CORETRUSTSEAL-CERTIFIED REPOSITORIES

Enabling Findable, Accessible, Interoperable, and Reusable (FAIR) Data

Mustapha Mokrane
Data Archiving and Networked Services
Netherlands
mustapha.mokrane@dans.knaw.nl
https://orcid.org/0000-0002-0925-7983

Jonas Recker
GESIS Data Archive for the Social Sciences
Germany
jonas.recker@gesis.org
https://orcid.org/0000-0001-9562-3339

Abstract – The preservation of research data to enable replication and reuse is critically dependent on efficient, effective and sustainable data stewardship by the research communities. The certification of trustworthy data repositories (TDRs) – custodian organizations that ensure data stewardship and long-term preservation – by means of a standard such as the CoreTrustSeal is an established and recognized procedure to support long-term access to reusable data. Likewise, the FAIR Guiding Principles and the developing FAIR metrics have largely codified the contemporary discourse and policies on research data management and stewardship. The proximity of objectives between the CoreTrustSeal certification of TDRs and the implementation of FAIR Principles calls for a close examination of their overlaps and complementarities. In particular, the concept of FAIR data cannot be detached from the characteristics of the data infrastructure, the environment in which FAIR data objects reside. It is therefore necessary to examine, under which circumstances the assessment of FAIRness should be carried out at collection- or repository-level, and to what extent CoreTrustSeal certification can be considered positioning TDRs as enabling FAIR data.

Keywords – FAIR Data, Data Preservation, Trustworthy Data Repositories, Certification.

Conference Topics – Exploring New Horizons; Building Capacity, Capability and Community

I. Introduction

Data repositories are key research infrastructures entrusted with the mission of managing research data assets and preserving their usefulness by ensuring accessibility, understandability and reusability over time. By deploying both human and technical capacities for data stewardship, repositories play a critical role in enabling reproducibility of research and data reuse for future discoveries. The mission of research data repositories is thus strongly aligned with the *FAIR Guiding Principles for Scientific Data Management and Stewardship* [1].

The FAIR Principles were published as a set of high-level aspirational principles describing four characteristics that data assets, tools, vocabularies and infrastructures should exhibit throughout their entire lifecycle: Findability, Accessibility, Interoperability, and Reusability. However, the principles do not explicitly describe how to achieve compliance or how to measure 'FAIRness'. While some FAIR principles address characteristics which are dependent on the (technical) environment in which a data object is stored and accessed (e.g. F1 on data identifiers or A1 on communication protocols) and thus can be usefully (and sometimes only) assessed at the level of the repository, other principles require a more detailed assessment at the level of the dataset. To support the measuring of FAIRness of data objects several ongoing initiatives have begun to explore the definition of FAIR metrics [2]. Increasingly, the FAIR Principles are also recognized and used as a benchmark to develop and improve research data infrastructure for maximizing the reuse of scholarly data.

In the *Turning FAIR into Reality* report [3] the European Commission Expert Group on FAIR Data states that "[t]he FAIR principles focus on access to the data and do not explicitly address the long-term preservation needed to ensure that this access endures. Data should be stored in a trusted and sustainable digital repository to provide reassurances about the standard of stewardship and the commitment to preserve" (p. 22). Accordingly, an important role of trustworthy repositories in this ecosystem is the provision of long-term stewardship of FAIR data objects, including curation activities to ensure that the data objects remain FAIR. This entails support and guidance for the data producers – e.g. advice on which metadata standards should be used – as well as for the data users. For the latter, a trustworthy repository serves as a guarantor that the data they download remain citable, accessible, and usable for the long term.

In this light, the importance of certification of repositories as trustworthy is twofold with regard to FAIR: Firstly, it can demonstrate to users that the repository enables FAIR data; secondly, certification of repositories as trustworthy may serve as a baseline for the evaluation of the FAIRness of datasets – namely if we assume that a correlation exists between a repository's trustworthiness, i.e. its demonstrated sustainability and capacity to perform data management appropriately, and its capacity to enable FAIR data.

