



CoreTrustSeal

Assessing Trustworthiness

Trustworthy Data Repositories Workshop
National Institutes of Health

Session 1: Trust Concepts and Standards
Monday, April 8, 2019; 11:00 a.m. – 12:00 a.m.



Ingrid Dillo
Deputy Director, DANS
The Netherlands

DANS is about keeping data FAIR



Mission: promote
and provide
permanent
access to digital
research
resources



DANS

Institute of
Dutch Academy
and Research
Funding
Organisation
(KNAW & NWO)
since 2005

First predecessor
dates back to
1964 (Steinmetz
Foundation),
Historical Data
Archive 1989

<https://dans.knaw.nl>

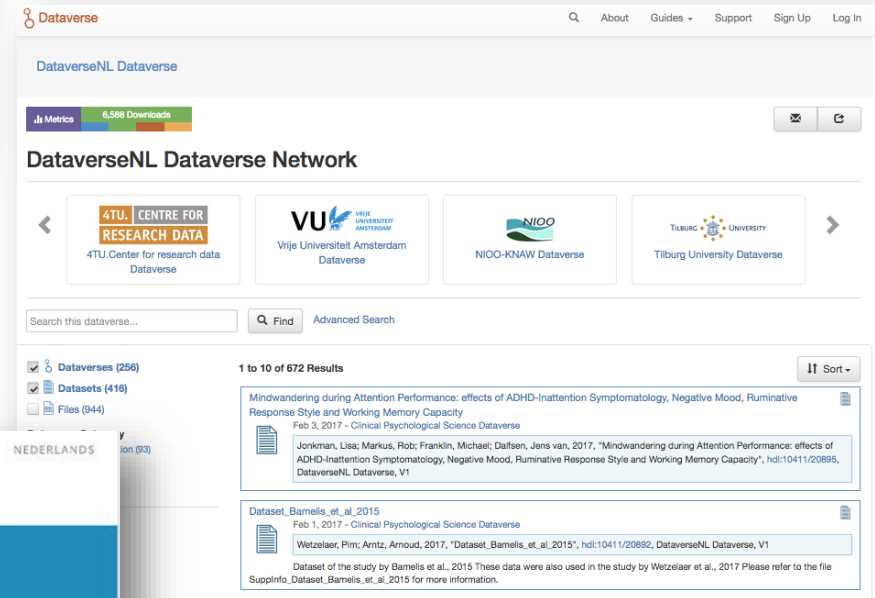
DANS Core Data Services



EASY: certified long-term repository



NARCIS: Gateway to scholarly information In the Netherlands

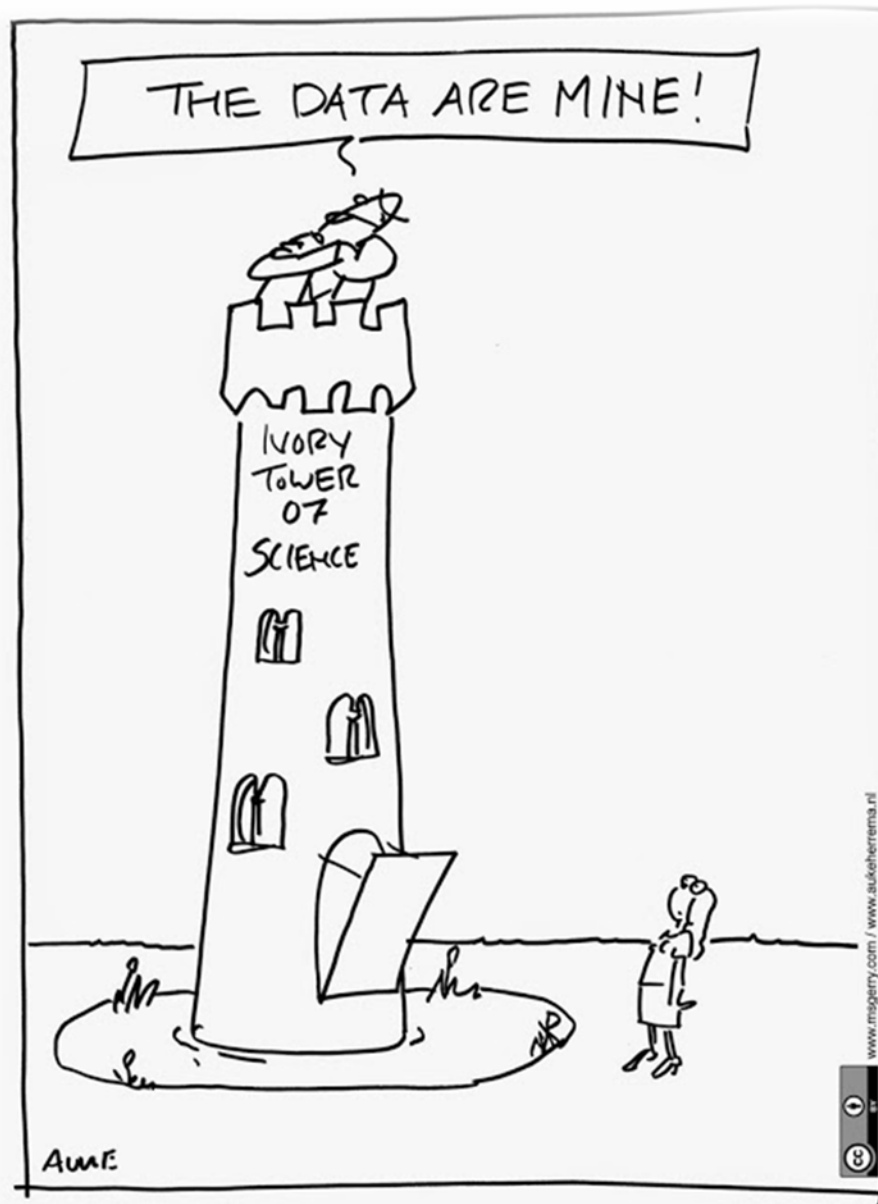


DataverseNL: data repository at universities and other institutions


Topics



- Requirements
- Process and procedures
- Current uptake
- Certification in practice
- Benefits of certification
- Future developments



SCENE FROM THE PAST ?



“Perhaps the biggest challenge in sharing data is trust: how do you create a system robust enough for scientists to trust that, if they share, their data won’t be lost, garbled, stolen or misused?”

