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

M2.7 ASSESSMENT REPORT ON 'FAIRNESS OF SERVICES'

Work Package	WP2 - FAIR practices: semantics, interoperability and services
Lead Author (Org)	Hylke Koers (SURFsara)
Contributing Author(s) (Org)	Morane Gruenpeter (INRIA), Patricia Herterich (DCC), Rob Hooft (DTL), Sarah Jones (DCC), Jessica Parland-von Essen (CSC), Christine Staiger (DTL)
Due Date	29.02.2020
Date	28.02.2020
Version	1.0
DOI	https://doi.org/10.5281/zenodo.3688762

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

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, we must look beyond 'just' research data: on the one hand we should include research software and other digital objects, and on the other hand we should lay out a vision and a roadmap for services and supporting infrastructure that fosters an optimal interplay between digital objects and services. The latter calls for an assessment framework for service owners which can be used to gauge where change is needed, together with actionable recommendations to drive incremental improvements.

However, while the development of various FAIR assessment frameworks for data and other digital objects has enjoyed substantial activity over the last few years, a FAIR assessment framework for data *services* is still lacking — a fact that was recently underlined again in the interim recommendations from the EOSC FAIR Working Group. FAIRsFAIR task 2.4, entitled 'FAIR services', aims to close this gap by developing a FAIR assessment framework for data services (alongside a similar framework for research software).

This report marks the first milestone of the task. It 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. We are seeking wide feedback on this report to inform subsequent work and, ultimately, feed into a FAIR assessment framework for data services that delivers clear direction and value to service owners and the community at large.

Versioning and contribution history

Version	Date	Authors	Notes
0.7	04.02.2019	Hylke Koers and all contributing authors	Reviewed content at task biweekly meeting, discussed last improvements.
0.9	24.02.2019	Hylke Koers and all contributing authors	Draft for internal review with selected FAIRsFAIR task and WP leaders
1.0	28.02.2019	Hylke Koers and all contributing authors	Incorporated feedback from PCO and FAIRsFAIR colleagues; ready to publish on Zenodo for wide community feedback.

Disclaimer

Abbreviations and Acronyms

EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Reusable
PID	Persistent Identifier

Table of contents

1. Introduction

Since their inception in 2014 (1) and formal publication in 2016 (2), the FAIR Data Principles have spurred tremendous activity and progress towards better research data management (3). As "guiding principles for scientific data management and stewardship", the FAIR Data Principles have proven to provide a valuable framework to organize the discussion amongst different communities and to develop standards, practices and services that support the creation and stewardship of research data. Yet, the FAIR Data Principles are an instrument — offering support and guidance along the way — rather than a destination in itself. As noted by the authors of Ref. (3) and many others, the principles are formulated as high-level guidelines that require further interpretation and definition by research communities to become truly actionable; a process that inevitably opens the door for ambiguity and subjectivity. Furthermore, and more specific to the scope of this report, the FAIR data principles focus primarily on data and not on services acting on data — though it is abundantly clear from the emphasis on machine actionability that the authors fully acknowledge the importance of such services. It is paramount to recognize that data and other digital objects can not be made FAIR without a number of enabling services that facilitate the provisioning of persistent identifiers (PIDs), provide indexable resources and support access, amongst other factors.

As articulated clearly and compellingly in the *Turning FAIR into reality* report (4), to realize the true value of the FAIR paradigm requires more than just FAIR data; what is needed is a 'FAIR ecosystem' that comprises of FAIR digital objects[1] (which could represent data, software, protocols or other research resources) together with services and an underpinning infrastructure. Especially in light of the necessarily distributed nature of such an ecosystem, a variety of services will be required to create, store, share and consume data and other FAIR digital objects. To do this successfully calls for an optimal interplay between digital objects and services — driving, in the words of Ref. (5), "maximal synergistic integrated utilization". Such an interplay should not be expected to just emerge organically but will require guidance, direction and fostering (building on an extensive legacy for both data and infrastructure). It is here that FAIRsFAIR Task 2.4[2] sees a role for itself: **by delivering an assessment framework for data services that will enable and stimulate such an interplay and thus help realize the full potential of a truly FAIR ecosystem.**

Since the publication of the FAIR Data Principles, many assessment frameworks[3] for FAIR datasets[4] have been put forward (see Annex B for a non-exhaustive overview; and note that in this report we will use 'assessment framework' as a broadly defined term that includes evaluation and certification tools). Typically, these provide further definition to what it means for data to be 'Findable', 'Accessible', 'Interoperable' and 'Re-usable' — sometimes complemented by additional

---

[1] In this report we will use the term 'digital objects' in a broad and inclusive way to refer to datasets as well as research software and other research artefacts where appropriate.
[2] This is one of the four tasks under FAIRsFAIR Work Package 2, see
https://www.fairsfair.eu/fair-practices-semantics-interoperability-and-services
[3] Note that in this report we will use 'assessment framework' as a broadly defined term that includes self-evaluation tools as well as more formal auditing and certification systems.
[4] Many of the existing FAIR assessment frameworks apply specifically to 'datasets' rather than the more broadly defined term 'digital objects'; where that is the case we follow existing nomenclature.

criteria around e.g. openness or quality aspects. Such tools typically give the user a means to quantify the level of 'FAIRness' of a dataset (thereby reducing some of subjectivity mentioned above) and provide recommendations to make them 'more FAIR'.

In addition to assessment frameworks for datasets and other digital objects, certification schemes for data repositories such as the CoreTrustSeal have been developed and are enjoying strong and growing adoption. For data services other than data repositories the current landscape is less populated: There are general quality and operational requirements stemming from IT service management, and there are several lists with criteria for services to be included in catalogues such as the European Open Science Cloud (EOSC) or the ELIXIR Core Data Resources — but to the best of our knowledge there is currently no assessment framework that can be used to gauge the level of 'FAIRness' of any data service and offer guidance as to how it can be improved. This gap was also noted in Ref. (4), which includes as Recommendation 9 to "Develop assessment frameworks to certify FAIR services" and as Recommendation 13 to "Develop metrics to certify FAIR services". Quoting from the report in more detail: "More work is needed to extend the FAIR data principles for application to a wide range of data services, including registries, Data Management Planning tools, metadata standards and vocabulary bodies, identifier providers, software libraries and other cloud services. Such extensions must take into account good management practice and sustainability. In doing so, the example of CoreTrustSeal and recommendations about business models and sustainability are good places to start". This task formulation is echoed in the recent report (6) from FAIRsFAIR Work Package 4 which states that "... a wider vision of FAIR ecosystem dependencies and interconnections is required" yet also signals that "The scope of and approach to FAIR services have not yet been fully defined". Similarly, Ref. (5) states that "... it is still unclear how to facilitate digital resource producers to define, assess, and implement FAIRness within their specialized specific projects". Finally, the need for certification of services and other elements of the FAIR ecosystem was recently underlined in a report with interim recommendations (7) from the EOSC FAIR Working Group.

The work being undertaken in FAIRsFAIR Task 2.4 aims to fill exactly this gap, leading up to the publication of a framework for assessing FAIRness of services at the close of the project in March 2022 (alongside a similar framework for FAIR software). The proposed Statement of Work for the task, pending approval from FAIRsFAIR project office after proposed modifications, is included as Annex A to this report.

The present report signifies a first milestone along this journey. It is meant to put a stake in the ground and set the scene for ongoing work by clarifying our starting points, intentions and ambitions. It will also put forward initial guiding principles for the assessment framework and present a number of case studies in which we analysed the 'FAIRness' of selected data services in a rather ad-hoc way (with the intention to learn from the process, feeding into a more formal methodology). Being a first publication for this Group, this report is also explicitly meant to connect with other groups and initiatives working on the assessment, certification or evaluation of data services — hopefully spurring collaboration and realizing synergies that not only save work but also underpin the development of a truly connected FAIR ecosystem.

Around the same time as the finalisation of this report, a related report Recommendations for Services in a FAIR data ecosystem (8) was published on Zenodo. This report presents a list of

recommendations, together with actions and priorities, for services in a FAIR ecosystem. These recommendations were sourced from the community in a series of workshops in an initially free-form and bottom-up way, rendering it complementary to the work presented here which attempts to structure the discussion, connect with earlier frameworks, and work towards an assessment framework in a more directed, top-down manner.

Finally, this report is structured as follows. Section 2 presents an overview and a brief discussion on existing FAIR assessment frameworks and related initiatives. It builds on an overview of such frameworks collected in Annex B. Section 3 discusses scoping and framing for the assessment framework. It also proposes a set of guiding principles and desiderata for the final assessment framework that this task will deliver. Section 4 presents three case studies on 'FAIRness of services' and, finally, a Conclusion section completes this report.

2. FAIR assessment at the start of 2020

2.1. Review of current FAIR assessment frameworks

At the time of compiling this report, there is already an impressive body of FAIR assessment frameworks — and there is ample activity to develop these further and create new frameworks in areas of (perceived) gaps. Annex B contains an overview of existing FAIR assessment frameworks. Building on overviews compiled by the RDA FAIR Data Maturity Model WG[5] and the FAIRassist tool developed by the FAIRsharing resource[6], this overview is meant as a baseline overview of the features most pertinent to the present work from the most relevant existing frameworks. (While the overview is not meant to be exhaustive, any feedback about substantial omissions will be very welcome.)

