



CoreTrustSeal - Earth, Space, and Environmental Science Repository Community Cohort

Rebecca Koskela
24 February 2021



Enabling FAIR Data Project - Objectives

- 🌐 FAIR-aligned data repositories
 - Add value to research data
 - Provide metadata and landing pages for discoverability
 - Support researchers with documentation guidance, citation support, and curation
- 🌐 FAIR-aligned Earth, space, and environmental science publishers
 - Align their policies to establish a similar experience for researchers.
 - Citations for data, software, technology resolve to repository landing pages
 - Provide availability statements
 - Data are not placed in the supplemental information



CoreTrustSeal Earth, Space, and Environmental Science (ESES) Repository Cohort

American Geophysical Union Enabling FAIR Data Project

- Cohort of data repositories assembled to apply for CoreTrustSeal certification
- Certification demonstrates to users and funders that they have been evaluated by an independent authority and endorsed for their trustworthiness



CoreTrustSeal ESES Cohort



Purpose

- Establish a group of interested Earth, space, and environmental science repositories to systematically work together through the tenets of the CTS requirements with the goal of each submitting a certification package with their individual responses to each criteria.



Expected Outcome

- Complete and submit their CTS certification applications within 3 years (February 2022) from the start date of the cohort



CoreTrustSeal ESES Cohort (cont.)


Venues

○ Webinars:

- 30 October 2018 <https://youtu.be/gS9HTzxbx0s>
- 26 November 2018 https://youtu.be/vV_3ZcDsagA
- 12 June 2020 https://youtu.be/Y8t8fJ0_piw
- 10 July 2020 - Panel discussion at EarthCube Council of Data Facilities Summer meeting
https://youtu.be/-qGr_ljkAhg

○ Meetings

- American Geophysical Union (AGU)
- European Geophysical Union (EGU)
- Earth Science Information Partners (ESIP)
- Research Data Alliance Plenaries

 **Shelley Stall**
@ShelleyStall
@CoreTrustSeal cohort in the Earth space and environmental sciences meeting tomorrow (Wed) at 2pm EST in Pod 4 @ #AGU18 #FAIRdata - Join Us!!



Successes

- 🌐 Lots of interest in joining the cohort (25 repositories - global)
- 🌐 Good attendance at webinars
- 🌐 Support for waiver of application fee
- 🌐 Submissions
 - Research Data Archive: certified
 - Arctic Data Center: certified
 - Biological and Chemical Oceanography Data Management Office (BCO-DMO): certified
 - Incorporated Research Institutions for Seismology Data Management Center (IRIS DMC): certified
 - IsoArch (Open-Access and Collaborative Isotope Database for Bioarcheological Samples): submitted



Challenges

- 🌐 Need for ongoing communication to cohort to keep interest
- 🌐 Resources at repositories to complete certification application
 - Needed for funder's renewal
 - Ability to fund a contractor to support the application process
- 🌐 Need to find a way to pair repositories starting process with those who have completed the process



Questions?



CoreTrustSeal 101

Certify Your Data Repository – Get the Support of a Cohort!
2021-02-24

Ocean Networks Canada, ROR: 05qknh003
University of Victoria, ROR: 04s5mat29

Reyna Jenkyns, ORCID: 0000-0001-6975-6816

A UNIVERSITY OF VICTORIA INITIATIVE

TRUST Principles

<https://www.nature.com/articles/s41597-020-0486-7.pdf>



Transparency	To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence.
Responsibility	To be responsible for ensuring the authenticity and integrity of data holdings and for the reliability and persistence of its service.
User Focus	To ensure that the data management norms and expectations of target user communities are met.
Sustainability	To sustain services and preserve data holdings for the long-term.
Technology	To provide infrastructure and capabilities to support secure, persistent, and reliable services.

“.. to make data **FAIR** whilst preserving them over time requires **trustworthy digital repositories** (TDRs) with **sustainable governance** and **organizational frameworks, reliable infrastructure, and comprehensive policies supporting community-agreed practices**. TDRs, with their clear remit to actively **preserve data** in response to changes in both technology and stakeholder requirements, play an important role in **maintaining the value of data**. They are held in a position of trust by their users as they **accept the responsibilities of data stewardship**.”