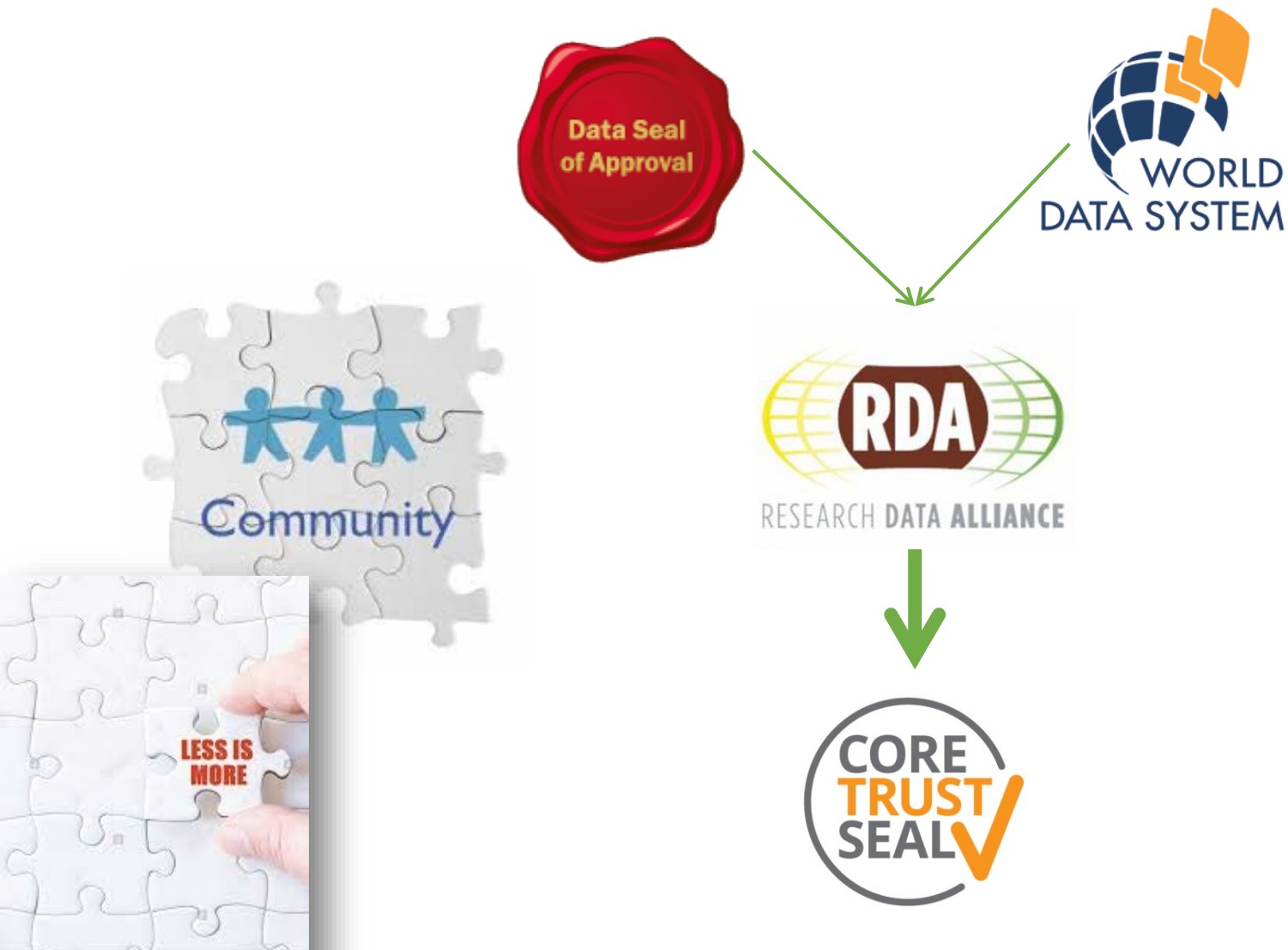
The Data Harvest:

How sharing research data can yield knowledge, jobs and growth

An RDA Europe Report

December 2014

CoreTrustSeal: a brief history



CoreTrustSeal

The objectives of the CoreTrustSeal are to safeguard data, to ensure high quality and to guide reliable management of data for the future without requiring the implementation of new standards, regulations or heavy investments.

CoreTrustSeal repository certification:

- Gives **data producers** the assurance that their data and associated materials will be stored in a reliable manner and can be reused;
- Provides **funding bodies** with the confidence that data will remain available for reuse;
- Enables **data consumers** to assess the repositories where data are held;
- Supports **data repositories** in the efficient archiving and distribution of data.





Requirements: background

The requirements are based on:

- OAIS reference model
- Designated community
- Sharing expertise and effort
- Stakeholder groups
 - Data producers (research integrity, formats, documentation)
 - Data consumers (terms of use, licences, citation)
 - Repositories (enabler; quality of storage, organisational processes, technical infrastructure, assurance of availability)

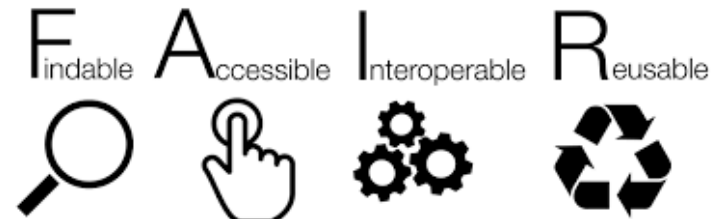


Requirements: background

Fundamental to the requirements are five criteria that together determine whether or not the digital data may be considered as sustainably archived:

- The data can be **found** on the Internet;
- The data are **accessible**, while taking into account relevant legislation with regard to personal information and intellectual property;
- The data are available in a **usable format**;
- The data are **reliable**;
- The data can be **referred** to (persistent identifiers).