As may be appreciated from the Annex, there is no shortage of assessment frameworks dealing with the level of FAIRness of datasets. While there is some diversity how these frameworks are designed — reflecting different target audiences or purpose — they are typically organised along the axes of "Findable", "Accessible", "Interoperable" and "Re-usable", sometimes extended with criteria pertaining to data quality, completeness, openness or trustworthiness which are complementary or, in some cases, implicit to FAIR. The assessment frameworks typically add specific meaning and definition to the (high-level) principles such that their users can readily score their dataset against them. The output is often a FAIRness indicator or score together with recommendations on how to improve the score and, thereby, the FAIRness of the dataset. As an illustration for this kind of scoring, the GARDIAN FAIR metrics assessment tool gives 0 points for "No PID, no metadata and/or documentation"; 2 points for "No PID but sufficient Metadata and/or documentation"; and 5 points for "(Meta)data are registered or indexed in a searchable resource".[7]

The majority of the analysed frameworks are manual, often taking the form of a self-evaluation tool which dataset owners can use themselves. While human interpretation is without doubt essential in the assessment of some subjective FAIR principles (for example 'Data are described with rich metadata'), assessing the FAIRness of digital objects at the kind of scale (and levels of abstraction) required for a FAIR ecosystem is simply not feasible by humans only, calling for an element of machine validation. Motivated by this challenge, Wilkinson et al. have proposed a more scalable and automated approach (9,10). Their proposed framework relies on research communities to develop Maturity Indicators and implement Compliance Tests that automatically test digital resources against those Maturity Indicators — thus collectively serving as a roadmap for incremental improvements on the FAIRness of a resource. At the time of writing, this approach is still actively being developed (for example through the FAIR Evaluation Services platform at https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/) and it will be interesting to monitor the success of this promising approach.

---

[5] https://www.rd-alliance.org/groups/fair-data-maturity-model-wg
[6] https://fairsharing.org/
[7] As this example immediately illustrates, even some relatively mature concepts such as PIDs are still not common practice (at least not in all domains).

It is worth noting that the framework proposed by Wilkinson et al. builds on the concept of a 'data resource'. While a formal definition is not provided by the authors, such a data resource can be thought of as a combination of a digital object together with services that act on the object to deliver it to the user in a suitable form (where 'user' may be a human or a machine agent, for example a Compliance Test). In other words, the proposed Compliance Tests typically measure 'FAIR at the point of consumption' by the user — which essentially is a convolution of the FAIRness of the digital object in itself (the static piece of content) and the services acting on it (which delivers the content to the user, for example including a process of content negotiation). This means that, if the level of 'FAIRness at the point of consumption' is measured to be low (e.g. no metadata with the digital object as delivered to the user), that could be due to low FAIRness of the digital object itself (e.g. because there is no metadata associated to it in the first place) or to a service breaking the data's intrinsic FAIRness (e.g. metadata is lost in delivering the data to the user). Practically, in such a case that would mean that further analysis will be necessary to disentangle the effect of data services from the intrinsic FAIRness of the original digital object.[8][9] This convolution also implies that the proposed approach cannot be readily applied to formulate an assessment framework for data services, though such an assessment framework can no doubt benefit from an automated and scalable process to assess the FAIRness of data resources.

The FAIRshake project (5) takes a similar approach towards community-driven, (semi-)automated FAIR assessment. Focusing at least initially on the biomedical research community, FAIRshake has developed a toolkit that enables the community to define manual or automated tests to assess the FAIRness of digital objects. The toolkit includes API's, browser extensions, a bookmarklet and a website. Finally, FAIRshake also supports insignias (badges) which can be embedded on websites hosting data objects thereby making the (measured) level of FAIRness very transparent to end-users in an easily digestible, visible way — and incentivising the organisations hosting such objects to deliver measurable improvements.
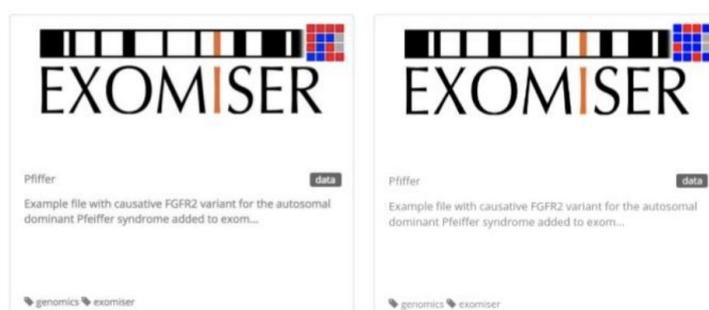
Figure 1: Illustration of how the 'FAIRness' of a digital resource can be visualized through online badges. The red/gray/blue boxes on the top-right display the outcome of certain (automated) tests on the EXOMISER resource. The left panel represents the state before the 2019 BioIT World NCBI FAIR Hackathon; the right panel represents the state thereafter. Copied from Ref. (5).

---

[8] As an amusing analogy (Mark Wilkinson & Hylke Koers, private conversation): Think of a process that creates sausages (data resources). If the sausage tastes bad (the test fails), that speaks to the state of the sausage at the point of consumption - which is a function of its ingredients (the data object) as well as the mixing, casing, distribution and all other processes (the data services) that have acted on them.

[9] Ideally, of course, services acting on a digital object should preserve their level of FAIRness, a principle that we will colloquially refer to as "FAIR-in-FAIR-out" but that still requires further definition and precision.

Moving beyond datasets (or data resources) to data services or other entities, the landscape becomes less populated as also suggested by the more concise listings in Annex B. Specifically for data services, we can group the assessment frameworks that exist into three main categories:

- Assessment frameworks, including more formal certification schemes, for data repositories. A well-known example of this category is the CoreTrustSeal developed by the DSA–WDS Partnership Working Group on Repository Audit and Certification, a Working Group (WG) of the Research Data Alliance[10]. Such frameworks do not automatically generalise in full to other data services, though they may serve as a valuable starting point.
- "Rules of Engagement" and other lists with criteria for services to be included in catalogues such as the European Open Science Cloud (EOSC) or the ELIXIR Core Data Resources (11). Typically these address various kinds of quality criteria, serving effectively as gatekeeper to preserve the value and trustworthiness of the catalog. In addition, they may safeguard some minimum level of interoperability by demanding the use of common standards. Or, in the words of Ref. (3), "FAIR guiding principles provide a scaffold for building such rules of engagement within each community."
- General service quality requirements stemming from IT service management, for example the FitSM framework.[11]

Beyond these categories, there appears to be a gap for a FAIR assessment framework for data services in general (see also Refs. (4,6)), underlining the need and urgency of the work carried out by this Group and reported here.

2.2. Other insights and recommendations from the literature

This section includes a brief review and discussion on relevant recent literature that offers insights or proposed recommendations for FAIRness of services that were, for various reasons, not included in the overview of assessment frameworks compiled in Annex B.

Firstly, Dunning et al. (12) have performed a detailed analysis of the 'FAIRness' of over 40 data repositories, scoring the individual data repositories against FAIR principles. While this analysis did not take the form of a formal assessment process, each repository was scored against a number of criteria on the basis of which the authors can draw conclusions on compliance with the various elements of FAIR, and compare compliance rates for different types of repositories. One of the main conclusions is that many repositories could make a big step in alignment with the FAIR principles by formulating and implementing a number of basic policies in areas such as:

- Creating a lasting policy for deploying PIDs
- Insisting on a minimum set of metadata, ideally coupled with the preferred use of semantic terms
- Having a clear licence
- Using HTTPS

---

[10] https://www.rd-alliance.org/groups/repository-audit-and-certification-dsa%E2%80%93wds-partnership-wg.html
[11] FitSM is a lightweight standard for IT service management developed for federated e-Infrastructures. See e.g. https://www.egi.eu/services/fitsm-training/

Furthermore, the authors plea for closer alliances between data archives and researchers, arguing that expertise and insights from both perspectives are needed to satisfy the FAIR principles.

More recently Hahnel and Valen (13) have investigated to what extent ten current, well-known data repositories (five disciplinary and five generalistic) align with the FAIR principles (see Ref. (13) for the results of the mapping). In addition to this mapping exercise, they have put forward a number of recommendations for data repositories "to help standardize the academy's data repository infrastructure and bolster interoperability". These recommendations may be summarized as follows:

- Persistent Identifiers, ideally including provisions for metadata indexing in central metadata stores such as provided by DataCite or Crossref DOI's.
- Application Programming Interface (API's), ideally following the OpenAPI specification
- Data Curation and Moderation Workflows, including human factors as well as the implementation of supporting technical workflows.
- Accessibility, including for people with disabilities, leveraging recommendations from WCAG (https://www.w3.org/WAI/standards-guidelines/wcag/)
- Licences for Reuse
- Sustainability, including technical provisions for disaster recovery as well as a scalable and sustainable business model for those elements of FAIR that require human capacity.

While these recommendations were not included in the overview of assessment frameworks in Annex B (because they are presented as general recommendations and are not directly linked to an assessment process by Hahnel and Valen), some elements should provide valuable input into the formulation of a FAIR assessment framework for data services.

Another recent paper worth mentioning is FAIR Ecosystem Components: Vision (6) authored by representatives of FAIRsFAIR work package 4.[12] Building on the FAIR ecosystem as described in *Turning FAIR into Reality* (4), it puts forward a high-level vision of the various components in a FAIR ecosystem with their dependencies and interactions. The paper places special emphasis on the dynamic nature of FAIRness, i.e. it is seen as a quality that requires long-term, ongoing investment by data stewards and researchers. These actors thus take a prominent role in the presented vision alongside Trustworthy Digital Repositories and several registries. Services (and software) are also included in the high-level vision diagrams, be it only in connection to repositories. This is an (intentional) limitation of the current paper — which focuses on repositories and made the assumption in their model that data is being held by a repository — which may be extended in future work.

Finally, as also mentioned in the introduction, a recent report entitled Recommendations for Services in a FAIR data ecosystem (8) — co-authored by FAIRsFAIR — is in many ways complementary to the present report. It presents a list of recommendations (together with actions and priorities) for services in a FAIR ecosystem, sourced from the community in a series of workshops in an initially free-form and bottom-up way.

---

[12] https://www.fairsfair.eu/fair-certification

2.3. Related initiatives

Some four years after the formal publication of the FAIR Principles (2), FAIR assessment is a subject of great interest with several groups and task forces working on the development of new assessment frameworks or maturing existing ones. Here we mention a number of initiatives most relevant to the scope of FAIRsFAIR Task 2.4.