The *Core Trustworthy Data Repositories Requirements* [4] were published as a universal catalog defining the minimum capacities research data repositories should achieve and the characteristics they must exhibit to be recognized as trustworthy. Because they were designed to assess the quality and performance of data management practices and compliance with internationally recognized and community-adopted standards, these requirements share a common objective and spirit with the FAIR Guiding Principles. The CoreTrustSeal certification based on these requirements provides a structured assessment of data repositories' trustworthiness. It is both a measure of sustainability and soundness of a data repository as an organization and a measure of its technical and technological reliability. The assessment also covers the management of digital objects in the repository and therefore sheds light on the overall quality of the data holdings. This certification scheme supported by a community effort (i.e. the CoreTrustSeal Foundation) is operational and open to data repositories worldwide. Over 140 data repositories have already been successfully certified as trustworthy by CoreTrustSeal and its precursors, the Data Seal of Approval and the World Data System Regular Members Certification [5].

It can be expected that if a data repository fulfills the CoreTrustSeal requirements, the data it holds will also meet a number of the FAIR criteria. Thus, the CoreTrustSeal certification may provide a good basis to assess FAIR compliance of datasets, at least for the FAIR principles directly linked to characteristics or capacities exhibited by the repositories holding the data. It may also provide a good proxy to assess compliance with other FAIR principles. Moreover, the CoreTrustSeal Requirements address other aspects such as maintaining the understandability and reusability of datasets over time (data curation and stewardship) which are not covered by the FAIR Guiding Principles but are extremely important to maintain the FAIRness of a data object.

While there is clearly some overlap between CoreTrustSeal requirements and FAIR Principles (see section IV. below), there is not yet a full understanding of the extent to which we can assume that data held by a CoreTrustSeal-certified repository comply with the FAIR principles. In this paper we therefore explore the extent to which the CoreTrustSeal certification can serve as a baseline to assess FAIR compliance of datasets and infrastructure. For this purpose, we will look at the FAIR Guiding Principles and their assessment in particular with relation to the sustainability and long-term preservation of datasets. We will then analyze CoreTrustSeal Requirements for Trustworthy Data Repositories (TDRs), their associated certification procedure and the relationship with the FAIR Principles. Subsequently a mapping between CoreTrustSeal Requirements and the FAIR Principles will be presented to discuss where and how they overlap. Based on this we will investigate the extent to which the CoreTrustSeal certification conducted at the level of the data repository can be used to assess the implementation of FAIR Principles at the level of datasets. We will conclude with considerations on the review of the CoreTrustSeal Requirements and how this process can incorporate the outcomes of relevant FAIR initiatives.

II. FAIR Guiding Principles and their Assessment

The FAIR Guiding Principles for scientific data management and stewardship define the characteristics that data resources on the one hand, and the tools, vocabularies, and infrastructures used for data management and sharing on the other should exhibit to assist discovery and reuse by third parties. As they were by design defined at a high level, the principles do not include an implementation framework nor an approach to assess datasets' FAIRness. The principles have a data-centric approach and the main focus in the current discussions about FAIR metrics and FAIR assessment is on the data objects without sufficient attention given to the characteristics of the environment in which the data are held, in particular data repositories. Thus, several initiatives have begun to develop tools to assess the FAIRness of datasets.[1] The FAIR Metrics Group has proposed metrics for the assessment of a dataset's FAIRness, with a strong emphasis on semi-automated

[1] The RDA FAIR Data Maturity Model Working Group created an overview of existing FAIR assessment tools and approaches (see https://docs.google.com/spreadsheets/d/14ojMSXVOITg3RoJn-PuDaPj8zuIGQz2Li-kl97HOBH4/edit).

procedures, as well as a framework for the definition of new, community-specific metrics [2]. Further discussion is now required to define an agreed core set of metrics as well as to come to an agreement about community-specific metrics suitable to measure the FAIRness of specific data types.

Characteristics that determine the FAIRness of a data object and which thus become the focus of an assessment can be intrinsic or extrinsic to the object, i.e. they can either be an integral part of the object, or the object derives these characteristics from the infrastructure in which it resides. For example, to check compliance with the FAIR Principle F1 (see Table I) we can either look at the data object itself to see if there is a unique and resolvable PID attached to it (e.g. as part of the metadata). However, it can also be verified globally for an entire class or collection of data objects by checking if the infrastructure holding the dataset assigns PIDs to its data assets. By contrast, FAIR Principle F4 refers to a characteristic extrinsic to the data object. It can only be verified by ascertaining that the infrastructure holding the data object implements such a registry or index. It follows that for some of the principles the FAIRness of a data object can be assessed at repository-level by looking at the policies and standards employed by the infrastructure holding the data object.

