D2.9: Certification Assessment Report

Author(s)	Heiko Tjalsma
Status	Final
Version	v1.0
Date	22/03/2018

Abstract: In this deliverable the question is answered *how* trustworthiness by means of DSA certification can be achieved within the EUDAT infrastructure. This turned out to be a complicated issue. A partial solution has been found for making the EUDAT services trustworthy for data repositories wanting to user these services for (long-term) storage.

EUDAT receives funding from the European Union's Horizon 2020 programme - DG CONNECT e-Infrastructures. Contract No. 654065

Document identifier: EUDAT2020-DEL-WP2-D2.9	
Deliverable lead	DANS
Related work package	WP2
Author(s)	Heiko Tjalsma (DANS)
Contributor(s)	
Due date	01/09/2017 - 01/03/2018
Actual submission date	22/03/2018
Reviewed by	Rob Baxter (EPCC)
Approved by	PMO
Dissemination level	PUBLIC
Website	www.eudat.eu
Call	H2020-EINFRA-2014-2
Project Number	654065
Start date of Project	01/03/2015
Duration	36 months
License	Creative Commons CC-BY 4.0
Keywords	DSA, CTS, certification, trust

TABLE OF CONTENT

EXECUTIVE SUMMARY

Challenges in the implementation of Data Seal of Approval certification (now merged with the ICSU World Data System and renamed CoreTrustSeal, CTS) across EUDAT's generic service providers were described in EUDAT2020 D2.7 *Certification Implementation Report*. These organisations, being more "generic" than a dedicated digital repository, are increasingly pursuing broader measures of trustworthiness through audited standards like ISO 9001 (service quality management) and ISO 27001 (information security management).

CTS is geared strongly towards data repositories, organisations with "*an explicit mission to provide access to and preserve data"* (CTS Requirement R1). However, it also offers a mechanism whereby certified repositories can make use of non-certified "outsourcing partners", provided that those partners can demonstrate the necessary levels of competence to the repositories' satisfaction. Certification under ISO 27001, for example, would almost certainly demonstrate this, and this could be captured in a suitable agreement between the repository and outsourcing partner. We suggest a form of words to record this within a service or operational level agreement.

Ultimately, while EUDAT's generic service providers have numerous ways to demonstrate their suitability as outsourcing partners under CTS (and a number have pursued or are pursuing such routes), formal recognition of this under CTS requires changes to the CTS certification approach which, while lobbied for by EUDAT, remain elusive.

1. INTRODUCTION

The previous report in this series, D2.7 *Certification Implementation Report* [1], highlighted the practical challenges of implementing the CoreTrustSeal (CTS [2]; formerly Data Seal of Approval, DSA[1]) for trustworthy digital repositories across a network of providers of digital data services of many different kinds. Two main issues arose from the certification implementation phase: EUDAT's core data management service B2SAFE is, by construction, distributed between multiple service providers, requiring an "N-body" level of coordination between different organisations in different jurisdictions to achieve CTS for one service; and, generic service providers are not specialised repositories and have competing interests and pressures which make signing up to CTS difficult or impossible.

In chapter 2 we highlight some of the alternative approaches to "trustworthy certification" taken by EUDAT CDI service providers. In chapter 3 we explore mechanisms through which we might apply alternative measures of trustworthiness like these to support CTS accreditation through outsourcing, and in chapter 4 we offer some ideas about possible modifications to the CTS which could enable its broader application. We conclude the "certification project within EUDAT2020" in chapter 5.

[1] Details of the transition from DSA to CTS are presented in Annex A, reproduced here from D2.7 [1].

2. SERVICE PROVIDERS VERSUS REPOSITORIES

It is easy to argue that the core of the EUDAT Collaborative Data Infrastructure (CDI) is formed from a strong network of generic service providers with deep engagement in data replication and sharing (see, for example, the EUDAT2020 Business Plan [3] for a description of the CDI at the end of the EUDAT2020 project). These centres have joined the CDI as "level 1" service providers under the CDI Agreement [4], tightly integrated and operating the core data replication and management backbone based on the B2SAFE service. These service providers form the core of the CDI because of the scale and IT infrastructure they possess through their roles (typically) as national or major regional computing and science support centres. They are not exclusively data repositories.

