

Research and Innovation Action

Social Sciences & Humanities Open Cloud

Project Number: 823782

Start Date of Project: 01/01/2019

Duration: 40 months

Report on **Milestone 47** Follow up of certification status

Dissemination Level	PU
Due Date of Milestone	31/12/21, M36
Actual Achievement Date	28/10/21, M34
Lead Beneficiary/LTP	1.6 CESSDA ERIC-FSD
Work Package	WP8 - Governance/ Sustainability/ Quality Assurance
Task	8.2 Trust & Quality Assurance, Impact
Version	V1.0
Number of Pages	p.1 – p.11

Abstract:

This document reports on the achievement of SSHOC Milestone 47, Follow up of certification status.

The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/ her sole risk and liability. This deliverable is licensed under a Creative Commons Attribution 4.0 International License.



Author List

Organisation	Name	Contact Information
CESSDA/FSD	Henri Ala-Lahti	henri.ala-lahti@tuni.fi
CESSDA/FSD	Mari Kleemola	mari.kleemola@tuni.fi
CESSDA/UKDS	Benjamin Mathers	bm21346@essex.ac.uk
CESSDA/UKDS	Hervé L'Hours	herve@essex.ac.uk
CNR/E-RIHS	Maurizio Sanesi	maurizio.sanesi.eng@gmail.com
CNR/E-RIHS	Emiliano Degl'Innocenti	emiliano.deglinnocenti@cnr.it

1. Introduction

This document reports on the status of the certification support provided to select SSHOC repositories based on the SSHOC D8.2 Certification plan for SSHOC repositories¹. The plan set out three main modes of support:

- Raising awareness of the relevance of certification of repositories and the role of the CoreTrustSeal in this via existing communication channels.
- Workshops and webinars on repository certification.
- Review of self-assessments based on the requirements of the CoreTrustSeal.

The status of these as well as the certification status of 49 repositories selected as candidates for certification support, and the support progress of 14 repositories are outlined in this document.

2. Description of the Milestone

2.1. Role of the Milestone

The certification support process benefits the participating repositories directly by improving their practices based on the CoreTrustSeal requirements and facilitating their journey towards a formal

¹ Kleemola, Mari, Alaterä, Tuomas J., Koski, Niko, Ala-Lahti, Henri, Jerlehag, Birger, L'Hours, Hervé, De Jong, Franciska, Van Uytvanck, Dieter, Parkola, Tomasz, Degl'Innocenti, Emiliano, Giacomini, Roberta, Sanesi, Maurizio, & Van Horik, René. (2020). *SSHOC D8.2 Certification plan for SSHOC repositories (v1.0)*. Zenodo. <https://doi.org/10.5281/zenodo.4558303>

CoreTrustSeal application. In addition, the support work allows exploring the benefits of certification and providing CoreTrustSeal feedback from SSH viewpoint based on the lessons learned. The support process also provides insight into the applicability of certification to different types of repositories as well as the differences and similarities between repositories within SSH.

On a more general level, certification improves the practices of repositories, as trusted digital repositories enable access to FAIR data over time. Certification also contributes to providing access to secure and trusted repositories by identifying these to researchers, funders and other stakeholders.

2.2. Means of verification

The means of verification for MS47, as stated in the GA, is “update of certification status shared within project”. This report on the certification support activities and the *SSHOC T8.2 Certification status spreadsheet*² (especially sheet *Certification status (MS47)*) shared with the project members serve as the update on the certification status. The table on the certification status (MS47) is also presented in Appendix 1 of this report.

In addition, T8.2 has an internal spreadsheet, *SSHOC T8.2 certification support progress*, restricted to the team members that outlines the discussions with the supported repositories and their progress in drafting their CoreTrustSeal applications. The reviewers update the spreadsheet and provide regular status updates on the progress in the monthly task meetings. The spreadsheet remains restricted and confidential, as agreed with the supported repositories.

Awareness about certification has been raised by communicating, with the help of Work Package 2, about certification and the work of T8.2 through the SSHOC Certification mailing list, SSHOC website³ as well as within the SSHOC network and informal contacts with other relevant networks, including the FAIRsFAIR and EOSC Nordic projects.

A webinar entitled *How to improve the quality of your repository? SSHOC and certification of repositories* was organised on 23 April 2020 and an online workshop entitled *Building Trustworthy Repositories: Introduction to CoreTrustSeal Certification* was organised with DARIAH on 14 October 2020.

² MS47 certification status spreadsheet is available within the SSHOC Shared Drive Repository available to all partners and maintained by the project coordinator during the duration of the project implementation [26.10.2021]

³ “SSHOC Trusted Repositories” <https://sshopencloud.eu/sshoc-trusted-repositories>,
“SSHOC webinar on certification of data repositories” <https://www.sshopencloud.eu/news/sshoc-webinar-certification-data-repositories>,

“Update from SSHOC Certification Support” <https://sshopencloud.eu/news/update-sshoc-certification-support>.