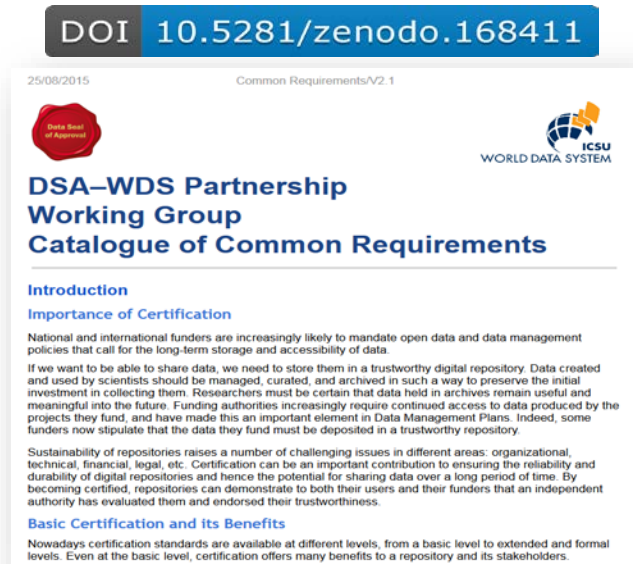
→ Strong link with:



16 Requirements

Categories:

- Background information (R0)
- Organizational infrastructure (R1-6)
- Digital object management (R7-14)
- Technology and security (R15-16)
- Applicant feedback



Compliance levels



- applicant to indicate a compliance level for each requirement:
 - 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
- applicants judged against statements supported by appropriate evidence; not against self-assessed compliance levels
- if a requirement is not applicable, the reason for this must be documented in detail
- certification can be granted even if some requirements are at level 3; requirements include an assumption of continuous improvement

Core TDR Requirements



Background information

R0 Please provide **context** for your organization

Organizational infrastructure

R1. The repository has an **explicit mission** to provide access to and preserve data in its domain.

R2. The repository maintains all applicable **licenses** covering data access and use and monitors compliance.

R3. The repository has a **continuity plan** to ensure ongoing access to and preservation of its holdings.

R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with **disciplinary and ethical norms**.

R5. The repository has **adequate funding** and sufficient numbers of **qualified staff** managed through a clear system of governance to effectively carry out the mission.

R6. The repository adopts mechanism(s) to secure ongoing **expert guidance** and feedback (either in-house, or external, including scientific guidance, if relevant).



Core TDR Requirements



Digital object management

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

R8. The repository accepts data and metadata based on **defined criteria to ensure relevance and understandability** for data users.

R9. The repository applies **documented processes and procedures** in managing archival storage of the data.

R10. The repository assumes responsibility **for long-term preservation** and manages this function in a planned and documented way.

R11. The repository has appropriate expertise to **address technical data and metadata quality** and ensures that sufficient information is available for end users to make quality-related evaluations.

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

R13. The repository enables users to **discover the data and refer to them in a persistent way** through proper citation.

R14. The repository enables reuse of the data over time, ensuring that **appropriate metadata** are available to support the understanding and use of the data.



Core TDR Requirements



Technology and security

R15. The repository functions on **well-supported operating systems and other core infrastructural software** and is using hardware and software technologies appropriate to the services it provides to its Designated Community.

R16. The technical infrastructure of the repository provides for **protection** of the facility and its data, products, services, and users.

Applicant feedback



Topics



- Requirements
- Process and procedures
- Current uptake
- Certification in practice
- Benefits of certification
- Future developments



Two step certification process

Self assessment based on 16 Requirements (written responses + URLs of documented public evidence + compliance level)

Peer review by two expert and independent reviewers under the responsibility of the CoreTrustSeal Standards and Certification Board

- Successful applications are made publicly available
- Administrative fee of 1,000 euro
- Certification valid 3 years



Application process

Apply for CoreTrustSeal Data Repository certification

CoreTrustSeal Application Management Tool (AMT)

Please click on the link below to access the AMT and create an account. After login, you will be asked to request an application form (setup your repository information), and the application process will then begin upon submission. We hope that the new Tool is fairly self-explanatory, but please do not hesitate to contact the [CoreTrustSeal Secretariat](#) with any questions or issues.

• CoreTrustSeal Application Management Tool

Note also that resources to help you complete the application can be found on the [Requirements page](#) of the [CoreTrustSeal website](#).


Important notice: Since 1 February 2018, an [administrative fee](#) of EUR 1,000 is charged for the review of a CoreTrustSeal certification. More information about the administrative fee is available on [this page](#).



Example:

XIV. Data reuse

R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.

Compliance Level 

Response

Guidance:

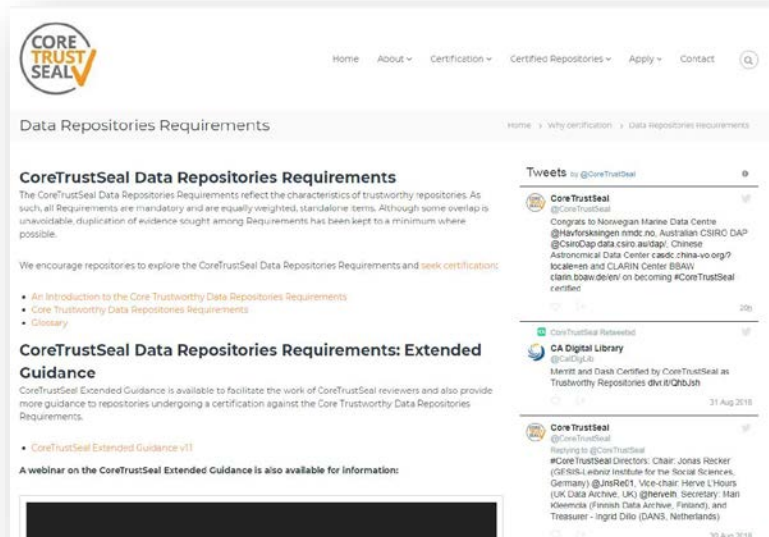
Repositories must ensure that data can be understood and used effectively into the future despite changes in technology. This Requirement evaluates the measures taken to ensure that data are reusable.

For this Requirement, responses should include evidence related to the following questions:

- Which metadata are required by the repository when the data are provided (e.g., Dublin Core or content-oriented metadata)?
- Are data provided in formats used by the Designated Community? Which formats?
- Are measures taken to account for the possible evolution of formats?
- Are plans related to future migrations in place?
- How does the repository ensure understandability of the data?

Reuse is dependent on the applicable licenses covered in R2 (Licenses).

Resources /help

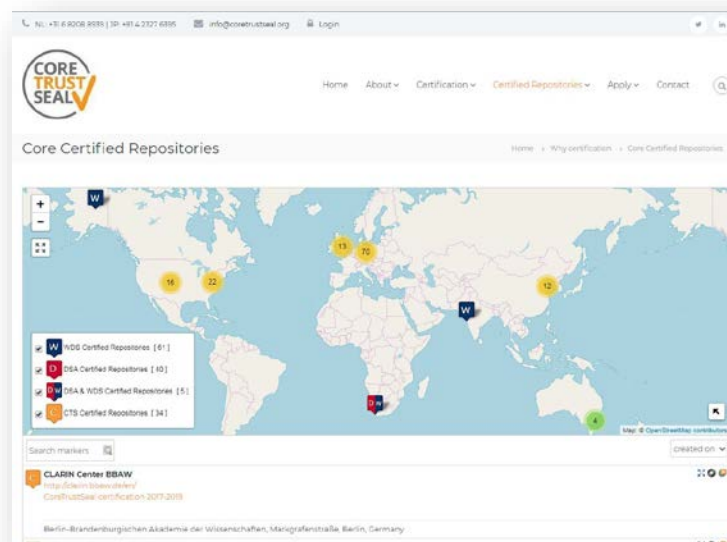


<https://www.coretrustseal.org/why-certification/certified-repositories/>

- Library of public applications; all are certified and so can be considered exemplars.


www.coretrustseal.org/why-certification/requirements/

- Extended Guidance and a webinar.
- The Extended Guidance is intended for reviewers, but is useful for applicants.



