise (technical and social)
  - ○ Have you ever done a FAIR assessment about for your service
    - ■ To what extent do you consider FAIR to be applicable to your service?
    - ■ Did you do a cost/benefit analysis for the FAIR principles?
    - ■ Did you consciously decide about not implementing one of the FAIR principles?
  - ○ What has your service done to enable FAIR?
    - ■ Examples?
    - ■ What actions did you take to make that happen?
    - ■ Did you see any benefits?
  - ○ Can you share some challenges and bottle-necks that you faced?
  - ○ What is the single most important tip?
    - ■ What would you have wanted to know yourself before you got started?
- ● Removing barriers (technical and social)
  - ○ Are there things that you could do to your service to better enable FAIR data?
  - ○ What is the single most important thing that keeps you from doing that?
  - ○ What would you need to overcome those barriers?

Part 4: FAIR assessment framework

Goal: Get explicit feedback on what would make the framework valuable for them. (Ask this only at the very end to not influence the earlier discussion about their service and their needs).

Open discussion: Explain the work we're doing and ask what – if anything – they would find useful to have included.