TABLE I
THE FAIR GUIDING PRINCIPLES. SOURCE: [1]

To be Findable:	
F1.	(meta)data are assigned a globally unique and persistent identifier
F2.	data are described with rich metadata (defined by R1 below)
F3.	metadata clearly and explicitly include the identifier of the data it describes
F4.	(meta)data are registered or indexed in a searchable resource
To be Accessible	
A1.	(meta)data are retrievable by their identifier using a standardized communications protocol
A1.1.	the protocol is open, free, and universally implementable
A1.2.	the protocol allows for an authentication and authorization procedure, where necessary
A2.	metadata are accessible, even when the data are no longer available
To be Interoperable:	
I1.	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
I2.	(meta)data use vocabularies that follow FAIR principles
I3.	(meta)data include qualified references to other (meta)data
To be Reusable:	
R1.	meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1.	(meta)data are released with a clear and accessible data usage license
R1.2.	(meta)data are associated with detailed provenance
R1.3.	(meta)data meet domain-relevant community standards

Assuming that the FAIRness of a data object can be assessed based on the FAIR Guiding Principles, if the data object meets these principles it is assigned a high score at the time of assessment. However, as Mari Kleemola points out, “[r]esearch data will not become nor stay FAIR by magic. We need skilled people, transparent processes, interoperable technologies and collaboration to build, operate and maintain research data infrastructures” [6]. That said, the current FAIR Principles neither cover data stewardship activities such as curation and long-term preservation nor strategies and procedures to promote the sustainability of the data repository, all of which ensure that the data objects remain FAIR over time. Consequently, the FAIRness score of a data object could decay over time and should be time stamped and updated regularly. An example is a data object to which a Digital Object Identifier (DOI) – a persistent, unique and resolvable identifier (PID) – is assigned, resulting in a positive assessment for FAIR Principle F1 (see Table I). If the data object is not managed and preserved in a TDR, the odds that the DOI no longer resolves to the data object are high because the persistence of the DOI depends entirely on the data custodian, i.e the data repository in most cases, to update the URL for the landing page. Hence global services for PID minting and resolution on their own can only enable persistence, but not guarantee it. Similarly, a data object that meets community and domain approved standards today because it uses a preferred file format may be assessed as FAIR regarding Principle

R1.3 (see Table I). Yet if the data object is not preserved in a TDR which ensures file formats migration in adherence with the needs of its designated community, it could become unreadable in the future, e.g. because the file format becomes obsolete or is deprecated in the research community.

Accordingly, a FAIR assessment that considers only the data object at a given point in time but which does not take into account characteristics of the infrastructure in which the object is stored is not sufficient to predict whether a data object will remain accessible and usable over time.

III. CoreTrustSeal requirements and certification

The examples provided earlier illustrate the risks of limiting the assessment of FAIRness to characteristics of data objects thus highlighting the importance of including an assessment of the environment in which the objects reside. In particular this concerns the quality and trustworthiness of the data repositories providing the key infrastructure for the dissemination and preservation of scholarly data.

The Core Trustworthy Data Repositories Requirements [4] define universal and essential (“core”) characteristics of trustworthy data repositories. The CoreTrustSeal Requirements are the result of integrating and improving two predecessor catalogs of criteria – the Data Seal of Approval (DSA) and the World Data System (WDS) Regular Members certifications, already used by many domain repositories in the natural and social sciences and humanities across the globe – into a set of universal requirements that can be applied to research data repositories generally. The CoreTrustSeal Requirements were developed by a DSA and WDS Partnership Working Group established under the umbrella of the Research Data Alliance.[2] In addition to unifying DSA and WDS catalogs of criteria, CoreTrustSeal requirements are aligned with the concepts defined in the Reference Model for an Open Archival Information System (OAIS), an international standard for data repositories also known as ISO 14721:2012 [7]. The requirements are also mapped to ISO 16363:2012, the international standard for Audit and Certification of Trustworthy Digital Repositories [8] as well as the German nestor Seal (based on DIN 31644) [9]. This facilitates the transition from a CoreTrustSeal certification to a DIN or ISO certification as agreed in the Memorandum of Understanding to create a European Framework for Audit and Certification of Repositories.[3]