Application Management Tool (ATM)

📍 The Hague, Netherlands 📞 +31 6 2386 3243 📍 Tokyo, Japan 📞 +81 4 2327 6395 ✉ info@coretrustseal.org [Twitter](#) [LinkedIn](#)



Data Repository Certification: Application Management Tool


Enter your Email Address and Password

Email Address:


Password:

Note: If you were redirected to this page, this means you need an account to login and complete your action. Please sign up for an account using the link below.

[Sign Up to CoreTrustSeal AMT](#)
[Forgot your password?](#)

 Except where otherwise noted, content on this website is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#) by CoreTrustSeal

📍 The Hague, Netherlands 📞 +31 6 2386 3243 📍 Tokyo, Japan 📞 +81 4 2327 6395 ✉ info@coretrustseal.org [Twitter](#) [LinkedIn](#)



[Rorie \(Applicant\)](#) [Log out](#)

Data Repository Certification: Application Management Tool

Info Welcome! You are now logged in.

Main Page

Welcome to the CoreTrustSeal Application Management Tool (AMT).

Start CoreTrustSeal Application Process

Please click on the below link to submit your organization's profile and begin the application process for CoreTrustSeal repository certification. Your application request will be sent to the CoreTrustSeal Administrators, who may contact you to ask for any clarifications and to provide you with additional information.

[Request Application Form](#)

[Application Form](#)


[Add / Remove Secondary Contacts](#)

Apply via: <https://www.coretrustseal.org/apply/>



Application Management Tool (ATM)

The Hague, Netherlands+31 6 2306 3243Tokyo, Japan+81 4 2327 6395info@coretrustseal.orgTwitterLinkedIn



Rorie (Applicant)Log out

Data Repository Certification: Application Management Tool

Organization Profile

Organization Name **Required**

Address **Required**

City **Required**

Country **Required**

No value

Website **Required**

Host Institute **Required**

DOI of re3data Entry

Please access <https://www.re3data.org/> to find the DOI of your entry, or submit an entry for your repository if it does not exist.

Representative Name **Required**


Email **Required**

Telephone


Brief Description of Repository

SubmitCancel

Move to top

Except where otherwise noted, content on this website is licensed under a [Creative Commons Attribution 3.0 Unported License](#) by CoreTrustSeal

The Hague, Netherlands+31 6 2306 3243Tokyo, Japan+81 4 2327 6395info@coretrustseal.orgTwitterLinkedIn



Rorie (Applicant)Log out

Data Repository Certification: Application Management Tool

InfoItem state changed

Main Page

A Repository

Application Review Progress

Draft

Submitted or Feedback Requested

Under Review

Feedback Available

Approved

Organization Profile

CoreTrustSeal Application Form

The online form for your repository to certify against the Core Trustworthy Data Repositories Requirements can be accessed below. The form does not need to be filled out on a single occasion, and you may save your draft, request simple feedback (assistance) from the CoreTrustSeal Admin, or submit the completed form for review.

Application Status: Draft

Please note: The form is editable **ONLY** when it has status 'Draft' or 'Feedback Available'.

Application Form


Contributors

Completing the CoreTrustSeal application will very likely require discussion with other data management and technical experts in your organization. To share the application with your colleagues, please add them as contributors using the link below. It is recommended to include at least one person as an additional contact for your organization.

Rorie (Applicant)[rorie.edmunds+user@iccu-wds.org]

Add / Remove Secondary Contact

Move to top

Except where otherwise noted, content on this website is licensed under a [Creative Commons Attribution 3.0 Unported License](#) by CoreTrustSeal



Recap application process



- Self-assessment based on requirements
- Online tool available
- Extended guidance (document and webinar) available
- Evidence provided:
 - Bullet lists of topics/questions for discussion and inclusion are neither exhaustive nor prescriptive. Highly preferred that responses are written in full prose.
 - URLs to evidence strongly encouraged
 - Maturity ratings strongly encouraged
 - Responses in English
 - Evidence documents in other languages need a short summary in English
- Review of the self assessment by two reviewers under the responsibility of the CTS Board



Governance

CoreTrustSeal Board 2018-2021

<https://www.coretrustseal.org/about/standards-and-certification-board/>



Jonas Recker (Chair)



Hervé L'Hours (vice chair)



Mari Kleemola (Secretary)



Ingrid Dillo (Treasurer)

Reyna Jenkins



Paul Trilsbeek

Bob Downs



John Faundeen



Mustapha Mokrane

Jon Crabtree



Wim Hugo

Dawei Lin



Ilona von Stein



Lindsey Callaghan



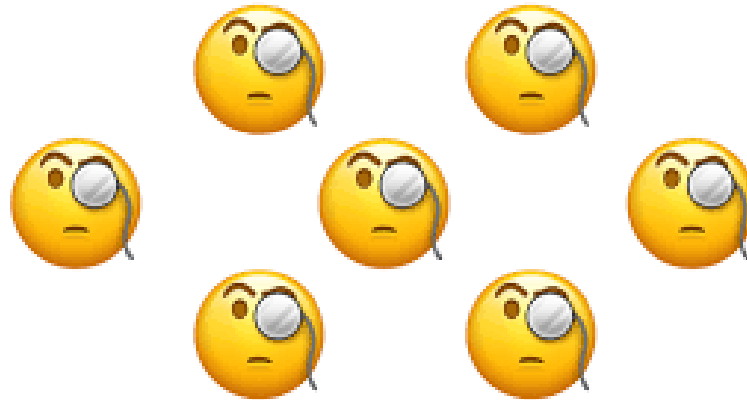
Rorie Edmunds



Governance

Assembly of reviewers:

<https://www.coretrustseal.org/about/assembly-of-reviewers/>



More information on governance in the Rules of Procedures:

