CESSDA Trust Group: Overview of Support Approaches

Possible Support Approaches for Trusted Infrastructures

Authors:

Hervé L'Hours. UK Data Service, UK Data Archive, University of Essex. Chair, CESSDA Trust WG

René van Horik. Data Archiving & Networked Services (DANS)

Mari Kleemola. Finnish Social Science Data Archive (FSD), Tampere University

Jonas Recker. GESIS Leibniz Institute for the Social Sciences

Janez Štebe. Arhiv družboslovnih podatkov [Social Science Data Archives], University of Ljubljana

Birger Jerlehag. Swedish National Data Service (SND)

Introduction

The members of the CESSDA Trust Working Group[1] have a variety of roles in related Trust activities including the CoreTrustSeal[2], Nestor Seal (DIN31644), RDA Repository Certification Interest Group[3] and trustworthy digital repository support on the SSHOC[4] project and FAIRsFAIR[5] project.

This text is a first draft overview of possible trust support approaches based on the experiences of the CESSDA Trust Working Group and its members. It supports internal discussion within CESSDA, particularly within the Trust Group but is also made available to other support providers for information and reference. At the time of writing, this includes Trust Support work planned within the SSHOC project and within the FAIRsFAIR project.

In the context of standards, especially community standards like the CoreTrustSeal 'support' is a peer-driven two-way process. Knowledge sharing comes from all participants and outcomes can feedback into future reviews of the standard requirements.

Support vs Training

Training and support are differentiated in this context. Training and support activities and materials should be mutually aware and closely linked, but the modalities and engagement approaches are not the same.

Training is delivered through the integration of prior work into materials designed to support various approaches to information delivery from self-learning, to webinars or training events. In contrast, support activities are designed around pre-existing standard requirements (e.g. legislation, trustworthy digital repository standards, FAIR principle or open data requirements) and are intended to provide interactive guidance based on the contextual situation of those receiving support. Support in this sense is not a 'one-way' process and may result in new or extended guidance, which could be used by other support providers or recipients. Support activity outcomes may also be fed back to the standards developers. In the context of CoreTrustSeal, this could include new approaches to developing or sharing evidence or feedback, which is used to clarify, correct, extend or otherwise improve the standard.

---

[1] See: https://www.cessda.eu/About/Working-Groups

[2] https://www.coretrustseal.org/wp-content/uploads/2019/11/2019-10-CoreTrustSeal-Extended-Guidance-v2_0.pdf

[3] https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html

[4] https://sshopencloud.eu/

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

Repositories vs Infrastructures

In referencing infrastructures we are not only referring to the existing ERIC infrastructures such as CESSDA[6], E-RIHS[7], CLARIN[8] and DARIAH[9] (all involved in the SSH elements of the EOSC through SSHOC), or to the broader range of ERICs. Infrastructure includes all of the people, processes and technologies that help deliver the full lifecycle data management necessary for science.

The 'traditional' repository model aligns with the 'OAIS' reference model[10] for the scope of an 'archive'. But archives and repositories have always offered a wide range of data support services (to depositors and data users) which go beyond the strict OAIS model. Repositories often offer support (and training) for potential depositors well before they engage with the repository deposit (ingest) process. Many repositories remain involved in the data lifecycle after access to data has been granted. This may include monitoring of data use to ensure that it complies with licence conditions (where possible) or monitoring data citations to demonstrate impact and to integrate knowledge about how data is used back into the repository services. Repositories will often mediate the use of the data they hold by providing analytics and visualisation services. For Big Data collections the repository may need to manage access to subsets of data if it is impractical for the user to simply 'download' data for use. Repositories holding personal data may need to offer specialist access facilities including secure remote access or access via safe rooms.

Not all infrastructure actors consider themselves to be repositories. But even those with no explicit preservation function often have a close alignment with the deposit, storage, access elements of trustworthy digital repository standards. In a full lifecycle context, the wider partnership and dependencies between actors are important. Not only to ensure trust (through transparency, consistency and interoperability) between those actors but also to maximize data integrity and provenance across the data flows.

---

[6] https://www.cessda.eu/
[7] http://www.e-rihs.eu/
[8] https://www.clarin.eu/
[9] https://www.dariah.eu/
[10] http://www.oais.info/

Stakeholders and Audiences

Stakeholder Ecosystem

The "Turning FAIR into Reality" report[11] envisions a wide range of potential stakeholders that are relevant to Trust.
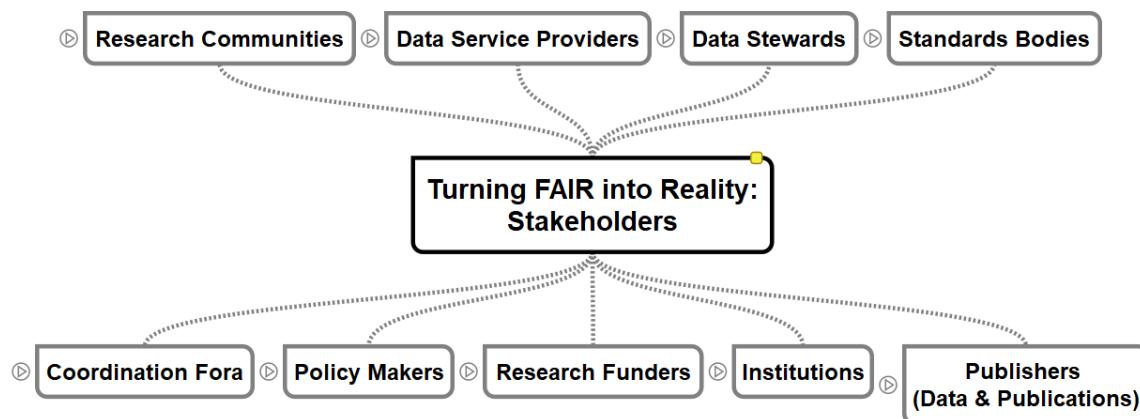

Figure 1: based on Turning FAIR into Reality: 8.3 Stakeholder Groups Assigned Actions

A wide range of these stakeholders may be candidates for Trusted Infrastructure Support. In practical terms, we also need to consider the focus and experience levels of organisations, partnerships and individuals receiving support.