In the CoreTrustSeal framework, the trustworthiness of a data repository is assessed through a formal certification process which starts with the submission of a self-assessment against the 16 CoreTrustSeal requirements via the Application Management Tool. This self-assessment is then peer-reviewed by two independent experts to verify that the repository meets the requirements and that there is sufficient public evidence supporting the claims made in the self-assessment. In the case of missing evidence or open questions the assessment is returned to the applicant with comments in an iterative process. A successful review results in the award of the CoreTrustSeal by the Board, signaling that the repository can be considered as trustworthy for a period of three years.

As indicated, the CoreTrustSeal certification is conducted at the repository level and the requirements are organized in three main categories addressing the context, structure and activities of a data repository in alignment with ISO 16363. In their self-assessment repositories provide evidence that

1. the *organizational infrastructure* is sound to ensure sustainability: this includes requirements on the mission and scope, licenses, continuity of access, confidentiality and ethics, funding, and expert guidance.
2. *Digital objects management* is performed according to standards to ensure understandability and reusability for the long term of datasets by the designated community: this includes requirements on data integrity and authenticity, appraisal, documented storage procedures, preservation plan, data quality, workflows, data discovery and identification, and data reuse.
3. *Technical infrastructure and security measures* are adequate to protect the data against loss and unauthorized and/or undocumented manipulation.

The CoreTrustSeal requirements were also heavily influenced by the discussions in the data management and data sharing communities, including the emerging consensus and momentum surrounding the FAIR Guiding Principles, which they consequently incorporate although with a different focus and slightly different terminology.

[2] https://rd-alliance.org/groups/repository-audit-and-certification-dsa%E2%80%93wds-partnership-wg.html

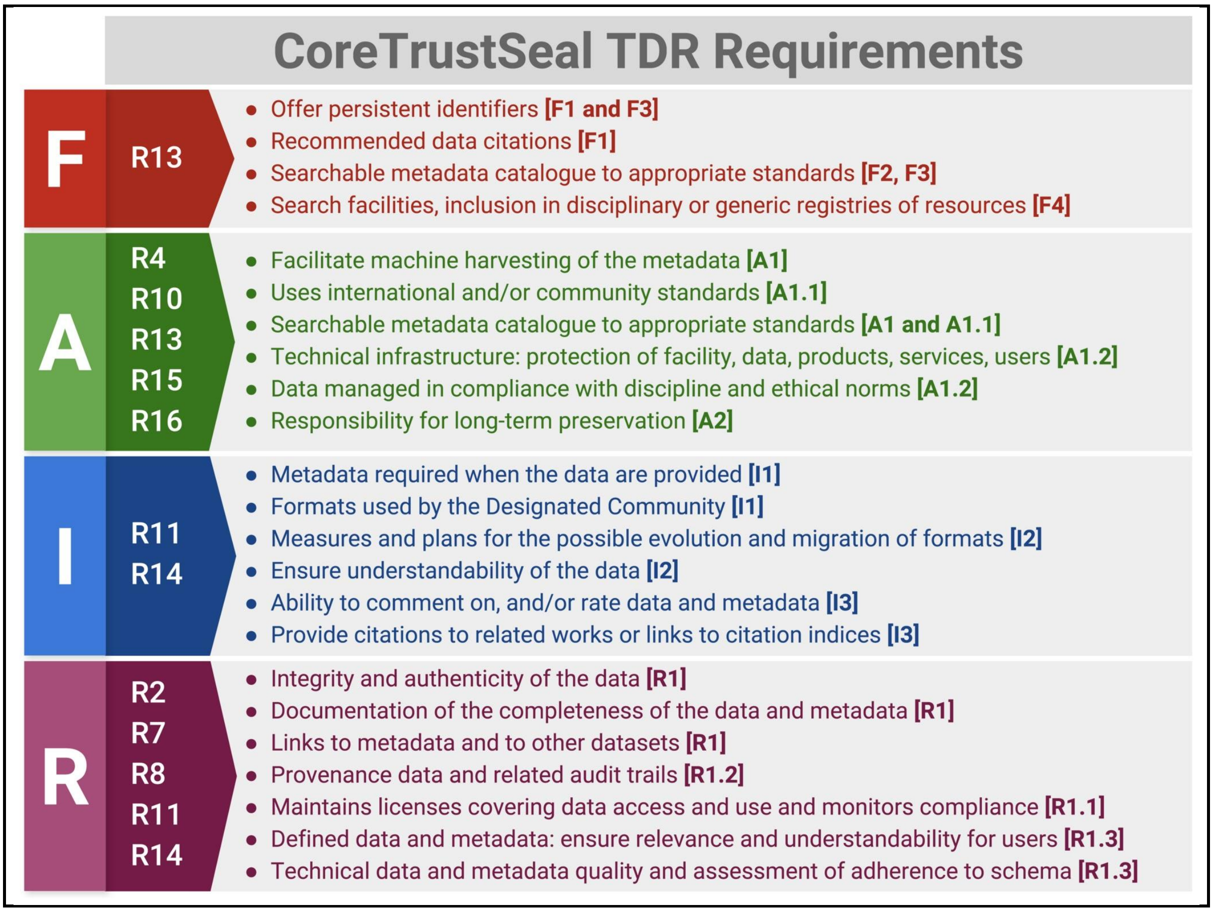
[3] http://www.trusteddigitalrepository.eu/Trusted%20Digital%20Repository.html