The events and communications were aimed at the whole SSHOC community even though the 49 repositories selected as candidates for certification support in the Certification plan were the main target group. The certification support program that includes the opportunity to receive information on the CoreTrustSeal certification process and detailed feedback on self-assessments was announced through an open call⁴ between June and August 2020.

Applications were received from 14 interested repositories, none of which had previously applied for CoreTrustSeal (one repository had previously received certification from CoreTrustSeal's predecessor, the Data Seal of Approval⁵). The applicants represented 13 different European countries. Six of them were among the 49 candidates listed in D8.2 and eight repositories came from outside the list which indicates wide interest in certification. All applicants were deemed in scope and accepted into the support program. The repositories are:

- CLARIN-LV [Latvia]
- Corpus OVI dell'italiano antico [Italy]
- Croatian Social Science Data Archive (CROSSDA)
- DAIS - Digital Archive of the Serbian Academy of Sciences and Arts
- DARIAH-DE Repository [Germany]
- Digital Library of University of Maribor [Slovenia]
- Digital Repository of Scientific Institutes [Poland]
- Historic Graves [Ireland]
- Lithuanian Data Archive for Social Sciences and Humanities (LiDA)
- mdw Repository [Austria]
- NAKALA [France]
- SB/CLARIN Repository (Språkbanken Text CLARIN Repository) [Sweden]
- Sciences Po, Center for Socio-Political Data (CDSP) [France]
- Slovak Archive of Social Data (SASD) [Slovakia]

The repositories were contacted in September 2020 and asked to complete a self-assessment template to gain insight into their level of preparedness. An invitation-only webinar was organised for supported repositories on 28 January 2021 where the participants were able to raise questions about certification and share their experiences. All were assigned two supporters from T8.2.

After the webinar, T8.2 has continued to support the repositories in individual meetings between the assigned reviewers and repository representatives. Typically, the repositories complete their self-

⁴ Call for Applications: <https://www.sshopencloud.eu/news/call-applications-sshoc-repository-certification-support-0>

⁵ Data Seal of Approval: <https://www.coretrustseal.org/about/history/data-seal-of-approval-synopsis-2008-2018/>

assessment texts against agreed CoreTrustSeal requirements before the meeting. In the meetings, T8.2 members coach them and help them move forward. Between the meetings, the T8.2 members have also provided comments on draft CoreTrustSeal texts. The repositories had varying goals ranging from aligning practices with the CoreTrustSeal requirements to full CoreTrustSeal certification. The certification support process benefits all the participants even if they are not going to apply for CoreTrustSeal at this time, as it has allowed them to improve their technical and data management practices, processes and documentation.

Regular meetings have been organised with nearly all of the participating repositories roughly once a month and the support work is ongoing. The support progress of each repository is followed in the monthly meetings of T8.2.

At least half of the supported repositories are poised to apply for CoreTrustSeal by the end of the project, although it is unlikely that any of them will receive the certification within the project timeline due to the response times of the CoreTrustSeal review process. The repositories that are close to applying will continue to be supported until the end of the project. The rest will be supported until the end of January 2022 or as agreed individually.

3. Conclusions and next steps

The support progress is on track and support for all participating repositories will be ongoing until the end of January 2022. The support of select repositories continues until the end of the project, April 2022. The certification status of the 49 candidates for certification listed in D8.2 will be reviewed in the final deliverable 8.3 of the task.

The support process is complemented by other activities of T8.2. These activities include: a survey charting the landscape of SSHOC repositories and the applicability of CoreTrustSeal certification for them; desk research examining the infrastructure and services of a wide range of repositories; collaborative work with EOSC Nordic and FAIRsFAIR; and a questionnaire to examine SSHOC stakeholder views on certification in relation to infrastructure.

The outputs of the survey and desk research will be published reports on the landscape of various SSHOC repositories. The findings from certification support work and possible certification solutions will be reported in the final deliverable 8.3.

4. Appendixes

Appendix 1: Certification status, MS47

The table lists 49 repositories selected as candidates for certification support in SSHOC D8.2 Certification plan for SSHOC repositories. The repositories that have achieved their certification goal by MS47 are highlighted in green. The repositories that are being supported in their certification process are highlighted in blue. As CoreTrustSeal is valid for three years from the certification date, the certificate in accordance with 2017-2019 requirements is still valid for many repositories that are seeking to renew their certificate. The status will be reassessed in D8.3.