Repository Actor Types

CoreTrustSeal cuts across a range of repository actors including:
- Managers: with a link to funders and related policy-makers (including senior management).
- Practitioners: those undertaking direct curation of the data or offering deposit or user support
- Technologists: developing and maintaining the technical elements of infrastructure

All of these parties need to be involved in a TDR assessment process. Alongside these, we have administrators, resource management and records managers who all have a role in providing and managing the evidence for assessment for certification. As certification is renewed over time these other actors have an important role to play in making sure that evidence of the operational practice is maintained to support quality services. Ideally, this should be sufficient to meet the relevant standard(s).

---

[11] https://doi.org/10.2777/1524

Audience Segmentation by Experience

Experience refers primarily to the knowledge base and skills of those staff receiving support..

Prior access to training events and materials means that most trust support should focus on the level above 'absolute beginner'.

Even experienced repository managers, practitioners and technologists may not be entirely familiar with trustworthy digital repository standards and evidence requirements. Experience levels might indicate support at various levels:
- Introductory: basic overviews of trust standards, their goals and benefits
- Intermediate: mapping trust standards to local practice
- Advanced: managing and improving evidence and efficiency over time.

Audience Segmentation by Maturity

Maturity refers primarily to the level of practice of the organisation receiving support.

There are numerous attempts to address maturity. In a model like CMMI[12] the approach is to evaluate 'capabilities' against different areas of business processes (incomplete, initial, managed, defined) and use them to assign a maturity level which supports future improvement (quantitatively managed, optimized etc).

Audience Segmentation by Goal

Support can be designed to meet a range of different goals, which might be based on the type of actor, the level of individuals' experience or the maturity of the repository processes.

- **Baseline**: familiarise the organisation with good repository operational practice based on standards
- **Integrated**: aligning the repository processes and data types with those standards
- **Assessment**: using internal assessment or peer review to evaluate the level of alignment between standard practice and local practice
- **Certification**: undertaking some formal review and recognition process by an independent third party
- **Renewal**: managing operational documentation to minimise the resource expenditure on maintaining certification. Change managing internal materials over time so they remain operationally useful and valid as evidence for certification. Adapting local materials in a managed way to address changes in standards.

---

[12] https://cmmiinstitute.com/learning/appraisals/levels

Support Providers

Organisations like CoreTrustSeal can offer direct support through the review process and are developing standard materials for wider support of applicants and reviewers. These formally approved materials provide a reference point for other support providers, including commercial consultancy.

ERIC infrastructures like CESSDA have developed internal Trust practice through a group of service provider representatives offering internal support to all service providers.

Through cross-ERIC work like the SSH-focussed SSHOC project support, there is the possibility of extending the CESSDA working group model to other infrastructures, but also to develop an 'umbrella' trust model where the SSH infrastructures share a common trust overview body.

Through projects like FAIRsFAIR repository support which aligns with evolving requirements for 'object FAIRness' will be provided. These repositories will seek to both be Trustworthy and to enable FAIRness.

Support Modalities

Outside of the standard training approach of consolidating information, presenting information and sometimes undertaking exercises, we have a number of possible approaches to offering support. Which is applied depends on the audience segmentation above and the resources available for support.

Support Routes

- Online Webinars
- Face to Face workshops
- Consultancy 'surgeries' for one-to-one support
- Helpdesk model: reactive to questions from multiple sources
- Direct engagement model: pro-active engagement with an applicant or group of potential applicants

Support Approaches

- Information delivery (broadcast)
- Roundtable (participant engagement)
- Q&A

Support Focus

- Trust overview
- Introduction to the CoreTrustSeal
- Changes to the CoreTrustSeal
- Addressing specific requirements
- Developing evidence statements
- Ideal evidence
- Evidence management
- Evidence alignment

Complying with Multiple Standards

One reason that it is important to focus on evidence management and change management is that many actors are subject to additional requirements whether due to local data sensitivity (ISO27001[13] for information security) or because additional specific requirements are in place (CESSDA Annex II[14] requirements for service providers). A managed evidence system:

- Allows the applicant to focus on the information necessary for high-quality service delivery
- Minimised the resource expenditure on meeting multiple standards
- Minimises the resource expenditure on change-managing evidence over time

In the context of CoreTrustSeal, there are evolving requirements that may need to be integrated into the central goal of delivering repository services for data access and long term preservation including:

- FAIR data[15]
- Open data[16]

Progress Monitoring

Support requires standard approaches to monitoring progress in terms of
- Baseline: starting point
- Goals: desired outcomes
- Outcomes

---

[13] https://www.iso.org/isoiec-27001-information-security.html

[14] https://www.cessda.eu/content/download/4015/36855/file/STATUTES%20of%20CESSDA%20ERIC_2017.pdf

[15] https://www.nature.com/articles/sdata201618

[16] https://ec.europa.eu/digital-single-market/en/open-data

Conclusion

This paper provides an overview of the approach to providing support as delivered by the members of the CESSDA Trust Working Group who all have experience with trustworthy repository certification. It has a particular focus on CoreTrustSeal but applies more generally to Trust-related research data management standards. It is provided for reference by other entities providing support and subject to review and revision based on feedback. The group can be contacts via trust@cessda.eu.

| v01.00 | Initial Release | 2020-01-22 |