IV. CORETRUSTSEAL ALIGNMENT WITH FAIR

As discussed, the FAIR Guiding Principles follow a data and metadata-centric approach with a focus on data discovery, reuse and machine readability, whereas the CoreTrustSeal Requirements are formulated following an infrastructure-centric perspective which incorporates the aspects addressed in the FAIR Guiding Principles but shifts the focus towards data preservation and organizational sustainability. In addition, unlike current approaches to determining FAIR metrics, the CoreTrustSeal criteria are not designed as a checklist of mandatory requirements that repositories and reviewers "tick off" to determine a repository's trustworthiness. While the accompanying guidance and extended guidance contain hints and suggestions as to what kind of technical implementations and evidence applicants are expected to provide, these are not fixed metrics that could be measured in a semi-automated fashion. Rather, reviewers and the CoreTrustSeal Board consider whether the evidence provided sufficiently demonstrates that the repository can be considered trustworthy in relation to its goals (e.g. the level of curation offered) and the context in which it operates (for example, the needs of the designated community for which the data are preserved): "Reviewers are looking for clear, open statements of evidence specific to the applicant. Not necessarily all bullet points in all requirements are mandatory; final judgment depends on the completeness and quality of the answer in the self-assessment of a specific Requirement" [10] (p. 5).

These differences in approach explain to a certain extent the different terminology between the CoreTrustSeal Requirements and the FAIR Principles and more importantly why they do not map in a one to one relationship.

A mapping between the FAIR Guiding Principles and the CoreTrustSeal Requirements is presented in Fig. 1 and will be discussed in more detail in the following sections.



CoreTrustSeal TDR Requirements

F	R13	● Offer persistent identifiers **[F1 and F3]** ● Recommended data citations **[F1]** ● Searchable metadata catalogue to appropriate standards **[F2, F3]** ● Search facilities, inclusion in disciplinary or generic registries of resources **[F4]**
A	R4 R10 R13 R15 R16	● Facilitate machine harvesting of the metadata **[A1]** ● Uses international and/or community standards **[A1.1]** ● Searchable metadata catalogue to appropriate standards **[A1 and A1.1]** ● Technical infrastructure: protection of facility, data, products, services, users **[A1.2]** ● Data managed in compliance with discipline and ethical norms **[A1.2]** ● Responsibility for long-term preservation **[A2]**
I	R11 R14	● Metadata required when the data are provided **[I1]** ● Formats used by the Designated Community **[I1]** ● Measures and plans for the possible evolution and migration of formats **[I2]** ● Ensure understandability of the data **[I2]** ● Ability to comment on, and/or rate data and metadata **[I3]** ● Provide citations to related works or links to citation indices **[I3]**
R	R2 R7 R8 R11 R14	● Integrity and authenticity of the data **[R1]** ● Documentation of the completeness of the data and metadata **[R1]** ● Links to metadata and to other datasets **[R1]** ● Provenance data and related audit trails **[R1.2]** ● Maintains licenses covering data access and use and monitors compliance **[R1.1]** ● Defined data and metadata: ensure relevance and understandability for users **[R1.3]** ● Technical data and metadata quality and assessment of adherence to schema **[R1.3]**