EUDAT's generic service providers typically provide data, computing and applications support for a wide audience of scientific groups and individual researchers. Trust and quality of service are both extremely important, but they need to be broader in scope than that of a data repository. Rather than pursue trustworthiness under an "ISO 16363" track (for trustworthy digital repositories), generic service providers in the CDI have sought, and are actively seeking, accreditation and external audit in more general standards of service quality (under ISO 9000) and information security (under ISO 27000).

A non-exhaustive survey of EUDAT CDI service providers (February 2018) identifies at least three generic service providers (CINECA, PDC, EPCC) which have achieved accreditation for service quality management (ISO 9001), at least three (CSC, SURFsara, CINECA) which have achieved accreditation for information security handling (ISO 27001), and at least one (EPCC) actively pursuing ISO 27001.

Achieving ISO accreditation is non-trivial and, arguably, to be done only once (there is a large degree of synergy between ISO 9001 and ISO 27001, making achieving both feasible). Having pursued and achieved one or more of these more general organisational measures of trust and quality, service providers have little incentive to invest more work in a more specialised certification for an area which is just one aspect of their day-to-day business. Is there, then, a way to leverage alternative accreditation or certification approaches like these in building the trustworthiness the CDI seeks?

3. TRUST THROUGH OUTSOURCING

The CTS does offer a mechanism which could provide a way for non-CTS-compliant service providers to engage fully in the CTS model. A non-CTS-compliant EUDAT service provider *is* permitted to offer services to an external CTS-compliant (or compliance-seeking) community repository without undermining the repository's certification; the repository can *outsource,* e.g., data storage, to the EUDAT service provider. The CTS requirements allow this:

> Requirement "R0 -Background Information" clearly indicates under which conditions this is accepted: ***Outsource Partners.*** *Please provide a list of Outsource Partners that your organization works with, describing the nature of the relationship (organizational, contractual, etc.), and whether the Partner has undertaken any trustworthy repository assessment. Such relationships may include but are not limited to: any services provided by an institution you are part of, storage provided by others as part of multicopy redundancy, or membership in organizations that may undertake stewardship of your data collection when a business continuity issue arises. Moreover, please list the certification requirements for which the Partner provides all, or part of, the relevant functionality/service, including any contracts or Service Level Agreements in place. ……. Qualifications/certifications …… are preferred for outsource partners****. However, it is not a necessity for them to be certified*** *(author's emphasis).*
>
> (page 6-7, Core Trustworthy Data Repositories Extended Guidance (v1.0, October 2017 [5])

From this it is clear that tasks can be outsourced, provided that there is unambiguous evidence, as captured in contracts or service level agreements, that the service provider meets the requirements. In the context of EUDAT, a clear task could be the outsourcing of physical storage and subsequent management of the data under B2SAFE. For this it is only necessary to "list the certification requirements for which the Partner provides all, or part of, the relevant functionality/service, including any contracts or Service Level Agreements in place."

These requirements would be included in the SLAs, OLAs and possible Terms of Use that would be agreed between the repository and the service providers in the EUDAT CDI. Wherever possible in these agreements, standard articles are preferable. The development of standard SLAs and OLAs for the EUDAT services is ongoing, affected noticeably by the creation of the CDI's legal persona of EUDAT ltd (see the EUDAT2020 Business Plan [3]), and the fact that CDI service providers are distinct legal entities distributed across different jurisdictions[2]. In support of this, we propose the following template as part of a SLA for complying with CTS Requirements. This template only deals with storing and giving access to data. These however are the main functions in this respect for B2SAFE and possibly B2SHARE. In the template below these requirements have been specified.

The advantage of this solution is that, if applied in this way, it does not create a distinction between EUDAT sites.

[2] One main conclusion to the work in EUDAT on service catalogues and service level management is that "standard" SLAs or OLAs defined by the project can only ever be templates; customization between customer, service provider and/or EUDAT ltd are inevitable.

TEMPLATE For Articles in SLAs to be used by repositories storing or making accessible data in EUDAT services

Storage

The site should guarantee reliable storage during the period for which the SLA is valid

The site should guarantee back-up storage for all the data covered by the SLA

The site is under no circumstances allowed to delete or modify any of the data covered by the SLA without consultation with and approval of the repository.

