Building trustworthy repositories: Introduction to CoreTrustSeal certification

Tomasz Parkoła[*†1], Rene Van Horik[*‡2], Birger Jerlehag[*3], and Daan Broeder[*4]

[1]Poznan Supercomputing and Networking Center – Poland
[2]Data Archiving and Networked Services – Netherlands
[3]Swedish National Data Service – Sweden
[4]KNAW Humanities Cluster/CLARIN ERIC – Netherlands

Abstract

Scholarly primitives, such as representing, discovering, sampling and referring are typically well supported by data repositories. However, these do not provide evidence that the data or features provided by the repository are managed and operated in a trustworthy and reliable manner. Assessment against an agreed standard is a way to evaluate the reliability and sustainability of data repositories. Certification helps to demonstrate this to the repository users and funders, thus ensuring them that available data and services are of a certain quality that is defined in corresponding requirements.

The CoreTrustSeal offers baseline certification for repositories and is playing an increasingly prominent role in the ecosystem of generating and serving FAIR data. For CoreTrustSeal certification, a repository first conducts an internal self-assessment against the 16 CoreTrustSeal requirements, which is then reviewed by community peers. Even outside of the formal certification framework the CoreTrustSeal criteria provide a demonstrable approach to internal and external review, supporting a benchmark for comparison and a means to determine the strengths and weaknesses of data repositories.

The main target audience of this workshop are the creators or managers of data repositories who are (planning to) seek the CoreTrustSeal certification or who want to understand how their data is managed in the context of CoreTrustSeal requirements. However, the workshop will be useful to anyone wishing to build digital research infrastructure in a sustainable way or understand the characteristics of such an infrastructure.

Participants of the workshop will have the opportunity to familiarize themselves with the CoreTrustSeal certification requirements and procedures, gaining a better understanding of the necessary steps for certification, and how compliance with the criteria may be met. CoreTrustSeal experts will share their experiences with certification and answer certification-related questions. Participants are asked to take notice of the "CoreTrustSeal Extended Guidance" and drafting a self-assessment response to at least 2-3 requirements.

We will also describe why and how the current SSH research infrastructures (CESSDA, CLARIN) use and recommend CoreTrustSeal as means for supporting sustainable FAIR science.

[*]Speaker
[†]Corresponding author: tparkola@man.poznan.pl
[‡]Corresponding author: rene.van.horik@dans.knaw.nl

Overall agenda of the workshop:
1. Introduction to CoreTrustSeal
2. CoreTrustSeal approach in CESSDA and CLARIN
3. Case studies / hands-on session

The workshop is part of the certification support activities of the Social Sciences & Humanities Open Cloud (SSHOC) project (https://sshopencloud.eu/) that will create the SSH part of the European Open Science Cloud. SSHOC has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-04-2018, grant agreement #823782.

Names of organizers:
René van Horik, Data Archiving and Networked Services
Tomasz Parkola, Poznań Supercomputing and Networking Center
List of speakers/instructors:
Birger Jerlehag, Swedish National Data Service
Daan Broeder, KNAW Humanities Cluster/CLARIN ERIC
Rene van Horik, Data Archiving and Networked Services
Tomasz Parkola, Poznań Supercomputing and Networking Center

Keywords: trustworthy repositories, data repositories, CoreTrustSeal requirements