<https://doi.org/10.5281/zenodo.1142960>



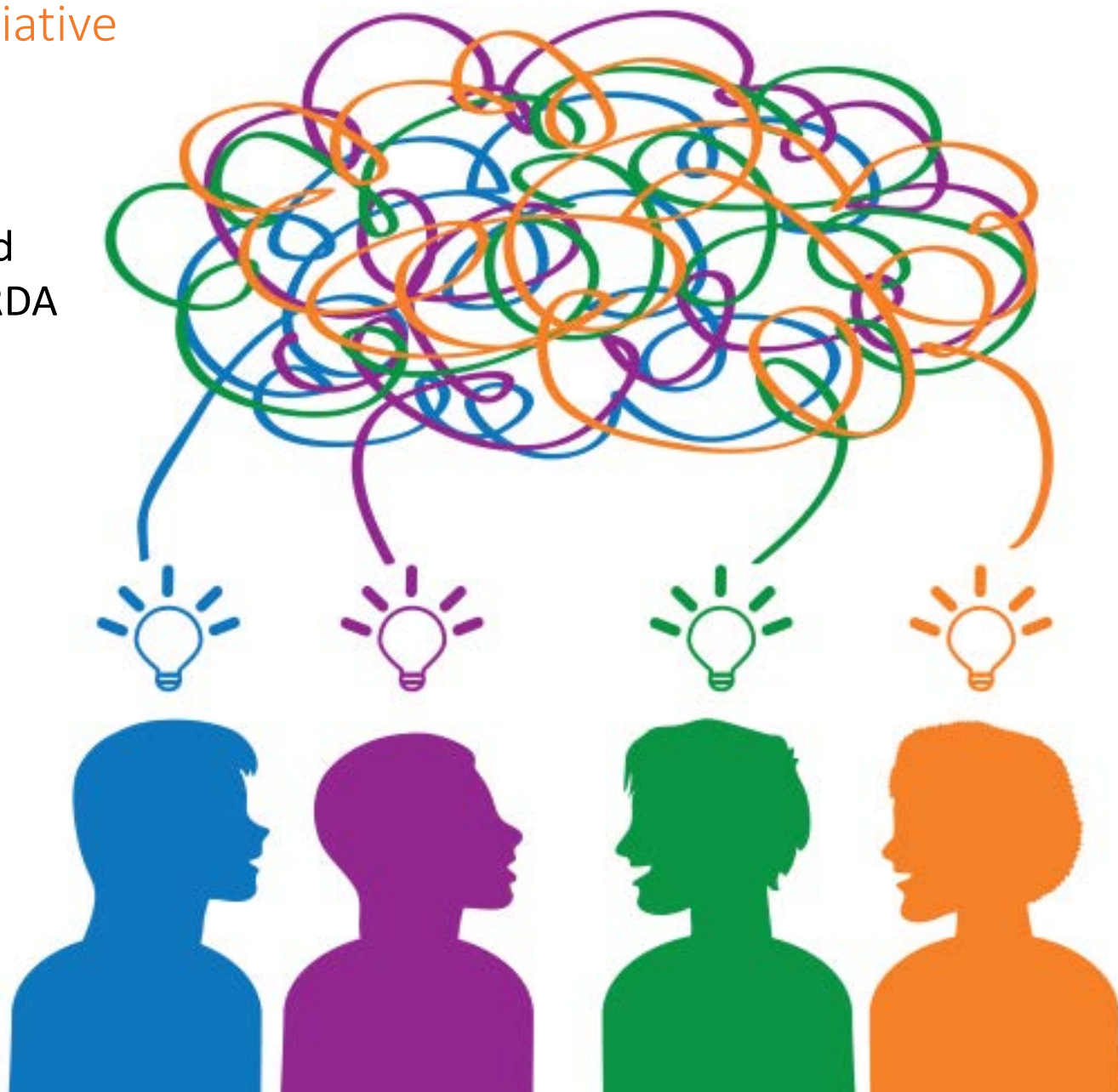


Members

1. **Alex de Sherbinin** (CIESIN/SEDAC)
2. **Alison Fernie** (UK National Geoscience Data Centre)
3. **Andrii Shelestov** (Ukrainian Geospatial Data Center)
4. **Aude Chambodut** (International Service of Geomagnetic Indices)
5. **Birger Jerlehag** (Swedish National Data Service)
6. **Dave Connell** (Australian Antarctic Data Centre)
7. **Dawei Lin** (ImmPort Repository, DAIT-NIAID-NIH, CoreTrustSeal Board)
8. **Françoise Genova** (Centre de Données astronomiques de Strasbourg: CDS)
9. **Garry Baker** (UK National Geoscience Data Centre)
10. **Guoqing Li** (World Data System)
11. **Hervé L'Hours** (UK Data Archive; CoreTrustSeal Board)
12. **Hiroaki Toh** (WDC – Geomagnetism, Kyoto)
13. **Ian Bruno** (The Cambridge Crystallographic Data Centre)
14. **Ilona von Stein** (Data Archiving and Networked Services: DANS; CoreTrustSeal Board)
15. **Ingrid Dillo** (Data Archiving and Networked Services: DANS; CoreTrustSeal Board)
16. **John Faundeen** (WDC – Earth Resources Observation and Science: EROS; CoreTrustSeal Board)
17. **John Howard** (Irish Social Science Data Archive)
18. **John Westbrook** (Worldwide Protein Data Bank)
19. **Jonas Recker** (GESIS Data Archive for the Social Sciences; CoreTrustSeal Board)
20. **Jonathan Crabtree** (Odum Institute Data Archive, CoreTrustSeal Board)
21. **Kehe Wang** (WDC – Space Weather, Australia)
22. **Lisa Johnston** (The Data Repository for University of Minnesota)
23. **Kerstin Lehnert** (Interdisciplinary Earth Data Alliance: IEDA)
24. **Maja Dolinar** (ADP Social Science Data Archives)
25. **Mari Kleemola** (The Finnish Social Science Data Archive; CoreTrustSeal Board)
26. **Marjan Grootveld** (Data Archiving and Networked Services: DANS)
27. **Monika Zarnitz** (ZBW – Leibniz Information Centre for Economics)
28. **Mustapha Mokrane** (World Data System; CoreTrustSeal Board)
29. **Niels H. Batjes** (ISRIC – WDC Soils)
30. **Olivier Rouchon** (CINES)
31. **Paul Trilsbeek** (Max Planck Institute for Psycholinguistics; CoreTrustSeal Board)
32. **Reyna Jenkyns** (Ocean Networks Canada; CoreTrustSeal Board)
33. **Robert R. Downs** (CIESIN/SEDAC; CoreTrustSeal Board)
34. **Rorie Edmunds** (World Data System; CoreTrustSeal Board)
35. **Steven McEachern** (Australian Data Archive)
36. **Toshihiko Iyemori** (WDC – Geomagnetism, Kyoto)
37. **Wim Hugo** (World Data System; CoreTrustSeal Board)
38. **Xiukuan Zhao** (WDC for Geophysics, Beijing)