The site should have a recovery plan in case of calamities in place

If the data contain personal data, the site complies with the European or national legislation on personal data

The site is only allowed to store data in external storing places, like Cloud storing solutions, if the contract between the site and the external (Cloud) provider respects the relevant European and national legislation on personal data

Access (if the EUDAT site would also provide access to the data to users)

The site should guarantee that both metadata and data are only made accessible under the conditions as specified by the repository

4. POSSIBLE AMENDMENTS OF THE CORETRUSTSEAL

One conclusion of our exploration of trustworthiness in EUDAT services is that out-sourcing as discussed in the previous chapter is a possibility for EUDAT sites to be able to offer trust to data repositories or researchers who want to deposit their data in these sites. This only works if the EUDAT site can comply with the conditions set by the CTS Requirements for outsourcing (requirement R0), as noted. If these conditions are met the external data repository can get CTS-certified.

Those EUDAT sites which have become CTS-compliant, as described in [1], could offer trustworthiness acceptable for the CTS certification. To offer this as part of an EUDAT service is not possible (as also explained in [1]).

Another, and perhaps more realistic, possibility in this respect would be that EUDAT services could become CTS certified if the CTS Board would allow this by either amending or extending its certification requirements, more specifically by enabling *partial certification*. Creating a kind of *CTS-outsource-partner* or *CTS-storage* seal would be a possibility for sites where full CTS is not reachable, and perhaps not even desirable. This level of certification would be naturally restricted to technical infrastructure, perhaps. However, it should also be noted that, for legal reasons, it seems inevitable that EUDAT service providers would also need to be aware (at least) of the legal framework concerning "sensitive" data, especially the risks of storing personal data. (This is described in EUDAT report *Guidelines on Open Access and Restricted Data* [6]). In response, one can easily argue that those service providers pursuing accreditation under ISO 27001, for instance, would have little difficulty in complying with such requirements.

Discussions with (formerly) the DSA and (latterly) the CTS Board on the subject of some kind of partial certification for external outsource partners have been held but have so far been inconclusive. This was possibly to do with the transition period the DSA Board had gone through from DSA to CTS consuming a huge amount of their time. The consequence of this inconclusiveness is that it makes it impossible to include a "CTS-lite" option for EUDAT services and service providers at the moment. Theoretically, though, this still remains an option.

5. CONCLUSIONS

The framing question for this activity within the EUDAT2020 project was how trustworthiness, by means of DSA certification, could be achieved within the EUDAT infrastructure. The question posed in this way at the beginning of the EUDAT2020 project transformed itself more and more into the question of *if* trustworthiness by means of DSA certification could be achieved within the EUDAT infrastructure.

Now, at the end of the task, we can say that the answer to the latter question is yes, but only in a limited way.

Recalling the two-track approach to implementation described in [1] – certification of EUDAT *sites*, and certification of EUDAT *services* – we can firstly summarise the work carried out on the first track. On the one hand, there are some clear results – some sites have either achieved or will achieve CTS certification soon; on the other, we can question the original premise that DSA/CTS accreditation for everyone was the appropriate way to build the sought-for trust in EUDAT services. The present requirements of the CoreTrustSeal make the latter impossible when contrasted with the nature of many of the core EUDAT service provider sites. As explained in [1], individual site certification cannot lead (automatically) to certification of the EUDAT services. It can of course be important for these certified sites themselves as they can function as a trustworthy data repository and can offer services in this field, but it is not a good fit for the general structure of the EUDAT CDI services.

The second track has limited value. Again, EUDAT services as such cannot obtain certification at the moment. It will only be possible for those sites offering EUDAT services, specially B2SAFE or B2SHARE, to act as *outsourced providers* of IT services – storage, for example – for external CTS-certified repositories. At the moment there is no formal mechanism under CTS to recognise this. In order to be recognised as such by the CTS Board this needs to be laid down in SLAs or contracts, as suggested in chapter 3. As noted, service providers pursuing broader certification in information security management would be well placed to meet these requirements. Indeed, the context of being an EUDAT CDI service provider could offer a useful framing use-case for sites pursuing, for example, ISO 27001 certification.