Within the FAIRsFAIR project, Task 2.3 focuses on FAIR data repositories. This Task's upcoming report *Set of FAIR data repositories features* (FAIRsFAIR D2.3) presents a number of policy, technical and other recommendations. While these recommendations are formulated for data repositories, it is expected that some of them may generalize to other data services — a possibility that will be actively explored between the two tasks going forward. FAIRsFAIR Work Package 4 focuses on the evaluation and certification of FAIR digital objects and FAIR-enabling repositories. The forthcoming report *Draft recommendations on requirements for FAIR datasets in certified repositories* (FAIRsFAIR D4.1), which presents a set of preliminary metrics corresponding to FAIR principles that can be used to assess digital objects, also discusses commonalities with Task 2.4 and presents a number of use cases that may prove fertile ground for further collaboration. In addition, the work within Task 3.4 entitled *Transitioning to FAIR data providers, stewardship and repositories* is expected to offer additional insight into certain aspects of implementing FAIR recommendations.

Within the broader context of the European Open Science Cloud, the EOSC FAIR working group[13] is tasked to provide recommendations on the implementation of 'open' and FAIR practices within the EOSC. The Working Group has recently published initial recommendations on FAIR service certification and FAIR metrics for EOSC (7,14). We expect significant potential for synergy with the work planned by this Group, which we will actively coordinate through individuals that are both members of FAIRsFAIR Task 2.4 and the EOSC FAIR WG. In particular, a jointly organised workshop on certification frameworks for services that enable FAIR is planned for Spring 2020.

Other groups and initiatives that will continue to interact and seek alignment with include EOSC Synergy WP3 (Fostering Service Integration), GO FAIR (in particular the GO BUILD Implementation Network[14]), FAIRsharing, the FAIRplus project (in particular WP3[15] which will be looking at tools that support FAIRification and work with metrics for database interoperability), EOSC Nordic (for example connection with ongoing work to carry out technical FAIR assessment of repositories) and the COAR Next Generation Repositories initiative. In this ongoing effort to align and collaborate, this Group expects to benefit from the Synchronisation Force within the FAIRsFAIR project and other communication and coordination activities.

Finally, in the context of related initiatives, it is worth noting that Plan S[16], though primarily addressing literature publications, places requirements on publication venues that are also of

---

[13] https://www.eoscsecretariat.eu/working-groups/fair-working-group
[14] https://www.go-fair.org/go-fair-initiative/go-build/
[15] https://fairplus-project.eu/about/how-project-organised#wp3
[16]
https://www.coalition-s.org/addendum-to-the-coalition-s-guidance-on-the-implementation-of-plan-s/principles-and-implementation/

relevance to research data and other digital outputs (e.g. "Linking to data, code, and other research outputs that underlie the publication and are available in external repositories."). Furthermore, by requiring that publication venues for more traditional research outputs use elements which are core to the FAIR principles — e.g. PID's, metadata and machine-readable license information — these services and platforms will get connected more tightly into the FAIR ecosystem.

3. Towards an assessment framework for 'FAIRness of services'

3.1. Data Services and the FAIR ecosystem

When FAIRsFAIR Task 2.4 kicked off in August 2019, the directive as stated in the FAIRsFAIR project plan — to "... propose ... a framework for assessing the FAIRness of services and set a series of recommendations for making services FAIR" — seemed a daunting objective. Some six months later it still is, though the many conversations with experts and various stakeholder groups have helped to better place the work in context and develop a view as to how to scope and frame the task such as to deliver a valuable contribution to the FAIR research data community.

Indeed, right from the start of the task, it was apparent that a naive (or, in the words of Ref. (15), 'slavish') mapping of the 15 FAIR principles (formulated for datasets) to data services would, despite being an entertaining exercise, probably not deliver actionable insights of real and lasting value. It was clear that a better starting point would be to step back to the original purpose of the FAIR paradigm and explore the role, challenges and opportunities for data services to help realize its potential. This broader picture is well laid out and discussed in the Turning FAIR into reality report (4) published in 2018 (and see also Refs. (6,15)). Central to this report is the concept of a 'FAIR ecosystem' that comprises of FAIR Digital Objects (which could represent data, software, protocols or other research resources) together with services and an underpinning infrastructure — and the notion that we need to develop such an ecosystem to realize the full potential of the FAIR paradigm. Data services are an essential component in this, yet there is limited tangible guidance on how to "make services FAIR" and become an integral part of the FAIR ecosystem. This gap is the starting point and motivation for our task to formulate recommendations and a FAIR assessment framework for data services.

3.2. Scope

The above discussed how data services are part and parcel of the FAIR ecosystem and how they are necessary to realize the potential of the FAIR paradigm. In this discussion, as in much of the underlying literature, a certain lexical ambiguity may be permitted but if we want to further crystallize the thinking the question as to what exactly is a data service (and what it is not) becomes more pressing. As a starting point, a definition was put forward in Ref. (16) as: "A means of delivering value to the producers and users of digital objects by facilitating outcomes they want to achieve without the ownership of specific costs or risks." FitSM does not offer a definition for data service, but it defines an IT service as a "Service that is enabled by the use of information technology (IT)", where the term 'service' is defined as a "Way to provide value to customers through bringing about results that they want to achieve".

While these definitions provide a basis in terms of associating 'services' with a way of delivering value to users/customers without the cost of ownership, the specialization to 'data' is less clear. We expect that existing models and frameworks such as the FAIR Digital Object model (see Ref. (17) and references therein), the Open Archival Information System (OAIS)[17], and the 'Internet of FAIR

---

[17] http://www.oais.info/

Data & Services' as promoted by GO-FAIR[18] will provide valuable starting points in this regard — though first of all further study and discussion will be necessary to determine the need for a more formal model describing data services, objects, resources, etc.
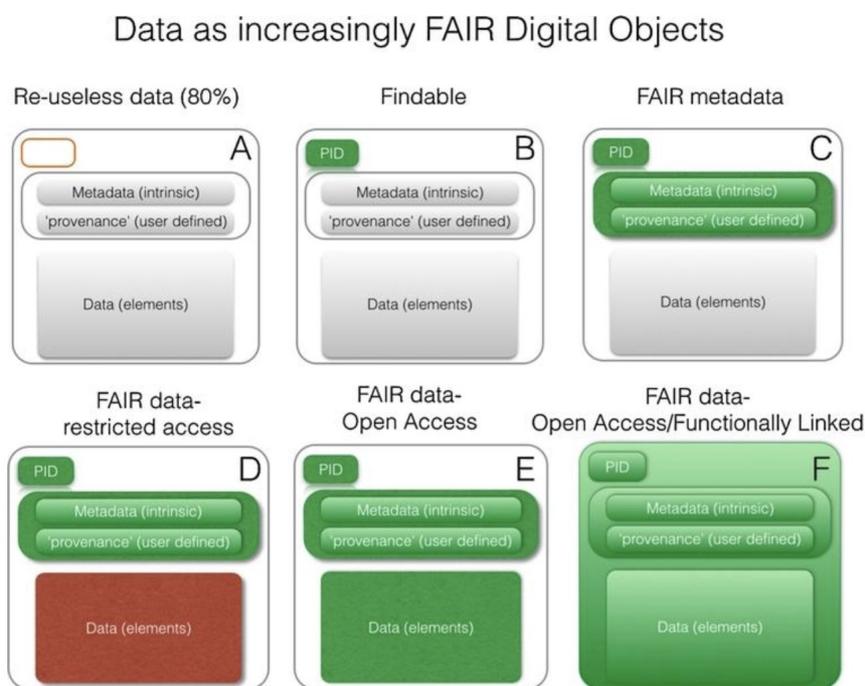

Figure 2: Illustration of FAIR applied to Digital Objects. Copied from Ref. (3).

As a working definition, we will consider here as a 'data service' any service that acts on at least one component of the 'holy trinity of data management': the bit sequence, the metadata and the PID of a digital object. This includes services that bind these components together (e.g. associating metadata with a bit sequence), services that deliver data to the user, services that automatically analyze or transform data, services that aggregate and index metadata, services that store or replicate data, etc. Also here, more work is needed to determine whether a more formal description of data services — e.g. a classification or a taxonomy — would be necessary or desirable.

3.3. Purpose

In order for the assessment framework to fuel progress towards realizing the FAIR ecosystem, it is essential that it gives guidance for those who are in a position to affect change. We therefore believe that the FAIR assessment framework, and the recommendations associated with it, should primarily be directed at organizations that develop and provide data services — with the purpose to **give insights and actionable recommendations that enable them to make incremental improvements to their services that support the emergence of a FAIR ecosystem.**

---

[18] https://www.go-fair.org/resources/internet-fair-data-services/

The landscape of organisations developing, managing and providing data services is diverse, ranging from academic parties to commercial businesses to national and European e-infrastructure providers, and from domain-specific to generalistic. All such parties are considered to be stakeholders and intended beneficiaries for the FAIR assessment framework.

3.4. Guiding principles

While an initial formulation of the assessment framework is due in October 2020 (see Annex A), in this initial report we would like to put forward some guiding principles and high-level requirements as currently perceived by this Group. We suggest that the FAIR assessment framework should:
- 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);[19]
- *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 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.

All of this with the ultimate goal, of course, of seeing adoption — both in name and in practice — by the European Open Science Cloud and associated communities.

Finally, as always when formulating measures and metrics, we will bear in mind Goodhart's law[20]— as history has demonstrated time and again that people will inevitably act to optimize for a target, rather than exhibit the kind of desirable behavior which the original measure (before it was made into a target) was meant to proxy for.

---

[19] 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.

[20] From Wikipedia (https://en.wikipedia.org/wiki/Goodhart%27s_law): "Goodhart's law is an adage named after economist Charles Goodhart, which has been phrased by Marilyn Strathern as "When a measure becomes a target, it ceases to be a good measure"."