CoreTrustSeal initiative

- Not for profit
- Community based
- Strong ties with RDA
- Global
- Domain agnostic



Topics



- Requirements
- Process and procedures
- **Current uptake**
- Certification in practice
- Benefits of certification
- Future developments

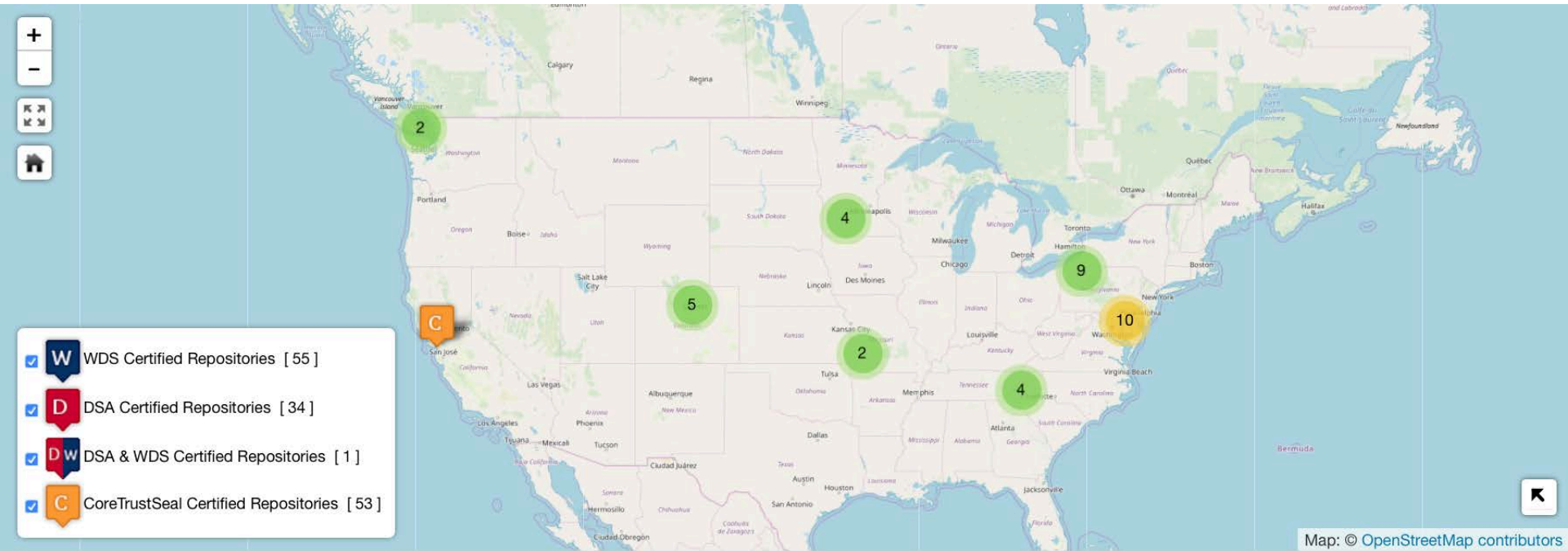
Current uptake



<https://www.coretrustseal.org/why-certification/certified-repositories/>



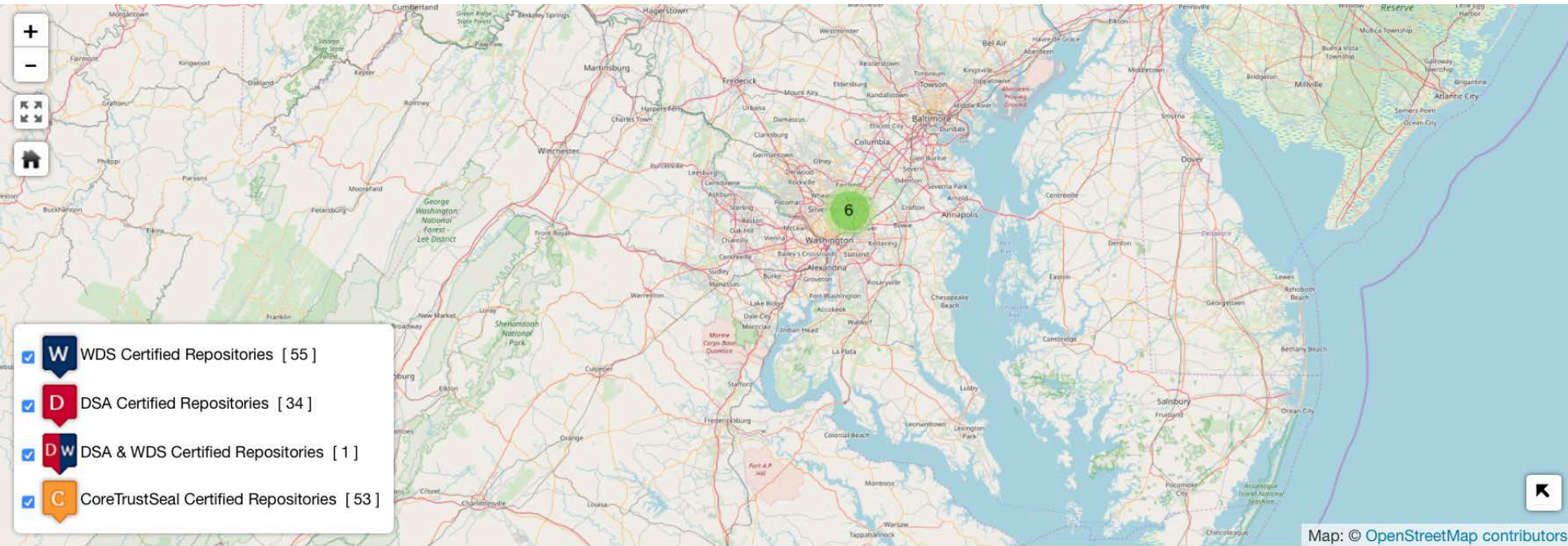
Current uptake



<https://www.coretrustseal.org/why-certification/certified-repositories/>



Current uptake



<https://www.coretrustseal.org/why-certification/certified-repositories/>



Topics



- Requirements
- Process and procedures
- Current uptake
- **Certification in practice**
- Benefits of certification
- Future developments

Certification in practice: broad scope

- Broad range of topics
 - organizational, staffing, financial and legal aspects, archival workflows, IT-infrastructure, risk management, etc.
- Properly describing policies, processes, etc.
- Development of missing policies, processes, IT- and infrastructural elements, etc.