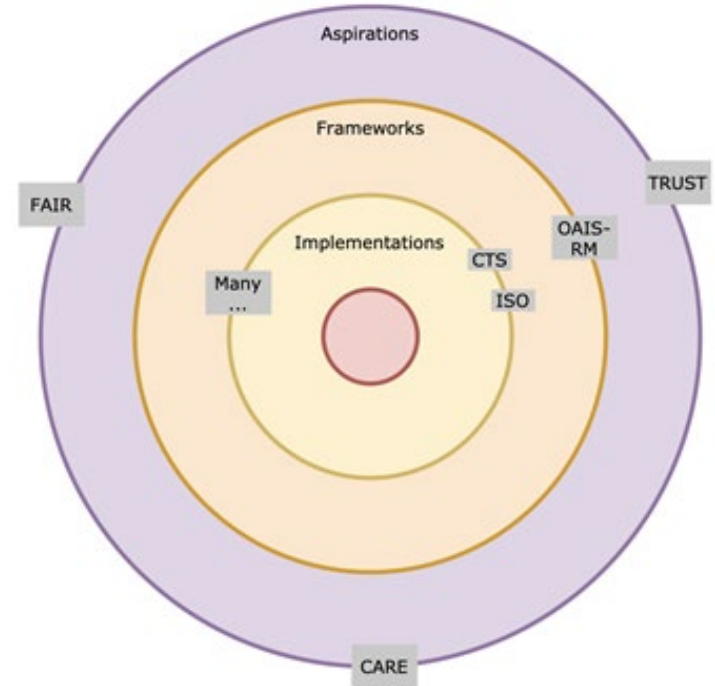
Lin, D., Crabtree, J., Dillo, I., Jenkyns, R., *et al.* The TRUST Principles for digital repositories. *Sci Data* 7, 144 (2020).
<https://doi.org/10.1038/s41597-020-0486-7>

CoreTrustSeal - History and Context

Data Seal of Approval Certification
of Trusted Data
Repositories



Research Data Alliance
Repository Audit and
Certification DSA-WDS
Partnership WG



Credit: Wim Hugo

CoreTrustSeal acts as a means towards the aspiration of the TRUST Principles. The implementation is done by the repository in fulfilling the CTS requirements.

CoreTrustSeal - Basics

- Certifications are valid for 3 years
- 16 Requirements self-assessed online, and two peer-reviews
- Applications become publicly available (great resource for new applicants!)
- 1000 euro fee
- Reliant on in-kind contributions from board members and application reviewers
- Compliance levels (must be at 3 or 4 for all 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

CoreTrustSeal - Requirements

Requirements are reviewed & updated if necessary every 3 years (current requirements effective 2020-2022)
 Website Information: <https://www.coretrustseal.org/why-certification/requirements/>

Organizational Infrastructure	R1 Mission/Scope R2 Licenses R3 Continuity of access R4 Confidentiality/ethics R5 Organizational infrastructure R6 Expert guidance
Digital Object Management	R7 Data integrity and authenticity R8 Appraisal R9 Documented storage procedures R10 Preservation plan R11 Data quality R13 Data discovery R14 Data reuse
Technology	R15 Technical infrastructure R16 Security

Example Requirement

Digital Object Management

7. Data integrity and authenticity

R7. The repository guarantees the integrity and authenticity of the data.

Compliance Level:

Response

Guidance:

The repository should provide evidence to show that it operates a data and metadata management system suitable for ensuring integrity and authenticity during the processes of ingest, archival storage, and data access. This Requirement covers the entire data lifecycle within the repository.

To protect the integrity of data and metadata, any intentional changes to data and metadata should be documented, including the rationale and originator of the change. Measures should be in place to ensure that unintentional or unauthorized changes can be detected and correct versions of data and metadata recovered.

Authenticity covers the degree of reliability of the original deposited data and its provenance, including the relationship between the original data and that disseminated, and whether or not existing relationships between datasets and/or metadata are maintained.

For this Requirement, responses on data integrity should include evidence related to the following:

- Description of checks to verify that a digital object has not been altered or corrupted (i.e., fixity checks) from deposit to use.
- Documentation of the completeness of the data and metadata.
- Details of how all changes to the data and metadata are logged.
- Description of version control strategy.
- Usage of appropriate international standards and conventions (which should be specified).

Evidence of authenticity management should relate to the following questions:

- Does the repository have a strategy for data changes? Are data producers made aware of this strategy?
- Does the repository maintain provenance data and related audit trails?
- Does the repository maintain links to metadata and to other datasets? If so, how?
- Does the repository compare the essential properties of different versions of the same file? How?
- Does the repository check the identities of depositors?