4. Case studies

4.1. Rationale and process

To complement the literature review and initial formulation of guiding principles for a FAIR assessment framework for data services — mostly a top-down approach — we have also performed a more practical, bottom-up exercise, focused on the guiding principle 'Be rooted in FAIR data' (introduced in the section above), which we expect will support the development of a FAIR assessment methodology for services. Taking the form of case studies, we have selected three existing data services from the EOSC catalogue and aimed to answer the question: "How is this service 'enabling FAIR' in the sense of increasing the FAIRness of digital objects that it acts on?". Starting from this question, we reviewed the selected services in detail and performed a deliberately ad-hoc analysis, i.e. we did not formulate a formal assessment methodology upfront; rather we took this to be a "learning by doing" exercise: performing the assessment using prior experience and common sense, and afterwards reflect on the process and outputs, formalize learnings, and use that to fuel the development of a formal assessment methodology. As a consequence, two caveats about the results presented here are important to note:

1. The results were generated using a methodology that is still very preliminary. **Hence no conclusions should be drawn about the selected services on the basis of our <u>findings</u>; rather we hope that it will trigger discussion and subsequent refinement on the <u>methodology</u>**. Findings are, in fact, only included to illustrate the methodology.
2. Even where we find that a service has limited FAIR enablement, such an observation does not make any statements on the value of a service to its users or its quality.

With the above objectives and caveats in mind, we proceeded to select three service form the EOSC catalogue:

- Identifiers.org: a PID resolver an a registry for PID schemas
  https://marketplace.eosc-portal.eu/services/identifiers-org
- B2FIND: discovery service https://catalogue.eosc-portal.eu/service/eudat.b2find
- Clowder: research data management system
  https://catalogue.eosc-portal.eu/service/vi-seem.v-seem_clowder

This particular set was selected so as to cover different parts of the data life cycle; focus on services that are well-defined and not too complex; and have a certain level of maturity (which we expected will be able to give us data on usage).

Below we present the results of the analysis and, thus, a first iteration of an assessment methodology. Deferring further reflection to the last part of this section, there are two points worth emphasizing before presenting the results. First, while the point of departure was the question *"How is this service 'enabling FAIR' in the sense of increasing the FAIRness of digital objects that it acts on?"*, we found that a purely functional FAIR-based analysis of the service gives too limited a picture to understand how the service is used and where it adds value. We have therefore also included more general service-level attributes such as users, purpose, adoption, documentation, and quality in our descriptive framework.

Second, as to the analysis of FAIR enablement in a functional sense, we soon realized that this is often not a simple matter of 'yes' or 'no'. Instead, we found it helpful to map a service onto one of four modes:

- ***Enable*** (color-coded in green): the service actively helps to realize this particular FAIR principle — for example by adding metadata or enabling discoverability;
- ***Respect*** (in blue): the service does not actively 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*** (in red): 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 not applicable; in white): This particular FAIR principle is not relevant for the service, or there was insufficient information to determine if the FAIR principle applies.

4.2. Case Study 1: B2FIND

Service Summary

B2FIND[21] is a metadata aggregator. The service harvests metadata from different community repositories and harmonises them such that users and services can search through the combined metadata. B2 FIND offers a rich faceted graphical search interface and a HTTP REST API that has been implemented in python for EUDAT's B2FIND Training[22]

URL: http://b2find.eudat.eu/   **EOSC:** https://marketplace.eosc-portal.eu/services/b2find

---

Users

The service targets two types of user groups:
- Scientific communities that can provide their metadata and integrate via the B2FIND service with other metadata
- Scientists who can employ the service to search for interesting research data across different communities simultaneously.

Purpose

B2FIND is a metadata aggregator. It gathers metadata from communities and repositories and integrates the different types of metadata. It provides a graphical user interface and an API to present the metadata and allows faceted searches across the metadata corpus.

Services
- Metadata harvesting and harmonisation to communities with a tool to search across the metadata for scientists.
- The relevant metadata of a DO is shown and a link to the metadata provenance is provided.

Target Digital Objects
- Metadata entries

Examples
- B2FIND entry (KONTROL 1984[23])
- OAI-PMH dataset's metadata[24]

-

Adoption

By now B2FIND hosts 824566 metadata entries harvested from 22 communities. We were unable to establish from the documentation how many users use B2FIND.

Documentation

EUDAT provides guidelines on how to use the B2FIND services[25] as well as detailed guidelines for harvesting and mapping metadata[26]

FAIR enablement mapping (see Annex C for details)

F1	F2	F3	F4	A1	A1.1	A1.2	A2	I1	I2	I3	R1	R1.1	R1.2	R1.3

---

[21] http://b2find.eudat.eu/dataset

[22] https://github.com/EUDAT-Training/B2FIND-Training

[23] http://b2find.eudat.eu/dataset/6d886106-fa05-5cf8-9651-c8a570a0d2be

[24] https://easy.dans.knaw.nl/oai/?verb=GetRecord&metadataPrefix=oai_datacite&identifier=oai:easy.dans.knaw.nl:easy-dataset:157981

[25] https://eudat.eu/services/userdoc/b2find-usage

[26] http://b2find.eudat.eu/guidelines/index.html

4.3. Case Study 2: Identifiers.org

Service Summary

Identifiers.org is a resolution service and registry for identifiers in the life sciences, including identifiers from different domains. Providing access to data in a consistent manner and compact identifiers, assigned unique prefix and a local provider designated accession number (prefix:accession)[27].

URL: http://identifiers.org/   **EOSC:** https://marketplace.eosc-portal.eu/services/identifiers-org

---

Users

Identifiers.org was developed by EMBL-EBI as part of the ELIXIR infrastructure for the scientific community, focusing on the Life Sciences domain.

Purpose

The service handles persistent identifiers in the form of URIs and CURIEs. This allows the referencing of data in both a location-independent and resource-dependent manner. The provision of resolvable identifiers (URLs) fits well with the Semantic Web vision[31] and the Linked Data initiative [32].

Services

- Request prefix: register a service provider's identifier schema by requesting a prefix.
- Registry: search and browse
- Resolver

Target Digital Objects

- Identifiers and identifiers schemas

Non-functional aspects

- The identifiers.org services are transparent and open source on this repositories: https://github.com/identifiers-org

Examples

- An example to an identifier schema entry:[28]
- The prefix form[29] and A resolved DOI[30]

Adoption

Part of the Elixir infrastructure and used by many services[33].

Documentation

There is very clear documentation on how to use the identifiers.org services and API[34] [35]

| **FAIR enablement mapping** (see Annex C for details) |

F1	F2	F3	F4	A1	A1.2	A1.3	A2[36]	I1	I2	I3	R1	R1.1	R1.2	R1.3

---

27 https://registry.identifiers.org
28 https://registry.identifiers.org/registry/swh
29 https://registry.identifiers.org/prefixregistrationrequest
30 http://identifiers.org/resolve?query=doi:10.5281/zenodo.3630224
31 https://www.w3.org/2001/sw
32 http://linkeddata.org
33 https://en.wikipedia.org/wiki/Identifiers.org
34 https://docs.identifiers.org/articles/services.html
35 https://docs.identifiers.org/articles/api.html
36 Depends on context: 'reduce' for ID Digital Objects; 'enable' for ID schema's.

4.4. Case Study 3: Clowder

Service Summary

Clowder is a research data management system designed to support any data format and multiple research domains. When new data is added to the system, metadata is extracted and different preprocessing steps are run based on the type of the data, for example to create previews. This raw metadata is presented to the user in the Clowder web interface.

URL: http://dchrepo.vi-seem.eu/   **EOSC:** https://marketplace.eosc-portal.eu/services/v-seem-clowder

Users

The primary users are researchers. The service is open and free for everyone. The use of some functionalities requires an approved account. It is unclear what service levels and policies around account approval are.

Purpose

This is a data management system intended to help users pre-process, process and visualise data from any domain in an extensible and accessible manner. It functions more as a virtual research environment than a long-term repository for data publication.

Services

- Research data management system
- Preprocess, process and visualise data

Target Digital Object

- Active data, i.e. in processing stage

Non-functional aspects

- Support is provided via VI-SEEM Helpdesk (https://support.vi-seem.eu).
- Access to the helpdesk is open and does not require any authentication.
- VI-SEEM funding ended on 30 September 2018, it is unclear how long the service will stay available.

Examples

- Dataset from the VI SEEM Clowder[37]

Adoption

This instance of CLOWDER is designed for the cultural heritage community and is made available via the three-year Virtual Research Environment (VRE) in Southeast Europe and the Eastern Mediterranean (SEEM) project. It is currently adopted by 20+ users from the VI-SEEM Digital Cultural Heritage community. There are 59 datasets in 8 collections.

Documentation

- There is a pdf guide on a google drive, which is not a persistent way to share documentation
- The API is accessible here: https://clowder.ncsa.illinois.edu/swagger/ (but note the link from the service's site was broken at the time of review)

FAIR enablement mapping (see Annex C for details)

F1	F2	F3	F4	A1	A1.1	A1.2	A2	I1	I2	I3	R1	R1.1	R1.2	R1.3

4.5. Learnings and discussion

After carrying out the case studies presented above and reflecting on the process, a number of observations and lessons can be drawn to feed into the formulation of a more formal assessment framework.

First, in terms of definition and precision of scope for the assessment, we see a need for a formalized model to describe data services and the digital objects which they act on. The reason for this is that some services act on the bit stream of a digital object, while other services act on metadata or PID's. Hence, to clearly assess how a given service is FAIR-enabling requires an unambiguous way to describe the entity which the service acts on. As mentioned in the section "Towards an assessment framework for 'FAIRness of services'" (in particular Fig. 2) above, we expect that the FAIR Digital Object framework may provide a useful starting point.

Second, while the authors of course fully acknowledge the value and importance of FAIR, it is good to bear in mind that not all what is FAIR is good, and not all what is good is FAIR. For example:
- In our assessments, we have observed that the FAIR principles seem to work best for services at the end of the life-cycle.
- Each FAIR principle can be interpreted in a more liberal or strict sense, and hence the interpretation of the FAIR principle should be stated with each analysis to enable proper comparison across services and principles.
- The FAIR principles do not speak to aspects of service quality or sustainability, while these are clearly quintessential for the long-term value for researchers.

Finally, on a more practical note, we have found that it is often challenging to assess a service due to limited documentation and/or barriers to gain access. Such practical limitations are, at present, a reality that needs to be considered in any effort for large-scale assessment of services.