Certification in practice: organizational aspects

- Responsibility for achieving the target on management level
- Core certification team: planning, discussing, monitoring and partly executing the work
- Many colleagues within the institute with specific expertise temporarily involved in the actual work



Certification in practice: the effort involved

- Highly dependent on your level of entry
- The effort will rise, if you still need to do real work in order to comply with the requirements
- The effort will rise when you chose more stringent certification standards like DIN31644 or ISO16363.



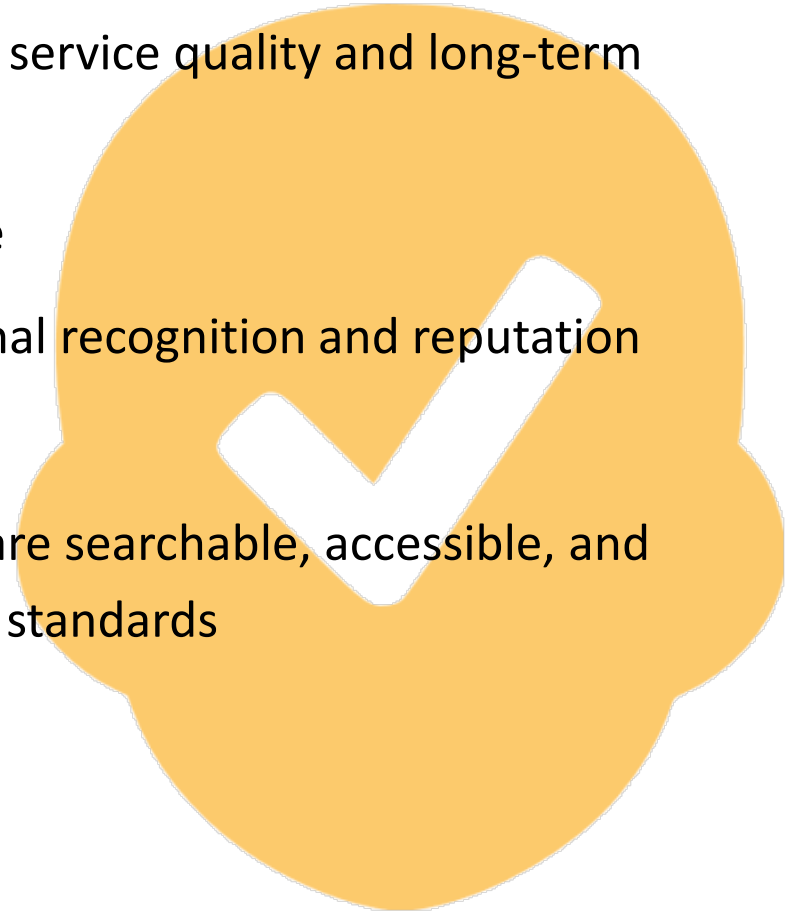
Topics



- Requirements
- Process and procedures
- Current uptake
- Certification in practice
- Benefits of core certification
- Future developments

Benefits of Core Certification: external

- Displays commitment to data and service quality and long-term data curation
- Heightens stakeholder confidence
- Increases national and international recognition and reputation
- Increases your visibility
- Show data holdings and services are searchable, accessible, and satisfy national and international standards



Benefits of Core Certification: internal

- Benchmark for comparison/ determine strengths and weaknesses
- Improves professionalism:
 - Checking, improving and updating policy and workflow documents
 - Re-evaluating and making improvements on our technical solutions and processes for long-term preservation
- Improves awareness and compliance with established standards
- Increases internal communication
- Good team building exercise
- Ensuring transparency





NCDD survey: perceived benefits

“The benefits propagated ... are in line with the perception of the respondents: this is most clearly the case with the stated benefit “**awareness raising** about digital preservation,” followed by “**stakeholder confidence**.”

“When queried about *other* perceived benefits, it is clear that the certification process not only led to external benefits but also to **improved internal processes, documentation** and opportunities to attract data producers as well as data consumers.”

“The majority of the respondents rated the **ratio** between investments and benefits as “adequate-rewarding” to “rewarding-excellent.”

[http://www.ncdd.nl/wp-content/uploads/2016/10/201611_DE_Ho
udbaar_Report_DSA-survey_2016.pdf](http://www.ncdd.nl/wp-content/uploads/2016/10/201611_DE_Houdbaar_Report_DSA-survey_2016.pdf)



Topics



- Requirements
- Process and procedures
- Current uptake
- Certification in practice
- Benefits of core certification
- Future developments

Review of TDR Requirements

- 3 year cycle of review (2017-2019)

[Home](#)[About ▾](#)[Certification ▾](#)[Certified Repositories ▾](#)[Apply ▾](#)[Contact](#)

Review of Requirements

[Home](#) > [Why certification](#) > [Review of Requirements](#)

A message to the CoreTrustSeal community:

A review of the CoreTrustSeal will take place in 2019 to define the Requirements for the period 2020–2023. This has no impact on the certifications of current CoreTrustSeal-certified repositories, which continue to run for three years from the date awarded.

The 2019 review process will focus on applicant feedback received during past reviews, other feedback received during communications and outreach activities, and an **open review period to run from 1 March 2019 to 30 April 2019**. Given the feedback received to date and the fact that a number of past WDS and DCA repositories are continuing to transition to the CoreTrustSeal Requirements, we request the members

Upcoming Events

NIH Workshop on Trustworthy Data Repositories for Biomedical Sciences

8 April @ 13:00 –19:00 UTC+0

[View All Events](#)



<https://www.coretrustseal.org/why-certification/meeting-community-needs/>

Increasing the scope of applicants

- Traditional focus on domain repositories
- Interest from:
 - national archives and libraries
 - infrastructure providers
 - repository software providers
 - bit-level replication services
 - commercial services

<https://www.coretrustseal.org/why-certification/meeting-community-needs/>



[Home](#) [About](#) [Certification](#) [Certified Repositories](#) [Apply](#) [Contact](#)



Meeting Community Needs

[Home](#) > [Why certification](#) > [Meeting Community Needs](#)

Exploring Opportunities for Expanding CoreTrustSeal Certification to Meet Community Needs

CoreTrustSeal is a community-based nonprofit organization that promotes sustainable and trustworthy data infrastructures by offering professional certification tools and services for data repositories and preservation-focussed institutions around the globe.