The idea of a *CTS-outsource-provider* or *CTS-storage* certification model still has to be developed; at the moment these do not exist. If the CTS Board should decide to develop such a model it could offer interesting chances for EUDAT to make its services trustworthy in a more integral way than is possible at the moment.

6. REFERENCES

[1] H. Tjalsma, *Certification Implementation Report*, EUDAT2020 D2.7, March 2018.
[2] CoreTrustSeal for data repositories, https://www.coretrustseal.org/
[3] R. Baxter, *EUDAT2020 Business Plan*, EUDAT2020 D2.10, February 2018.
[4] D. Lecarpentier et al, *EUDAT Collaborative Data Infrastructure Agreement and Annexes*, version 1.2, 10 June 2016.
[5] CoreTrustSeal, *Core Trustworthy Data Repositories Extended Guidance*, v1.0, October 2017, https://www.coretrustseal.org/wp-content/uploads/2017/01/20171026-CTS-Extended-Guidance-v1.0.pdf
[6] R. Baxter et al, *Guidelines on Open Access and Restricted Access Data (final)*, EUDAT2020 D2.8, August 2017.

ANNEX A. BACKGROUND OF THE CHANGE FROM DSA TO CORETRUSTSEAL CERTIFICATION

As mentioned in the text the DSA certification has changed during the course of the EUDAT2020 project into the CTS CoreTrustSeal certification standard. This was the result of the DSA entering into a partnership with ICSU, the World Data System. The aim of the partnership was to create a set of harmonized common procedures for certification of repositories at the basic level. The evolution of standards and processes to apply trustworthy digital repository (TDR) status has met an acknowledged need for standardization and formalization in the area of data repository practice. The ISO 16363 - *Audit and certification of trustworthy digital repositories*, DIN 31644 - Criteria for trustworthy digital archives, NESTORseal, and the Data Seal of Approval (now CoreTrustSeal) have provided a number of OAIS-related approaches to certifyi trustworthiness.

By 2014 it was considered that the provision of a single 'core' TDR certification approach was better for the data management community as a whole and between 2014 and 2016 the DSA and WDS reviewed their goals, processes and procedures in the context of a Research Data Alliance (RDA) working group. Group members came primarily from the DSA Board and WDS Membership Committee but participation was open and progress reports were provided to the wide-ranging membership of the Repository Certification Interest Group of the RDA. The group released a set of common requirements and common procedures for public comment and undertook internal testing (by both WDS and DSA members) of the revised requirements. After a long open comment period, the new criteria were adopted by both DSA and WDS in 2016.

In the CTS, there are still 16 requirements as in the DSA, though their structure and content had changed to provide clearer language and a greater alignment with the other mentioned TDR Framework standards. There were changes to the structure of the requirements: six on organisational infrastructure, eight on digital object management and two on technology. Extensive care was taken to ensure consistency between the old and new approaches such that prior WDS or DSA recipients would be able to renew to the new requirements. However, the requirements are not identical, with more detail provided on documented public evidence requirements and a greater focus on technical infrastructure and appropriate information security.

During the self-assessment, each applicant indicates a compliance level for each of the requirements:

0 – Not applicable

1 – The repository has not considered this yet

2 – The repository has a theoretical concept

3 – The repository is in the implementation phase

4 – The guideline has been fully implemented in the repository

The compliance levels provided by the applicants will be judged against the given evidence by the reviewers.

Since July 2017, repositories applying for DSA or WDS certification are certified by the CoreTrustSeal. More information can be found at https://www.coretrustseal.org/about/. In August 2017, the two bodies constituted an interim DSA-WDS Board to manage the process and requirements. The CTS governance will be formalised in a legal structure early 2018.

A new common tool and a new CoreTrustSeal brand to support a new independent board will replace the interim board when statues and business model considerations are agreed.

ANNEX B. GLOSSARY

Term	Explanation
DSA	Data Seal of Approval
CTS	CoreTrustSeal
OAIS	Open Archival Information System
SLA	Service Level Agreement
OLA	Operational Level Agreement
ToU	Terms of Use
AIP	Archival Information Package (OAIS)
SIP	Submission Information Package (OAIS)
DIP	Dissemination Information Package (OAIS)
WDS	World Data System