5. Conclusion

This report represents the first milestone of FAIRsFAIR Task 2.4 which is tasked with developing a framework for FAIR assessment of data services. Such an assessment framework should build on and extend the FAIR guiding principles for datasets and other digital objects to data services with the ultimate aim to support the development of a FAIR ecosystem as envisioned in the Turning FAIR into reality report (4) and recently reaffirmed in the interim recommendations from the EOSC FAIR WG (7,14)

With this report, we hope to have achieved four objectives. Firstly, when we began this work we thought it would be of value — also for the work to come — to create an overview of current FAIR assessment frameworks for datasets, services and other entities. In doing so we found that, while there is no shortage for FAIR assessment frameworks for datasets, fragmentation and scalability still pose significant challenges. We also found that FAIR assessment for data services (as opposed to datasets or other digital objects) is still rather premature, though there is existing work (e.g. on certification for data repositories and on quality criteria for EOSC-hub) that can be built upon.

Second, we set out to critically review the motivation for his task — in brief, the lack of guidance on how to develop or evolve data services for an optimal interplay with FAIR digital objects — and wanted to lay out our proposed approach to address this challenge. We explicitly formulated scope, purpose and a number of guiding principles to govern the development of a FAIR assessment framework for data services over the next 18 months. As the primary audience for this framework, we propose to focus on organizations that develop and provide data services because they are well-suited to affect change. With the framework, we intend to arm them with the insights and actionable recommendations that they need to make incremental improvements to their services that support the emergence of a FAIR ecosystem.

Thirdly, we wanted to add a more practical element to this report and start exploring FAIR assessment of services by doing. To this end we selected three data services from the EOSC catalogue and subjected them, as case studies, to a critical appraisal of 'FAIRness'. We did not agree on a fixed methodology upfront, but rather choose a 'learning by doing' approach: first perform ad-hoc analyses, then reflect and identify valuable elements of the approach, and finally abstract those into a more formal methodology.

The fourth and final objective for this report is to seek interaction and collaboration with other working groups and activities that operate in this space. This Group has connections with several initiatives, and has several others in its radar, but it is hard to maintain a full overview in this active and rapidly-changing environment. We hope that this report outlines our ambitions and starting points, and encourage others who find this work interesting to connect with us and provide feedback.

Over the course of the next 18 months, we will be working towards the publication of a *Framework for assessing FAIR services* at the close of the project in March 2022 (alongside a similar framework for FAIR software). This assessment framework should not only give an objective and transparent appraisal of the 'FAIRness' of a service, but also provide the service owner with actionable

recommendations to make it better. Ideally such an assessment framework will be (partially) automated though it is not clear if this will be feasible in the available time given the current state of play.

This report uncovered several open questions that will need to be addressed as we develop the assessment framework. In our view the most pertinent question at this time is: *How do we make sure that the assessment framework will really meet the needs of service owners, i.e. organizations that develop and provide data services?* Beyond this immediate question as to value for the intended users, there are a number of open questions related to definitions, representations and modelling: *How do we exactly define a 'data service' vis-à-vis other (non-data) services but also in relation to data resources, digital objects, etc.? Do we need a more formal architectural model to represent all building blocks and their interrelations[38]? Do we need a taxonomy of data services?* Suggestions and recommendations on how to address these questions from other tasks, working groups, initiatives or individuals will be most welcome as feedback to this report.

---

[38] Perhaps a 'FAIR Ecosystem Blueprint Architecture' akin to the AARC Blueprint Architecture (https://aarc-project.eu/architecture/) for federated access management?

Annex A: FAIRsFAIR Task 2.4 Statement of Work

FAIRsFAIR Task 2.4 (T2.4), as stipulated in the FAIRsFAIR project proposal, is tasked with *"extending the FAIR concept currently applied to data to the range of data services needed to enable and support FAIR data, and to software"*. This ambition naturally splits into two related but distinct topics, namely (i) 'FAIRness of services' and (ii) 'FAIR principles for (research) software'. T2.4 will be working on these topics alongside each other, seeking synergies where possible but also mindful of intrinsic differences that warrant a parallel approach.

In terms of the first of these two topics, 'FAIRness of services', T2.4 will be considering how services can make data (more) FAIR. This formulation respects and builds upon the original FAIR principles, which were articulated specifically for research data objects (and not for services or software). Taking these principles as starting point, T2.4 will be delivering an assessment framework that can be used to gauge how a given service acting on a data object makes that data object 'more FAIR', 'less FAIR' or 'equally FAIR'. In formulating such a framework, it is anticipated that some of the FAIR principles may apply to services as they do to data objects (e.g., "being registered and indexed in a searchable resource" under Findability). Equally there will be FAIR principles for data objects that do not translate to services, and there will be criteria for services that do not directly map onto one of the original FAIR principles (e.g., quality measures or warranties specific to services such as availability or trustworthiness). In other words, the directive of this task goes beyond a naïve mapping of the FAIR principles for data objects to services; rather, it aims to support an optimal interplay between services and research (data) objects to realize a 'FAIR ecosystem' (as articulated in Ref. (4)).

Two remarks are in order. First, the proposed assessment framework will be normative in the sense that it scores against a desired future state, i.e. it is constructed relative to a set of desired features and qualities for services to have. The task will thus also provide concrete, actionable recommendations for services to increase their level of 'FAIRness'. Such desiderata will be defined from community input about the current state (including good practices and current pain points), desired state, and recommendations to close the gap between current and desired state. The second remark is about scope: The task considers all services that create, read, update or delete data at any point in the data life cycle[39]. For the sake of focus, it will primarily concentrate on digital services with a strong IT component, i.e. strongly relying on technology to deliver value users.

The second objective of the task pertains to research software, i.e. software artefacts that are the output of a research activity. T2.4 will deliver recommendations on how to apply or adapt the FAIR principles, formulated for data objects, to software artefacts. Here it is expected that a naïve mapping of the original FAIR principles may already provide a useful starting point (in that such a straightforward application "may act as a guideline for those wishing to enhance the reusability of their research software holdings", wording adapted from Ref. (2)); yet adaptations and/or

---

[39] As working definition for the concept of a 'service' in the context of research data, we will adopt the formulation put forward by the ICSU-WDS/RDA Publishing Data Workflows WG (16): "A means of delivering value to the producers and users of digital objects by facilitating outcomes they want to achieve without the ownership of specific costs or risks"

extensions will likely be required to account for the special nature of research software (such as its dynamic nature with large numbers of versions and library dependencies). Also for this objective, community input about the current state, desired state and recommendations to close the gap will be central in formulating the recommendations.

While the subject matter and intended approach for both objectives within T2.4 have their differences, they are joined at the hip by the overarching ambition to create a 'FAIR ecosystem' which comprises of FAIR Digital Objects — including data and software — together with relevant services and infrastructure (4). This suggests that there will be ample connections and inter-related questions between the objectives, which justifies addressing them in parallel within the task.

The explicit connection between T2.4 and the notion of a 'FAIR ecosystem' also signals that the work carried out by the task will not stand in isolation. T2.4 will seek coordination and collaboration with a number of relevant projects and organization, including (but not limited to) the EOSC FAIR WG, GO FAIR, FAIRsharing, the joint FORCE11 & RDA Software Identification WG, the RDA Software Source Code IG, and FAIRsFAIR WP4 around FAIR certification.

Finally, the approach taken within T2.4 will be guided by the ambition to deliver concrete, reasonable and actionable outputs that are rooted in real-life problems and ready to be adopted by the various stakeholders – and hence will err on the side of 'progress' over 'perfection'.

Milestones & Deliverables

M12	Feb 2020	Milestone M2.7	Assessment report on FAIRness of services *(this report)*
M16	June 2020	Milestone TBC; pending review	Assessment report on FAIRness of research software
M20	Oct 2020	Milestone M2.10	Report on basic framework on FAIRness of services
M30	Aug 2021	Deliverable D2.7	Framework for assessing FAIR services

Annex B: FAIR assessment frameworks

This annex contains a set of 'cards' summarizing existing FAIR assessment frameworks, paying special attention to their positioning and purpose as well as the overall structure that was followed. This list builds on, and extends, a recent overview compiled by the RDA 'FAIR Data Maturity Model' WG (18) as well as the FAIRassist tool[40] developed by the FAIRsharing resource. While an effort was made to include the most relevant current assessment frameworks, this overview is not meant to be exhaustive. Suggestions for additions will be most welcome as feedback to this report.

B.1. FAIR assessment for datasets and/or FAIR digital objects

5-star Data Rating tool	
Created by: **CSIRO (OzNome)** Links: ● Checklist: http://oznome.csiro.au/5star/ ● Documentation: https://confluence.csiro.au/display/OZNOME/Data+ratings	
Scope: Datasets	**Elements**
Purpose: "To help users understand how mature some data or a service is."	● Published ● Hosted ● Curated ● Updated, maintained ● Licensed ● Citable ● Described ● Findable ● Loadable ● Useable ● Comprehensible ● Connected, linked ● Assessable ● Trusted
Target audience (if specified)	
Type: Online checklist	
Status: Live	

Checklist for Evaluation of Dataset Fitness for Use

Created by: **WDS/RDA Assessment of Data Fitness for Use WG**

Links:
- Checklist: Google docs[41]
- RDA WG webpage: https://www.rd-alliance.org/groups/assessment-data-fitness-use

	Elements
Scope: Datasets	- Dataset and Assessor Identification - Repository Certification - Curation - Metadata Completeness - Accessibility - Data completeness and correctness - Findability & interoperability
Purpose: "For use by repository managers or a similar external entity to evaluate the data holdings of repository for fitness for use". (".. meant to supplement the CoreTrustSeal Repository Certification process.")	
Target audience (if specified): Primarily for data librarians and IT staff; additionally for software engineers and researchers	
Type: Online checklist	
Status: Live	

FAIR data maturity model indicators

Created by: **RDA FAIR Maturity WG**

Links:
- Draft list of indicators: Google docs[42]
- RDA WG webpage: https://www.rd-alliance.org/groups/fair-data-maturity-model-wg

	Elements
Scope: Datasets	- Findable - Accessible - Interoperable - Reusable
Purpose: ".. (to develop) a common set of core assessment criteria for FAIRness and a generic and expandable self-assessment model for measuring the maturity level of a dataset."	
Target audience (if specified)	
Type: List of indicators (optional / recommended / mandatory)	
Status: Under development	

---

[41]

https://docs.google.com/forms/d/1p3iLmF_tSXWRy9LI66TjUcwdq5CacA-_cHL81Jx5gVE/viewform?edit_requested=true

[42]

https://docs.google.com/spreadsheets/d/1mkjElFrTBPBH0QViODexNur0xNGhJqau0zkL4w8RRAw/edit#gid=1325892715

FAIR enough?