Upcoming Events

NIH Workshop on Trustworthy Data Repositories for Biomedical Sciences

8 April @ 13:00 -19:00 UTC+0

[View All Events](#)



European ICT Technical Specification



- The rules on European standardisation allow the European Commission to identify ICT technical specifications - that are not national, European or international standards - to be eligible for referencing in public procurement.



- Thorough external evaluation by European Multi Stakeholder Platform on ICT Standardisation based on very precise requirements



FAIRytale?

FAIR-CoreTrustSeal complementarity



“Research data will not become nor stay FAIR by magic. We need skilled people, transparent processes, interoperable technologies and collaboration to build, operate and maintain research data infrastructures.”

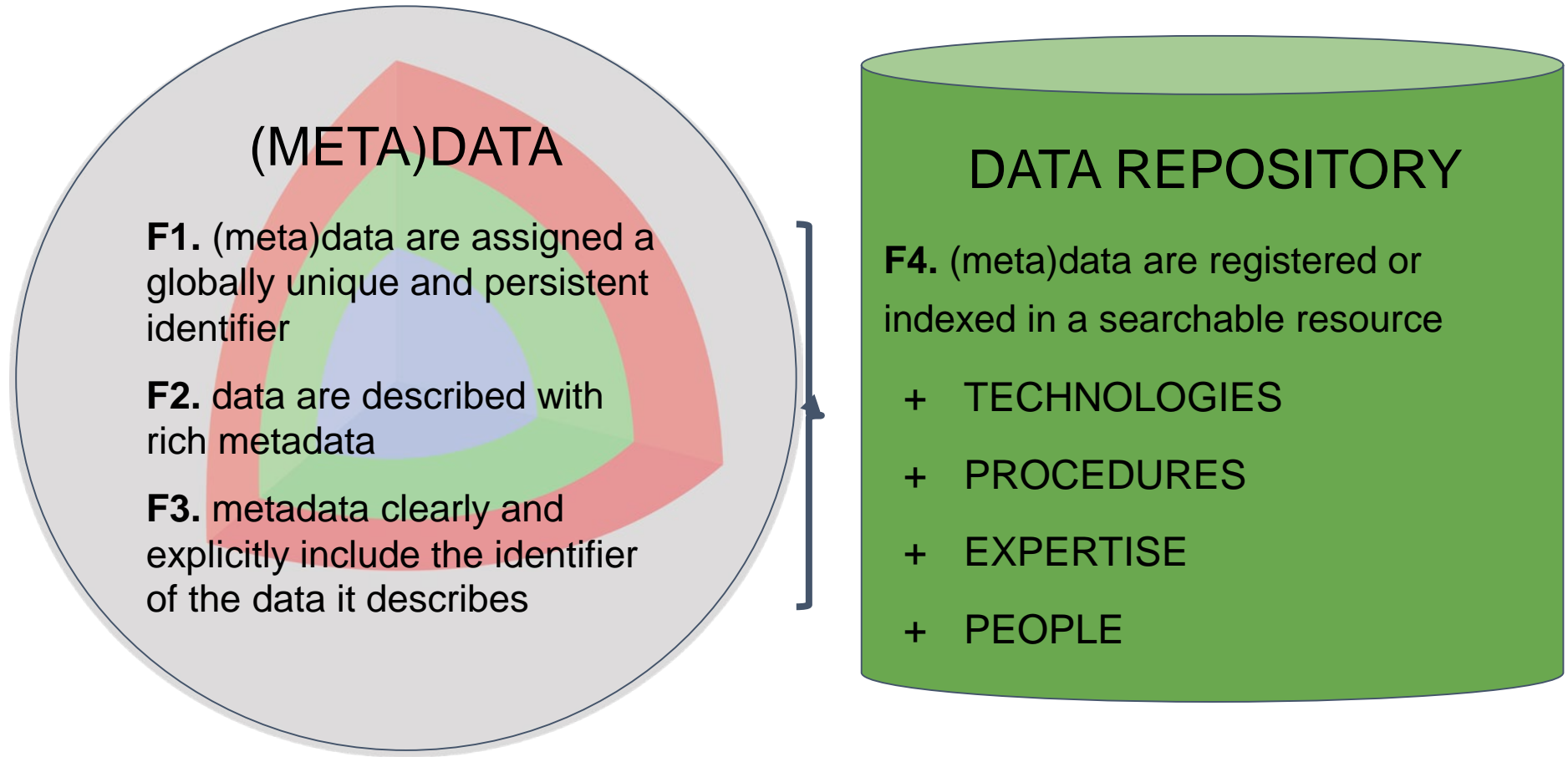
Mari Kleemola, Finnish Social Science Data Archive/CoreTrustSeal Board, Secretary

<https://tietoarkistoblogi.blogspot.com/2018/11/being-trustworthy-and-fair.html>





FAIR data assessment: levels



TDR to guarantee baseline data FAIRness

- Majority of CoreTrustSeal requirements (indirectly) refer to the FAIRness of the repository holdings
- Baseline of data FAIRness, but:
- Some data will be more FAIR than others!



TRUST Principles

- **FAIR** defines the properties of data and metadata
- **TRUST** describes the characteristics of data repositories that are responsible for managing and disseminating the data over a long period of time
- **FAIR** data in repositories we **TRUST**

T - Transparency is achieved by providing publicly accessible evidence of the services that a repository can and can not offer.

R - Responsibility is a commitment to provide high (technical) quality data services.

U - User community is the focus on the uses and potential uses of the data and services offered.

S - Sustainability is the capability to support long-term data preservation and use.

T - Technology is the infrastructure and capabilities to support the repository operations.



TRUST Principles White Paper

The intent is to invite stakeholders' input to develop concise and measurable guidelines for those wishing to build and sustain a trustworthy data repository that provides services to make datasets FAIR, as open as possible and as closed as necessary, and citable.

- Version 0.01
- Dawei Lin, Jonathon Crabtree, Ingrid Dillo, Robert R. Downs, Rorie Edmunds, Wim Hugo, and Mustapha Mokrane
- Link: <https://bit.ly/2lh7g8F>



Thank you for listening



ingrid.dillo@dans.knaw.nl

www.dans.knaw.nl

www.coretrustseal.org