CoreTrustSeal - Requirements

Recommend to use the *Extended Requirements*, as these give more clarifying context to each requirement. For example:

Workflows R12. *Archiving takes place according to defined workflows from ingest to dissemination.*

Response Guidance: To ensure the consistency of practices across datasets and services and to avoid ad hoc actions, workflows should be defined according to the repository's activities and clearly documented. Provisions for managed change should be in place. The OAIS reference model can help to specify the workflow functions of a repository. For this Requirement, responses should include evidence related to the following:

- Workflows/business process descriptions.
- Clear communication to depositors and users about handling of data.
- Levels of security and impact on workflows (guarding privacy of subjects, etc.).
- Qualitative and quantitative checking of outputs.
- The types of data managed and any impact on workflow.
- Decision handling within the workflows (e.g., archival data transformation).
- Change management of workflows. This Requirement confirms that all workflows are documented.

Extended Guidance: The reviewer is looking for evidence that the applicant takes a consistent, rigorous, documented approach to managing all activities throughout their processes and that changes to those processes are appropriately implemented, evaluated, recorded, and administered. The Requirement does not demand detailed descriptions of workflows, but seeks evidence of how and where these workflows are documented.

CoreTrustSeal - Clarifications

French support:

- Requirements: An unofficial French translation was created a couple years ago (can be made available upon request), but it is recommended to use the English one as a primary resource.
- Applications: The submission responses must be in English. If supporting evidence is not in English, it must be summarized in English in the response statement.

Review process:

- 2 reviewers from existing CTS-certified repositories are assigned to each submission
- the board reviews the reviews to ensure consistency and for added discussion before responding to applicant
- applicants can re-submit applications by addressing any feedback (this may take a couple rounds)
- if certification awarded, applicant is notified and the final application is posted to the CTS website

CoreTrustSeal - Clarifications

Designated Communities:

- From Requirements: “A clear definition of the Designated Community demonstrates that the applicant understands the scope, knowledge base, and methodologies—including preferred software/formats—of the user community or communities they are targeting.”
- From Extended Requirements: “it is possible...to have a Designated Community composed of different ‘sub-communities’...To serve its Designated Community well, a repository has to have a deep understanding of the Designated Community’s composition, skills, knowledge base, and needs, and how these may transform over time. Throughout the application, evidence should demonstrate an understanding of what the curation requirements are ... to best serve the Designated Community ... alongside a demonstration that the applicant monitors and responds to changes in the needs of the Designated Community.”

The CoreTrustSeal has recently completed community consultation for a certification strategy that is inclusive of [specialist, generalist and technical repository service providers](#). The board has been finalizing analysis of comments, with intention to complete and release final revisions in the near-term.

CoreTrustSeal - Certification Repositories

100th certified repository milestone reached with the Odum Institute Data Archive! (now 104)

- [Press release](#)
- [Application](#)



Home About ▾ Certification ▾ **Certified Repositories ▾** Apply ▾ Contact ▾

Core Certified Repositories

Home ▸ Why certification ▸ Core Certified Repositories

Applications are made public only once certification of a data repository has been approved by the CoreTrustSeal Board. Certification is against the version of the Core Trustworthy Data Repositories Requirements named in the link to the public application (e.g., 2017-2019). The CoreTrustSeal for Data Repositories is valid for three years from the certification date listed within the public application.



CoreTrustSeal - Re3Data

The screenshot shows the Re3Data website interface. On the left is a 'Filter' sidebar with categories: Subjects, Content Types, Countries, AID systems, API, Certificates, Data access, Data access restrictions, Database access, Database access restrictions, Database licenses, and Data licenses. The main content area shows a search bar, a search button, and a 'Toggle short help' link. Below the search bar is a pagination control showing '1' selected. The search results show 'Found 95 result(s)' and a result for 'Archaeology Data Service' (ADS) with subject tags like 'Ancient Cultures' and 'History'.

The screenshot shows the 'Repository details' page for the 'Australian Antarctic Data Centre'. The page has tabs for 'General', 'Institutions', 'Terms', and 'Standards'. The 'General' tab is active, showing the following information:

- Name of repository: Australian Antarctic Data Centre
- Additional name(s): Data management and spatial data services, AADC
- Repository URL: <https://data.aad.gov.au/>
- Subject(s): Biology, Geosciences (including Geography), Physics, Life Sciences, Natural Sciences
- Description: The Australian Antarctic Data Centre (AADC) provides data collection and data management services in Australia's Antarctic Science Program. The AADC manages science data from Australia's Antarctic research, maps Australia's areas of interest in the Antarctic region, manages Australia's Antarctic state of the environment reporting, and provides advice and education and a range of other products.
- Content type(s): Images, Scientific and statistical data formats, Plain text, Standard office documents
- Certificates and Standards: CoreTrustSeal
- Keyword(s): Antarctic Treaty System, Antarctic Data, Polar Data
- Persistent identifier(s) of the repository: RRID:nlx_152019, FAIRsharing_doi:10.25504/FAIRsharing.t11vm9, RRID:SCR_006320
- Repository type(s): disciplinary
- Research data repository language(s): eng
- Data and/or service provider: dataProvider, serviceProvider