Created by: **DANS**
Links:
- Checklist: [Google docs](https://docs.google.com/forms/d/e/1FAIpQLSf7t1Z9IOBoj5GgWqik8KnhtH3B819Ch6lD5KuAz7yn0I0Opw/viewform)[43]

	Elements
Scope: Datasets (and, to a minimal extent, data repositories)	- Trustworthiness of data repository - Findable - Accessible - Interoperable - Reusable - Openness
Purpose: To support researchers in depositing their data in a FAIR way	
Target audience (if specified): Researchers	
Type: Online checklist	
Status: Live	

FAIR metrics Gen1: Initial questionnaire

Created by: **FAIR metrics group**
Links:
- Questionnaire: [Github](https://github.com/FAIRMetrics/Metrics/blob/master/Evaluation_Of_Metrics/Supplementary%20Information_%20FM%20Evaluation%20Results.pdf)[44]
- Publication: A design framework and exemplar metrics for FAIRness (9)

	Elements
Scope: Digital resources, datasets	- Findable - Accessible - Interoperable - Reusable
Purpose: Intermediate step in the development of automated, scalable FAIR maturity indicators	
Target audience (if specified)	
Type: Self-assessment questionnaire	
Status: Published	

---

[43] https://docs.google.com/forms/d/e/1FAIpQLSf7t1Z9IOBoj5GgWqik8KnhtH3B819Ch6lD5KuAz7yn0I0Opw/viewform
[44]

https://github.com/FAIRMetrics/Metrics/blob/master/Evaluation_Of_Metrics/Supplementary%20Information_%20FM%20Evaluation%20Results.pdf

FAIR metrics Gen2: Bootstrap FAIR maturity indicators

Created by: **FAIR metrics group**

Links:

- Prototype evaluator software: https://w3id.org/AmIFAIR
- Publication: Evaluating FAIR maturity through a scalable, automated, community-governed framework (10)
- Documentation and code: https://github.com/FAIRMetrics/Metrics

Scope: Digital resources (including both data and software)	**Elements** Maturity Indicators and Compliance Tests that assess certain elements or specifications of the FAIR guidelines
Purpose: "(To provide) a roadmap for incremental improvements in the FAIRness of a resource"	
Target audience (if specified)	
Type: Automated tests	
Status: Live (initial release)	

FAIR self-assessment tool

Created by: **ARDC**

Links:

- Checklist: https://www.ands-nectar-rds.org.au/fair-tool
- Documentation: https://www.ands.org.au/working-with-data/fairdata/fair-data-self-assessment-tool

Scope: Datasets	**Elements** - Findable - Accessible - Interoperable - Reusable
Purpose: "... enables you to assess the 'FAIRness' of a dataset and determine how to enhance its FAIRness (where applicable)."	
Target audience (if specified): Primarily for data librarians and IT staff; additionally for software engineers and researchers	
Type: Online checklist	
Status: Live	

'FAIRdat' FAIR data assessment tool

Created by: **DANS**

Links:

- Checklist: https://www.surveymonkey.com/r/fairdat
- Documentation: Google docs[45] and
  http://blog.ukdataservice.ac.uk/fair-data-assessment-tool/

Scope: Datasets	**Elements** - Findable - Accessible - Interoperable
Purpose: ""To score the 'FAIRness' of a dataset."	
Target audience (if specified): "People who have recently re-used a dataset from a trusted repository and who now want to provide a review on their experience with the data. Additionally, archive/repository staff who wish to score a dataset are also able to use this document to establish the 'FAIRness' rating of a dataset."	
Type: Online checklist	
Status: Prototype	

FAIRshake

Created by: **Daniel Clarke et al. (funded by NIH)**

Links:

- Primary interface to toolkit: https://fairshake.cloud
- Publication: *FAIRshake: Toolkit to Evaluate the FAIRness of Research Digital Resources* (5)
- Documentation and code: https://github.com/MaayanLab/FAIRshake

Scope: Digital resources (including datasets, tools, repositories, APIs) / digital objects[46]	**Elements** Metrics and Rubrics (collections of FAIR metrics) and assessment capabilities for certain elements or specifications of the FAIR guidelines
Purpose: "... To promote the FAIRification of digital objects produced by research projects"	
Target audience (if specified): Biomedical research community	
Type: Toolkit that accommodates both manual and automated tests	
Status: Live	

---

[45] https://docs.google.com/document/d/1bRQDN_VFSPSMnsADLyzky-sbd6ZPArsHOcYhERdyrL8/edit
[46] The authors use both the terms 'digital object' and 'digital resource' in what appears to be an interchangeable way.

GARDIAN FAIR metrics

Created by: **CGIAR Platform for Big Data in Agriculture**

Links:

- Checklist: https://gardian.bigdata.cgiar.org/metrics.php#!/
- Documentation:
  http://aims.fao.org/activity/blog/put-fair-principles-practice-and-enjoy-your-data

Scope: Datasets	**Elements**
Purpose: Assessing the FAIRness of datasets	- Findable - Accessible - Interoperable
Target audience (if specified): Agriculture	
Type: Self-assessment checklist (scores between 0 and 16.5 for F, A and I)	Note: Reusable is not tested independently, but a score is assigned based on scores for F, A and I.
Status: Published	

How FAIR are your data?

Created by: **Sarah Jones & Marjan Grootveld for EUDAT**

Links:

- Checklist: How FAIR are your data? (19)

Scope: Datasets	**Elements**
Purpose: "... to discuss how FAIR the participant's research data were and what measures could be taken to improve FAIRness."	- Findable - Accessible - Interoperable - Reusable
Target audience (if specified): EUDAT summer school students	
Type: Self-assessment checklist (4 tick-boxes per element)	
Status: Published	

Simplified FAIR criteria assessment grid

Created by: **RDA-SHARC (Sharing Rewards and Credit) IG**

Links:

- Poster: Comment opérationnaliser et évaluer la prise en compte du concept 'FAIR' dans le partage des données: vers une grille simplifiée d'évaluation du respect des critères FAIR. (20)
- RDA IG webpage: https://www.rd-alliance.org/groups/sharing-rewards-and-credit-sharc-ig

Scope: Datasets	**Elements**
Purpose: "To establish if data are compliant to [sic] the FAIR principles"	● Findable ● Accessible ● Interoperable ● Reusable
Target audience (if specified)	
Type: Checklist (Never/NA; If Mandatory; Sometimes; Always)	
Status: First draft presented as poster	

Stewardship Maturity Matrix

Created by: **NOAA**

Links:

- Publication: A Unified Framework for Measuring Stewardship Practices Applied to Digital Environmental Datasets (21)
- See also: https://esip.figshare.com/articles/MM-Serv_ESIP_2018sum_v2r1_20180709_pdf/6855020

Scope: Datasets (data products)	**Elements**
Purpose: "It can be used to create a stewardship maturity scoreboard of dataset(s) and a roadmap for scientific data stewardship improvement or to provide data quality and usability information to users, stakeholders, and decision makers."	● Preservability ● Accessibility ● Usage ● Production Sustainability ● Data Quality Assurance ● Data Quality Control / Monitoring ● Data Quality Assessment ● Transparency / Traceability ● Data Integrity
Target audience (if specified): Environmental sciences	
Type: Maturity matrix	
Status: Published	

<table>
<tr>
<td colspan="2">

WMO Stewardship Maturity Matrix for Climate Data (SMM-CD)

Created by: **The SMM-CD Working Group**

Links:
- Maturity Matrix: The template for the WMO-Wide Stewardship Maturity Matrix for Climate Data (22).
- Documentation: The guidance booklet on the WMO-Wide Stewardship Maturity Matrix for Climate Data (23)

</td>
</tr>
<tr>
<td>

Scope: Datasets

Purpose: "enable data providers to assess and rate their datasets quantifiably based on internationally validated data stewardship best practices."

Target audience (if specified): Climate data

Type: Maturity matrix (five levels per subelement)

Status: Published

</td>
<td>

Elements

- Data Access
- Usability & Usage
- Quality Management
- Data Management

</td>
</tr>
</table>

B.2. FAIR assessment for data repositories and other data services

<table>
<tr>
<td colspan="2">

CoreTrustSeal

Created by: **DSA–WDS Partnership Working Group on Repository Audit and Certification, a Working Group (WG) of the Research Data Alliance**

Links:
- Requirements and documentation: https://www.coretrustseal.org/why-certification/requirements/
- RDA WG webpage: https://www.rd-alliance.org/groups/repository-audit-and-certification-dsa%E2%80%93wds-partnership-wg.html

</td>
</tr>
<tr>
<td>

Scope: Data repositories

Purpose: "... ensuring the reliability and durability of data repositories and hence the potential for sharing data over a long period of time."