Figure 1 Mapping between FAIR and CoreTrustSeal.

A. *Findability*

The FAIR Findability principle has four sub-principles (FAIR-F1 to -F4; see Table 1) covering persistent identifiers, rich metadata including a reference to the identifier, and registration in a searchable resource. In the CoreTrustSeal (CTS) Requirements, this is mostly covered by CTS-R13 requesting evidence that the repository "enables users to discover the data and refer to them in a persistent way" [4]. The additional guidance provided to data repositories for this requirement makes it clear that data discovery is key to data sharing, and that datasets should be citable including with persistent identifiers to ensure that data can be accessed into the future. The CoreTrustSeal reviewers will particularly look for evidence that the repository offers search facilities, which covers FAIR-F4, and maintains a searchable metadata catalog to appropriate (internationally agreed) standards addressing FAIR-F2 and -F3. The reviewers will confirm if the catalogue is registered in one or more disciplinary or generic registries of resources (FAIR-F4), offers recommended data citations (FAIR-F1) and makes use of persistent identifiers (FAIR-F1 and -F3) .

It is worth noting that FAIR-F2 explicitly requires rich metadata describing the data (as defined in FAIR-R1: "(meta)data are richly described with a plurality of accurate and relevant attributes"). The CoreTrustSeal Requirements focus on the availability of metadata for discovery and the use of domain-specific and international standards without going (much like the FAIR Principles) into the details of necessary attributes. Such attributes to enrich the (meta)data which include controlled vocabularies and ontologies are not necessarily included in discovery or domain-specific standards. What exactly constitutes "rich" metadata requires further discussion in the FAIR context as well [11], in particular as "richness" of metadata will mean different things in different scientific communities. Thus, while the FAIR-Findability principles can be assessed at repository level and would therefore lend themselves to using CoreTrustSeal certification as a baseline for measuring FAIRness, further discussion is required to determine what "richness" means in different contexts and disciplines and how to measure it.

B. *Accessibility*

The FAIR Accessibility principles prescribe the use of standardized, open, free, and universally implementable communication protocols allowing authentication and authorization where necessary to retrieve data and metadata. They also require that metadata remain accessible even when the data are no longer available. At the repository level these principles are either explicitly or implicitly covered by several CoreTrustSeal Requirements.

The guidance for CTS-R13 explicitly covers FAIR-A1 by indicating for instance that the repository should facilitate machine harvesting of the metadata. The CoreTrustSeal guidance also mentions that the repository should maintain a searchable metadata catalogue to appropriate (internationally agreed) standards, which implicitly covers FAIR-A1.1. Similarly, CTS-R15 addresses the technical infrastructure of data repositories to ensure that it is appropriate and in particular that the standards used are relevant to their designated community of users. Although free, open and universally implementable communication protocols are not mentioned explicitly, they are implicitly required to implement a searchable and machine-harvestable metadata catalogue. Thus HTTP/HTTPS and the Open Archives Initiative Protocol for Metadata Harvesting OAI-PMH [12] are among the international standards widely used by data repositories.