Re3Data has integrated a feature to filter on and recognize CoreTrustSeal certified repositories

CoreTrustSeal - Related Analysis

Johnston, Lisa R. (2021). **Level of curation self-reported by 100 CoreTrustSeal certified repositories (2017-2019)**. Retrieved from the Data Repository for the University of Minnesota, <https://hdl.handle.net/11299/218332>.

- A. Content distributed as deposited (**1**)
- B. Basic curation – e.g., brief checking, addition of basic metadata or documentation (**14**)
- C. Enhanced curation – e.g., conversion to new formats, enhancement of documentation (**27**)
- D. Data -level curation – as in C above, but with additional editing of deposited data for accuracy (**51**)
- some blank...(7)

Lindlar M, Schwab F. **All that work ... for what? Return on Investment for Trustworthy Archive Certification Processes—a Case Study**. Proceedings of the 15th International Conference on Digital Preservation. Boston, MA; 2018. pp. 1–10. <https://phaidra.univie.ac.at/o:922206>

- “Stakeholder confidence, especially that of funders, improved and transparent documentation and process improvement are the three most important benefits to us. ”
- Other benefits also evaluated in the paper

CoreTrustSeal - Research Data Alliance

RDA/WDS Certification of Digital Repositories IG <https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html>

Upcoming session at the next RDA plenary in April
<https://www.rd-alliance.org/plenaries/rda-17th-plenary-meeting-edinburgh-virtual/principles-metrics-evaluation-increasing-trust>



Ocean Networks Canada

- Currently working on our own application for the CTS to renew our World Data System membership, as our original certification was through the WDS requirements in 2014
- Provides a means to communicate rationale for continually improving our own services, processes and documentation
- Utilized as a motivating factor for the recently completed CANARIE-funded MINTED (Making Identifiers Necessary for Tracking Evolving Data) project (see [Portage Network webinar recording from 2020-11](#))
- Certification has helped to promote stakeholder confidence in our capabilities
- World Data System has allowed us to be part of a community to share and disseminate ideas



A large school of fish, likely salmon, swimming in dark blue water. The fish are densely packed and moving in various directions, creating a sense of dynamic movement. The lighting is dim, highlighting the silvery scales of the fish against the deep blue background.

OCEAN
NETWORKS
CANADA

Questions?

reyna@uvic.ca

NDRIO- Portage CoreTrustSeal Certification Cohort & Funding

portage

SERVICES PARTAGÉS POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA





CoreTrustSeal Certification Cohort & Funding

The call for applications is now open!

What are we offering:



- A chance to be part of an initial cohort of Canadian repositories working together to achieve CoreTrustSeal certification;
- Training and facilitated workshops to ease the process, including feedback from peers and experts;
- CAD \$10,000 per participating repository to fund the work required and the application fee.



CoreTrustSeal Certification Cohort & Funding

Overview of the application process:

- Ensure your repository is eligible to apply for CoreTrustSeal certification.
- Ensure you meet the **eligibility** criteria for this call.
- Ensure you can provide evidence of the **scoring** criteria for this call.
- Prepare and submit your application!



CoreTrustSeal Certification Cohort & Funding

Eligibility criteria (mandatory):

- Be a Canadian repository;
- Have a mission that includes the stewardship and retention of data;
- Have identified a designated community;
- Provide, or intend to provide, curation of data.



CoreTrustSeal Certification Cohort & Funding

Scoring criteria (used to assign a **score** for selection decisions):

- Evidence of clear, senior management commitment to CoreTrustSeal certification.
- Identification of an office, team, or individual(s) who will participate in cohort activities and have lead writing responsibilities.
- Evidence that the repository serves its designated community.
- A clearly-articulated rationale for achieving CoreTrustSeal certification.



CoreTrustSeal Certification Cohort & Funding

Additional considerations (may be considered to assemble a diverse and balanced cohort):

- Regional distribution;
- Broad disciplinary coverage;
- Inclusion of repositories having previously achieved a repository certification.



Questions?

Contact us at portage@carl-abrc.ca