Target audience (if specified): Data repositories

Type: List of requirements (compliance levels & narrative), self-assessment followed by external review for certification

Status: Published

</td>
<td>

Elements

- Organizational infrastructure
- Digital Object Management
- Technology

</td>
</tr>
</table>

Criteria for possible inclusion in the EOSC Service Portfolio

Created by: **EGI**
Links:
- Draft list of criteria: EOSC-hub wiki[47]

Scope: Data repositories	**Elements**
Purpose: "To understand which services should be onboarded (to the EOSC service portfolio"	(n/a)
Target audience (if specified): Data service providers	
Type: List of requirements	
Status: Under development	

Data Repository Selection: Criteria That Matter

Created by: **Susanna-Assunta Sansone and others (including FAIRsharing, DataCite & Publishers)**
Links:
- Criteria and documentation: Data Repository Selection: Criteria That Matter (24)
- Project webpage: https://osf.io/n9qj7/

Scope: Data repositories	**Elements**
Purpose: ".. a set of criteria that journals and publishers believe are important for the identification and selection of data repositories, which can be recommended to researchers when they are preparing to publish the data underlying their findings."	Essential: ● Repository Status ● Data Access Conditions ● Data Reuse Conditions ● Data Deposition Condition ● Data Preservation Policy ● Persistent Identifiers for Data ● User Support (plus 11 'Desirable' criteria)
Target audience (if specified): Researchers, journals, publishers	
Type: Set of criteria	
Status: Under development	

---

[47] https://wiki.eosc-hub.eu/display/EOSC/Criteria+for+possible+inclusion+in+the+EOSC+Service+Portfolio

ELIXIR Core Data Resource indicators

Created by: **ELIXIR**

Links:
- Indicators and documentation: Identifying ELIXIR Core Data Resources (11)
- See also: https://elixir-europe.org/platforms/interoperability/rir-selection

	Elements
Scope: Data services	• Scientific focus and quality of science
Purpose: " .. (to) guide and inform the managers of Emerging Services in the development of their Resource towards an 'ELIXIR Service' status."	• Community served by the resource • Quality of service • Legal and funding infrastructure, and governance
Target audience (if specified): Life Sciences	• Impact and translational stories.
Type: List of indicators (qualitative and quantitative)	
Status: Published	

SHARP services

Created by: **EGI**

Links:
- Presentation:
  https://repository.eoscsecretariat.eu/index.php/s/XLXpPsJ5HZGqq3M#pdfviewer

	Elements
Scope: Data services	• Sustainable
Purpose: "To professionalize IT services in support of enabling and supporting open data for EOSC"	• Helpful • Accessible • Reusable
Target audience (if specified)	• Professional
Type	
Status: Under development	

B.3. FAIR assessment for data stewardship, processes and organizational maturity

---

Data stewardship wizard

Created by: **ELIXIR, DTL and others**

Links:

- DZ wizard tool: https://demo.ds-wizard.org/questionnaire
- Documentation and more information: https://www.dtls.nl/2017/11/11/data-stewardship-wizard-enhance-quality-data-management-plan

	Elements
Scope: Data management processes	- Administrative detail - Re-using data - Creating and collecting data - Processing data - Interpreting data - Preserving data - Giving access to data
Purpose: "The Data Stewardship Wizard converts a lengthy data management questionnaire into an effective flowchart, saving you research time and money, and enhancing the quality of your Data Management Plan."	
Target audience (if specified)	
Type: Online questionnaire (wizard)	
Status: Live	

FAIR Capability Maturity Model

Created by: **FAIRsharing**

Links:

- Presentation Google drive[48]

	Elements
Scope: Organizations	- FAIR data sets - FAIRification process - Foundational components
Purpose: "FAIR CMM is a tool to develop, assess and refine a strategic FAIR data transformation program for organizations."	
Target audience (if specified): Organizations, projects and teams	
Type	
Status: Under development	

---

[48] https://drive.google.com/file/d/16QaBaoGXLJnJtZTjWlzWvw_Pkwneu8Z9/view

Annex C: FAIR assessment case studies

C.1. Analysis according to the FAIR principles

This annex contains the detailed FAIR assessment conducted on the three services that were selected as case studies: B2FIND, identifiers.org and Clowder. Each service was analyzed on how it serves the digital object that it acts on in respect to the FAIR principles, i.e. we set out to answer the question: *"How is this service 'enabling FAIR' in the sense of increasing the FAIRness of digital objects that it acts on?"*

We divided the analysis to four tables for each service (Findable, Accessible, Interoperable and Re-usable) and, for each of the FAIR principles, mapped the functionality of the service with regard to that principle to one of four modes: 'Enable', 'Respect', 'Reduce' or 'N/A' as described in the section 'Case studies' in the main text.

C.2. FAIR Enablement mapping: B2FIND

<table>
<tr><td colspan="5" align="center">B2FIND / Findable</td></tr>
<tr><td>FAIR: Findable</td><td>reduce</td><td>respect</td><td>enable</td><td>comment</td></tr>
<tr><td>F1. (meta)data are assigned a <u>globally unique and eternally persistent identifier.</u></td><td></td><td>F1</td><td></td><td>B2FIND relies on the harvested repositories to attach a PID to its records and expose that as part of the metadata. If the provided metadata contains a PID B2FIND represents this PID and uses it to link the harvested metadata to the original data object in the repository.</td></tr>
<tr><td>F2. data are described with <u>rich metadata.</u></td><td></td><td></td><td>F2</td><td>B2FIND relies on the provided metadata by the harvested repositories. B2FIND does not further enrich metadata.
enhances F2 with citation metadata</td></tr>
<tr><td>F3. (meta)data are <u>registered or indexed in a searchable resource.</u></td><td></td><td></td><td>F3</td><td>By being on B2FIND the (meta)data is registered and indexed and can be found through search.</td></tr>
<tr><td>F4. metadata <u>specify</u> the data identifier.</td><td></td><td></td><td>F4</td><td>In the current data infrastructure landscape B2FIND takes the position of a metadata registry and indexer. It extends the repository service with a good searching interface.</td></tr>
<tr><td colspan="5" align="center">B2FIND / Accessible</td></tr>
</table>

FAIR: Accessible	reduce	respect	enable	comment
A1 (meta)data are <u>retrievable by their identifier</u> using <u>a standardized communications protocol.</u>		A1		If repositories mention a PID in their metadata record, B2FIND uses that identifier to link back to the original data object and thus fulfills A1 (see also F2/F3). B2FIND relies on that metadata to be accessible.
A1.1 the <u>protocol</u> is open, free, and universally implementable.			A1.1	B2FIND relies on a REST API, using OAI-PMH (an open standard)
A1.2 the <u>protocol</u> allows for an authentication and authorization procedure, where necessary.			A1.2	The API allows for an authentication for uploading https://github.com/EUDAT-Training/B2FIND-Training/blob/master/uploading.py
A2 <u>metadata are accessible</u>, even when the data are no longer available.			A2	Once B2FIND harvested, i.e. copied metadata from repositories, the service can potentially keep the metadata independently from whether the repository deletes the original data object including all metadata. In that sense, B2FIND serves A2 in the data infrastructure. However, if metadata is kept and for how long is subject to SLAs and single agreements between the repositories agreeing to being harvested and the service B2FIND.

B2FIND / Interoperable

FAIR: Interoperable	reduce	respect	enable	comment
I1. (meta)data use a <u>formal, accessible, shared, and broadly applicable language</u> for knowledge representation.			I1	Metadata can be explored through a graphical user interface or by APIs. The response format is JSON.
I2. (meta)data use <u>vocabularies that follow FAIR principles.</u>			I2	B2FIND maps harvested metadata to Dublin core, OAI-PMH and more. Upon harvesting B2FIND checks the validity of metadata and evaluates whether it conforms to the above mentioned standards. B2FIND in general improves the interoperability of metadata coming from different repositories by mapping them to the same standards and search facets.

| | | | I3 | If references exist in the metadata |
| I3. (meta)data include <u>qualified references</u> to other (meta)data. | | | I3 | If references exist in the metadata standards in a PID form, B2FIND will produce the direct link to the resource |

B2FIND / Re-usable				
FAIR: Re-usable	**reduce**	**respect**	**enable**	**comment**
R1. meta(data) are richly described with a <u>plurality of accurate and relevant attributes.</u>		R1		Only if repository provides them or if they are subject to the standards mentioned in I2.
R1.1. (meta)data are released with a clear and accessible data usage license.		R1.1		Only if the repository provides the license or usage rights.
R1.2. (meta)data are associated with their <u>provenance.</u>			R1.2	B2FIND tracks the provenance of the metadata
R1.3. (meta)data <u>meet domain-relevant community standards.</u>		R2		It depends on the repository and metadata entry. Entries can meet community standards if they were before the ingestion by B2FIND

C.3. FAIR Enablement mapping: Identifiers.org

Identifiers.org / Findable				
FAIR: Findable	**reduce**	**respect**	**enable**	**comment**
F1. (meta)data are assigned a <u>globally unique and eternally persistent identifier.</u>			F1	Identifiers.org has only entries that have an identifier.It is a registry for identifier schema and a resolver for known schemas.
F2. data are described with <u>rich metadata.</u>		F2		Identifiers.org resolves an identifier to a metadata page where only the identifier's provenance metadata are accessible. It does not manipulate the metadata, but it doesn't show all metadata associated with a resolved PID.
F3. (meta)data are <u>registered or indexed in a</u>			F3	The identifier or the identifier schema is searchable in this case, the data is the

searchable resource.				identifier and identifier schema)
F4. metadata specify the data identifier.			F4	Metadata do link to the resource. The metadata, when a PID is resolved, will be the specific location of the data.

Identifiers.org / Accessible

FAIR: Accessible	reduce	respect	enable	comment
A1 (meta)data are retrievable by their identifier using a standardized communications protocol.			A1	Identifiers.org as a resolver does retrieve the identified object to the user.
A1.1 the protocol is open, free, and universally implementable.			A1.1	Identifiers.org provides a REST API, using the HATEOAS[49] standard (Hypermedia as the Engine of Application State) Open, free and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary.		A1.3		There is no authentication, even on prefix request, where only requester details are asked, but this doesn't reduce accessibility.
A2 metadata are accessible, even when the data are no longer available.	A2 for ID DO		A2 for ID schemas	It seems that the metadata is not kept and archived by identifiers.org and thus. if the identified resource is no longer available, the identifier won't be resolvable and no metadata will be accessible. On the other hand, identifiers schemas that aren't in use, will be visible with the browse and search identifiers schemas service.

Identifiers.org / Interoperable

FAIR: Interoperable	reduce	respect	enable	comment
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.			I1	Metadata can be explored through a graphical user interface or by APIs. The response format is JSON.