For CTS-R4 (and CTS-R9 to some extent) reviewers will check that the repository's data – in particular personal data with a disclosure risk – are created, curated, accessed, and used in accordance with disciplinary and ethical norms. Evidence must include availability of human expertise and technical capacities, for example for anonymization and secure access. Similarly, CTS-R16 stipulates that the technical infrastructure of the data repository must provide for protection of the facility and its data, products, services, and users. Both of these requirements thus mirror FAIR-A1.2

FAIR-A2 requires that metadata should remain accessible, even when the data are no longer available. This is an area where CoreTrustSeal-certified repositories excel by definition because they commit to preserving data and metadata for the long term. CTS-R10 addresses the responsibility for long-term preservation and reviewers will look for evidence that the repository manages this function well. As a consequence, data in general – and metadata in particular – can be expected to continue to be accessible in the case of TDRs (within the boundaries of the data retention policies the TDR set itself).

C. *Interoperability and Reusability*

Reference [3] stipulates that Interoperability and Reusability depend on the FAIR Digital Objects being "represented in common – and ideally open – formats, and [being] richly documented using metadata standards and vocabularies adopted by the related research community" (p. 12). With regard to interoperability, in many disciplines the necessary frameworks already exist. However, due to the increasingly interdisciplinary nature of research, "attention needs to be paid to the extremely challenging task of developing FAIR data frameworks across disciplines and for interdisciplinary research" [3] (p. 11).

The same is true for Reusability, which strongly depends on the use of community- agreed file formats and software as well as on the description of the data objects with standardized metadata and documentation. While community-specific standards and agreements exist in this regard, which formats and metadata should be used for a given class of data objects to facilitate cross-disciplinary reuse depends on the context and thus has to be decided with regard to specific use scenarios.

It follows that the assessment of both Interoperability and Reusability is impossible without taking into account the purpose for which a given community seeks to use a data object. It does not seem feasible to assess this at the level of individual objects but on the level of collections and repository-level, making the CoreTrustSeal certification a potential tool to support this assessment.

Two CoreTrustSeal Requirements particularly relevant to Interoperability and Reusability – CTS-R8, "Appraisal," -R11 "Data quality," and -R14 "Data reuse" – will always be assessed in relation to the repository's scope, preservation goals, and the needs of the designated community.[4] Reviewers will particularly focus on the question of if and how the repository seeking certification ensures that the data objects deposited can be rendered by and are understandable to the intended user community, and that all metadata deemed necessary for this purpose are of sufficient quality.

V. Discussion

Datasets in a TDR meeting the CoreTrustSeal Requirements are managed, curated, and archived in such a way that they are useful and meaningful (FAIR enough) for the repository's designated community and remain so in the future.

We therefore assert that CoreTrustSeal certification of data repositories facilitates the FAIRness assessment of data objects by providing proxy information to evaluate compliance with many FAIR Principles. Certified trustworthy repositories enable a baseline FAIRness level to the datasets they hold and contribute to maintain or even increase the level of FAIRness over time through appropriate data curation and stewardship services.

An automated assessment of the FAIR Findability and Accessibility Principles can rely on machine-readable metadata and testable data services. In contrast, an assessment of Interoperability and Reusability is more difficult as it requires domain expertise to evaluate for example conformance with community standards or the completeness of data and metadata content. For sensitive data in particular, automatic assessment of anonymization is hardly possible which means that a data curator will always be required. In these cases, a possible FAIR assessment procedure cannot rely on comparably simple metrics that could be assessed semi-automatically; instead, the evaluation of an infrastructure's ingest and quality control procedures, for example, has to take into account a complex set of community-specific conditions that does not lend itself to automated assessment easily.

CoreTrustSeal TDR certification addresses the FAIR Interoperability and Reusability requirements at the level of the repository – for example by ensuring that sufficient levels of data curation are applied and that procedures for checking the quality of the data and metadata are in place, in accordance with the needs of the repository's designated community. Therefore, a FAIR assessment for data objects could usefully build upon the certification status of the repository holding the object to make assumptions on its interoperability and reusability.

FAIR Guiding Principles are still being assimilated in the various research communities and their implementation will affect data infrastructures at large. Like many other certification processes, the CoreTrustSeal Requirements are reviewed regularly to incorporate feedback received from certified data repositories and to account for the evolution of practices of the data community. As part of this evolution process the CoreTrustSeal Board has to reflect on how to incorporate references to FAIR Principles as well as to FAIR-enabling standards and technologies (e.g. ontologies and controlled vocabularies) and their implementation in the (extended) CoreTrustSeal guidance.