TABLE 1 - REPOSITORY CANDIDATES

Repository name	Country	Community	Certificate during D8.2	Goal (D8.2)	Certificate during MS47	Status (MS47)	Supported by SSHOC
ACDH - A Resource Centre for the HumanitiEs (ACDH-ARCHE)	Austria	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2020-2022	Renewed	no
ASV Leipzig	Germany	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Austrian Social Science Data Archive (AUSSDA)	Austria	CESSDA	Not certified	B	CoreTrustSeal v2017-2019	Certified	no
Bayerisches Archiv für Sprachsignale (BAS)	Germany	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no

Berlin-Brandenburg Academy of Sciences and Humanities (BBAW)	Germany	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Center of Estonian Language Resources (CELR-EKK)	Estonia	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
CLARIN-PL Language Technology Centre	Poland	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
CLARIN.SI Language Technology Centre (CLARINSI)	Slovenia	CLARIN	DSA	B	CoreTrustSeal v2017-2019	Certified	no
CLARINO Bergen Center	Norway	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
CMU-TalkBank (CMU)	USA	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2020-2022	Renewed	no
Corpus testuale OVI	Italy	E-RIHS	Not certified	B	Not certified	Not certified	yes
Czech Social Science Data Archive (CSDA)	Czech Republic	CESSDA	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Danish National Archives (DNA)	Denmark	CESSDA	Not certified	B	Not certified	Not certified	no
Data Archiving and Networked Services (DANS)	Netherlands	CESSDA	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2020-2022	Renewed	no

Data Center Serbia for Social Sciences (DCS)	Serbia	CESSDA	Not certified	B	Not certified	Not certified	no
Digital Repository of Ireland (DRI)	Ireland	DARIAH	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Digital Repository of Scientific Institutes (DRSI)	Poland	DARIAH	Not certified	B	Not certified	Not certified	yes
Eberhard Karls Universität Tübingen (EKUT)	Germany	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Finnish Social Science Data Archive (FSD)	Finland	CESSDA	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2020-2022	Renewed	no
Geisteswissenschaftliches Asset Management System (GAMS)	Germany	DARIAH	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
GESIS - Leibniz Institute for the Social Sciences	Germany	CESSDA	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Greek research infrastructure for the social sciences (So.Da.Net)	Greece	CESSDA	Not certified	B	Not certified	Not certified	no
Hamburger Zentrum für Sprachkorpora (HZSK)	Germany	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no

Institut für Deutsche Sprache (IDS)	Germany	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Institut für Maschinelle Sprachverarbeitung (IMS)	Germany	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Instituut voor de Nederlandse Taal (IVDNT)	Dutch Language Union	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
LINDAT/CLARIN (LINDAT)	Czech Republic	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Meertens Instituut/HUC (MI)	The Netherlands	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
MOBILE-laboratory Visualization DATA (MOVIDA)	Italy	E-RIHS	Not certified	B	Not certified	Not certified	no
MPI for Psycholinguistics (MPI-PL)	The Netherlands	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
NAKALA	France	DARIAH	Not certified	B	Not certified	Not certified	yes
Norwegian Centre for Research Data (NSD)	Norway	CESSDA	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no

Piattaforma Lessicografica Unica del Tesoro delle Origini (PLUTO)	Italy	E-RIHS	Not certified	B	Not certified	Not certified	no
Portuguese Social Information Archive (APIS)	Portugal	CESSDA	Not certified	B	CoreTrustSeal v2020-2022	Certified	no
PORTULAN CLARIN	Portugal	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
PROGEDO Research Infrastructure	France	CESSDA	Not certified	B	Not certified	Not certified	yes
Slovak Archive of Social Data (SASD)	Slovakia	CESSDA	Not certified	B	Not certified	Not certified	yes
Social Science Data Archives (ADP)	Slovenia	CESSDA	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Social Sciences and Humanities Data Archive (SOHDA)	Belgium	CESSDA	Not certified	B	Not certified	Not certified	no
Språkbanken, The Swedish language bank	Sweden	CLARIN	DSA	B	Not certified	Not certified	yes
Swedish National Data Service (SND)	Sweden	CESSDA	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
Swiss Centre of Expertise in the Social Sciences (FORS)	Switzerland	CESSDA	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no

Tárki Data Archive	Hungary	CESSDA	Not certified	B	Not certified	Not certified	no
Tesoro della Lingua Italiana delle Origini (TLIO)	Italy	E-RIHS	Not certified	B	Not certified	Not certified	no
The CLARIN Centre at University of Copenhagen (CLARIN-DK-UCPH)	Denmark	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
The ILC4CLARIN Centre at the Institute for Computational Linguistics (ILC4CLARIN)	Italy	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
The Language Bank of Finland (FIN-CLARIN)	Finland	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no
UK Data Service (UKDS)	United Kingdom	CESSDA	DSA	B	CoreTrustSeal v2017-2019	Certified	no
Universität des Saarlandes (UdS)	Germany	CLARIN	CoreTrustSeal v2017-2019	A	CoreTrustSeal v2017-2019	Not yet renewed	no