---

[49] https://restfulapi.net/hateoas/

I2. (meta)data use <u>vocabularies that follow FAIR principles.</u>	I2			Metadata isn't using a standard vocabulary, which might reduce interoperability of metadata
I3. (meta)data include <u>qualified references</u> to other (meta)data.		I3		If the identifier schema has a home url or optional references, it will be referenced in the schema entry and it is qualified metadata

<table>
<tr><td colspan="5" align="center">Identifiers.org / Re-usable</td></tr>
<tr><td>FAIR: Re-usable</td><td>reduce</td><td>respect</td><td>enable</td><td>comment</td></tr>
<tr><td>R1. meta(data) are richly described with a <u>plurality of accurate and relevant attributes.</u></td><td></td><td>R1</td><td></td><td>A fixed set of metadata is asked in the request prefix form.</td></tr>
<tr><td>R1.1. (meta)data are released with a clear and accessible data usage license.</td><td>R1.1</td><td></td><td></td><td>No license is requested on the schema identifier request and the license of a DO which resolved by a PID is not accessible on identifiers.org</td></tr>
<tr><td>R1.2. (meta)data are associated with their <u>provenance.</u></td><td></td><td></td><td>R1.2</td><td>Identifiers.org tracks the provenance of resolved PID</td></tr>
<tr><td>R1.3. (meta)data <u>meet domain-relevant community standards.</u></td><td></td><td>R1.3</td><td></td><td>The identifiers schemas are a product of domain community standards, which resulted in a specific PID registered in identifiers.org. Even though the identifier registry does not verify compliance with community standards.</td></tr>
</table>

C.4. FAIR Enablement mapping: Clowder

<table>
<tr><td colspan="5" align="center">Clowder / Findable</td></tr>
<tr><td>FAIR: Findable</td><td>reduce</td><td>respect</td><td>enable</td><td>comment</td></tr>
<tr><td>F1. (meta)data are assigned a <u>globally unique and eternally persistent identifier.</u></td><td></td><td>F1</td><td></td><td>The service applies checksums per file but not PIDs as far as we can tell; the system assigns an internal UUID that might be globally unique</td></tr>
</table>

F2. data are described with <u>rich metadata.</u>	N/A			We have only found examples of metadata added by the owner and this is not extensive. There doesn't seem to be any examples of extraction. The software does allow for extending the basic metadata schema with rich tailored metadata, as search is restricted it's unclear if any additional metadata can be added.
F3. (meta)data are <u>registered or indexed in a searchable resource.</u>	N/A			As F2.
F4. metadata <u>specify</u> the data identifier.	N/A			API is not working so metadata is unlikely to be harvested by a third party source.

Clowder / Accessible	

FAIR: Accessible	**reduce**	**respect**	**enable**	**comment**
A1 (meta)data are <u>retrievable by their identifier</u> using <u>a standardized communications protocol.</u>	N/A			API is not available so it is unclear if any of the identifiers can be used to retrieve metadata.
A1.1 the <u>protocol</u> is open, free, and universally implementable.	N/A			N/A
A1.2 the <u>protocol</u> allows for an authentication and authorization procedure, where necessary.	N/A			N/A There is an authentication mechanism following the account creation. Unfortunately, the requested accounts are still on standby so that the protocols couldn't be tested.
A2 <u>metadata are accessible</u>, even when the data are no longer available.	A2			Terms of Service explicitly state that they may shut down the service at any time with no alternatives given to host the metadata or data

Clowder / Interoperable	

FAIR: Interoperable	**reduce**	**respect**	**enable**	**comment**

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.		I1		Metadata schema itself is JSON pulling in Dublin Core and can be extended to suit a services needs: http://sead-data.net/managing-metadata-terms. For the instance tested by us, as searching is only accessible when logged in, it was not clear whether any additional standardised terms are used.
I2. (meta)data use vocabularies that follow FAIR principles.	N/A			No information about any ontologies provided
I3. (meta)data include qualified references to other (meta)data.	N/A			Only metadata field allowing for references is an unqualified "Related Publications", "uri":"http://sead-data.net/terms/Related Publications"

Clowder / Re-usable

FAIR: Re-usable	reduce	respect	enable	comment
R1. meta(data) are richly described with a plurality of accurate and relevant attributes.	N/A			Publicly accessible metadata is very limited
R1.1. (meta)data are released with a clear and accessible data usage license.	R1.1			All examples for datasets that were used in this assessment were licensed "all rights reserved" and not available for download, so overall, reduce is the most appropriate review.
R1.2. (meta)data are associated with their provenance.	N/A			Publicly accessible metadata is very limited
R1.3. (meta)data meet domain-relevant community standards.	N/A			

References

1. Hagstrom S. The FAIR Data Principles [Internet]. FORCE11. 2014 [cited 2020 Feb 17]. Available from: https://www.force11.org/group/fairgroup/fairprinciples
2. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar 15;3(1):160018.
3. Mons B, Neylon C, Velterop J, Dumontier M, da Silva Santos LOB, Wilkinson MD. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. Inf Serv Use. 2017 Jan 1;37(1):49–56.
4. 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. [Internet]. Brussels: Directorate-General for Research and Innovation (European Commission); 2018 Nov [cited 2019 Nov 19]. Available from: https://doi.org/10.2777/1524
5. Clarke DJB, Wang L, Jones A, Wojciechowicz ML, Torre D, Jagodnik KM, et al. FAIRshake: Toolkit to Evaluate the FAIRness of Research Digital Resources. Cell Syst. 2019 Nov 27;9(5):417–21.
6. L'Hours H, von Stein I. FAIR Ecosystem Components: Vision [Internet]. FAIRsFAIR project; 2019 Dec [cited 2020 Jan 22]. Available from: https://doi.org/10.5281/zenodo.3565428
7. Genova F, Aronsen JM, Beyan O, Harrower N, Holl A, McIntosh L, et al. Interim recommendations on certifying the services required to enable FAIR research outputs within EOSC [Internet]. 2020. Available from: https://repository.eoscsecretariat.eu/index.php/s/zCnHTcytBHaLjRp#pdfviewer
8. Bangert D, Hermans E, van Horik R, de Jong M, Koers H, Mokrane M. Recommendations for Services in a FAIR data ecosystem [Internet]. 2019. Available from: https://doi.org/10.5281/zenodo.3585742
9. Wilkinson MD, Sansone S-A, Schultes E, Doorn P, Bonino da Silva Santos LO, Dumontier M. A design framework and exemplar metrics for FAIRness. Sci Data. 2018 Jun 26;5(1):1–4.
10. Wilkinson MD, Dumontier M, Sansone S-A, Bonino da Silva Santos LO, Prieto M, Batista D, et al. Evaluating FAIR maturity through a scalable, automated, community-governed framework. Sci Data. 2019 Dec;6(1):174.
11. Durinx C, McEntyre J, Appel R, Apweiler R, Barlow M, Blomberg N, et al. Identifying ELIXIR Core Data Resources [version 2; peer review: 2 approved]. F1000research [Internet]. 2017 [cited 2020 Feb 2];5 (ELIXIR)(2422). Available from: https://f1000research.com/articles/5-2422/v2#B1
12. Dunning A, De Smaele M, Böhmer J. Are the FAIR Data Principles fair? Int J Digit Curation. 2017;12(2):177–95.
13. Hahnel, Mark, Valen, Dan. How to (easily) extend the FAIRness of existing repositories. Data Intell. 2020;2(1):192–8.
14. Genova F, Aronsen JM, Beyan O, Harrower N, Holl A, McIntosh L, et al. Interim recommendations on FAIR Metrics for EOSC [Internet]. 2020. Available from: https://repository.eoscsecretariat.eu/index.php/s/C3a5WkpsFHL6GD3#pdfviewer
15. Hodson S, Jones S, Collins S, Genova F, Harrower N, Mietchen D, et al. FAIR Data Action Plan: Interim recommendations and actions from the European Commission Expert Group on FAIR data. 2018 Jun 7 [cited 2020 Jan 22]; Available from: https://doi.org/10.5281/zenodo.1285290
16. Dallmeier-Tiessen S, Khodiyar V, Murphy F, Nurnberger A, Raymond L, Whyte A. Connecting Data Publication to the Research Workflow: A Preliminary Analysis. Int J Digit Curation.

2017;12(1):88–105.

17. Wittenburg P. From Persistent Identifiers to Digital Objects to Make Data Science More Efficient. Data Intell. 2019 Mar 1;1(1):6–21.

18. Bahim C, Dekkers M, Wyns B. Results of an Analysis of Existing FAIR Assessment Tools [Internet]. Research Data Alliance; 2019 May [cited 2020 Jan 22]. Available from: https://doi.org/10.15497/RDA00035

19. Jones S, Grootveld M. How FAIR are your data? [Internet]. Zenodo; 2017. Available from: https://doi.org/10.5281/zenodo.1065991

20. David R, Mabile L, Yahia M, Cambon-Thomsen A, Archambeau A-S, Bezuidenhout L, et al. Comment opérationnaliser et évaluer la prise en compte du concept 'FAIR' dans le partage des données: vers une grille simplifiée d'évaluation du respect des critères FAIR. [Internet]. Zenodo; 2018. Available from: https://doi.org/10.5281/zenodo.1995646

21. Peng, Ge, Privette JL, Kearns EJ, Ritchey, Nancy A., Ansari, Steve. A Unified Framework for Measuring Stewardship Practices Applied to Digital Environmental Datasets. Data Sci J. 13:231–53.

22. Lief C, Peng, G. The template for the WMO-Wide Stewardship Maturity Matrix for Climate Data [Internet]. [cited 2020 Feb 2]. Available from: https://doi.org/10.6084/m9.figshare.7003709.v6

23. Group TS-CW. The guidance booklet on the WMO-Wide Stewardship Maturity Matrix for Climate Data [Internet]. figshare; 2019. Available from: https://doi.org/10.6084/m9.figshare.7002482.v6

24. Sansone S-A, McQuilton P, Cousijn H, Cannon M, Chan WM, Carnevale I, et al. Data Repository Selection: Criteria That Matter [Internet]. [cited 2020 Feb 2]. Available from: https://osf.io/m2bce/