In March 2019, the CoreTrustSeal Board initiated an open review of the CoreTrustSeal Requirements to define the requirements for the period 2020–2023.[5] The Board anticipates a certain stability of the requirements, yet it makes a commitment to consider the requirements in the light of FAIR Principles implementation.

[4] Further relevant CoreTrustSeal requirements include CTS-R2 "Licenses," mapping to FAIR-R1.1 and CTS-R7 "Authenticity and Integrity" mapping to FAIR-R1.2.

[5] http://www.trusteddigitalrepository.eu/Memorandum%20of%20Understanding.html

The CoreTrustSeal Requirements will also gain from the work and outputs of many initiatives worldwide aiming at making research data FAIR. The "FAIRsFAIR: Fostering FAIR Data Practices in Europe" project in particular will contribute to the adoption FAIR Principles in practice and will cater mainly to the European Open Science Cloud project which brings together European research communities, infrastructure providers and practitioners.[6] It is expected that the outcomes of the FAIRsFAIR project will be directly relevant to the CoreTrustSeal Requirements and will most certainly be considered in the next scheduled review of CoreTrustSeal Requirements.

The CoreTrustSeal, unlike other certification frameworks, emerged directly out of the community of research data repositories. As a "core" certification it provides an entry-level procedure to help data repositories continuously improve and demonstrate their trustworthiness. To be able to continue fulfilling this role, the CoreTrustSeal Board considers it an important task to take into account the current developments around FAIR: to ensure that data repositories are recognized – by researchers, publishers, and funders – as both trustworthy and as enabling FAIR data, now and in the future.

ACKNOWLEDGMENT

The authors would like to acknowledge the contributions to this paper from Dr Peter Doorn, Dr Eliane Fankhauser as well as members of the CoreTrustSeal Board.

REFERENCES

[1] M. D. Wilkinson, et al., "The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data* 3, Article number:160018, March 2016. https://doi.org/10.1038/sdata.2016.18

[2] M. D. Wilkinson, et al., "A design framework and exemplar metrics for FAIRness," *Scientific Data 5*, Article number:180118, June 2018. https://doi.org/10.1038/sdata.2018.118

[3] S. Collins, et al., "Turning FAIR into reality. Final report and action plan from the European Commission Expert Group on FAIR data," European Commission, November 2018. https://doi.org/10.2777/1524

[4] R. Edmunds, et al., "Core Trustworthy Data Repositories Requirements," Zenodo, November 2016. https://doi.org/10.5281/zenodo.168411

[5] CoreTrustSeal Certified Repositories. https://www.coretrustseal.org/why-certification/certified-repositories/

[6] M. Kleemola, "Being trustworthy and FAIR requires people, processes, technologies and collaboration." Tietarkisto Blogi, 29 November 2018. https://tietoarkistoblogi.blogspot.com/2018/11/being-trustworthy-and-fair.html

[7] "Reference Model for an Open Archival Information System: OAIS," CCSDS Secretariat, June 2012. https://public.ccsds.org/Pubs/650x0m2.pdf

[8] "Space data and information transfer systems — Audit and certification of trustworthy digital repositories," International Organization for Standardization, February 2012. https://www.iso.org/standard/56510.html

[9] "nestor Seal for Trustworthy Digital Archives," nestor Working Group "Certification", 2013. https://www.langzeitarchivierung.de/Subsites/nestor/EN/Siegel/siegel_node.html

[10] Core Trustworthy Data Repositories Extended Guidance. https://www.coretrustseal.org/wp-content/uploads/2017/01/20180629-CTS-Extended-Guidance-v1.1.pdf

[11] M. Wilkinson, et al. "FAIRMetrics/Metrics: FAIR Metrics, Evaluation results, and initial release of automated evaluator code (Version v1.0.3)," Zenodo, July 2018. http://doi.org/10.5281/zenodo.1305060

[12] The Open Archives Initiative Protocol for Metadata Harvesting (June 2002) [online] Available at: http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm.

[6] https://www.fairsfair.eu/, https://dans.knaw.nl/en/current/news/european-commission-awards-grant-to-fairsfair-project, https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud