CoreTrustSeal
Enabling FAIR Data Policies

e-IRG Workshop May 2019
Content Block A: Research Data Alliance
CERN, Geneva

Ingrid Dillo
Deputy Director, DANS
The Netherlands

DANS is about keeping data FAIR

https://dans.knaw.nl

Topics

- Data sharing and trust
- CoreTrustSeal catalogue and procedures
- Benefits of certification
- Future developments

Data sharing is important

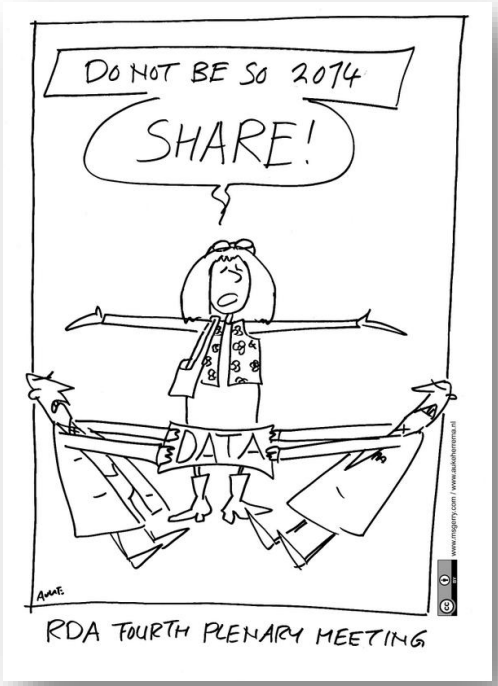
Growing recognition of the value of data:

- Replication and validation of research outcomes: transparency and integrity of science
- Re-use of data: efficiency in research , return on public investment

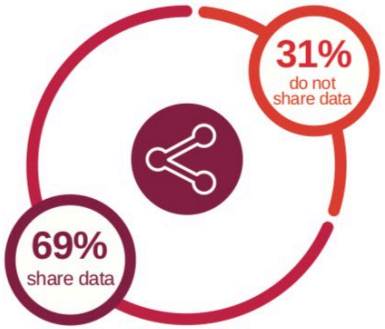
- Funder requirements: open data
- Publisher requirements: DAPs

CORE TRUST SEAL

Sharing practice



Data sharing in 2016



More than two thirds of Wiley researchers reported they are now sharing their data. Though this varies geographically and across research disciplines we are seeing that more researchers are sharing their data and taking efforts to make it reproducible. Archiving in institutional repositories, public repositories, and personal web pages has almost doubled since 2014.

Ways data is shared

41%
As supplementary material in a journal

10%
Discipline-specific data repository (e.g. GenBank, OpenEI, Protein Data Bank, TreeBASE)

29%
Personal, institutional, or project webpage

6%
General-purpose data repository (e.g. Dryad, figshare)

25%
Institutional data repository (i.e. university or institute-sponsored)

Researchers also report sharing their data in other ways including: 49% are sharing their data at conferences while 34% of researchers share their data upon informal request (email, direct contact, etc).

https://authorservices.wiley.com/asset/photos/licensing-and-open-access-photos/Wiley%20Global%20Data%20Sharing%20Infographic%20June%202017.pdf

CORE TRUST SEAL

Sharing practice

”36% of respondents have lost data on which they were working and there is, unsurprisingly, a high correlation between the vehicle for storing data and where it was lost - computer hard drives were the most common culprit here.”

Science, Digital; Hahnel, Mark; Treadway, Jon; Fane, Briony; Kiley, Robert; Peters, Dale; et al. (2017): The State of Open Data Report 2017. figshare. Paper. https://doi.org/10.6084/m9.figshare.5481187.v1

Enabling researchers

- Awareness raising
- Training (RDM)
- Infrastructure (VREs, TDRs, ..)


“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

Data repositories

re3data.org

Search | Browse | Suggest | Resources | Contact | DataCite

Filter

- Subjects
- Content Types
- Countries
- AID systems
- API
- Certificates
- Data access
- Data access restrictions
- Database access
- Database access restrictions
- Database licenses
- Data licenses
- Data upload
- Data upload restrictions
- Enhanced publication
- Institution responsibility type
- Institution type
- Keywords
- Metadata standards
- PID systems

Search... Search

Toogle short help

← Previous 1 2 3 4 5 6 7 ... 92 Next →

Sort by

Found 2299 result(s)

CancerData.org

Sharing data for cancer research

Subject(s)	Basic Biological and Medical Research, Medicine, Biology, Life Sciences
Content type(s)	Standard office documents, Databases, Images, Structured graphics, Scientific and statistical data formats, Raw data, Plain text, Archived data, other
Country	Netherlands

The CancerData site is an effort of the Medical Informatics and Knowledge Engineering
Our activities in the field of medical image analysis and data modelling are visible in a
datasets. They are grouped in collections and can be public or private. You can search
image archives without logging in.

Claremont Colleges Digital Library

CORE TRUST SEAL

Pillars of trust

- actions and attributes of the trustee (integrity, transparency, competence, predictability, guarantees, positive intentions)

- external acknowledgements:
  - reputation (researchers)
  - third party endorsements (funders, publishers)

Different assessments available

CORE TRUST SEAL

Topics

- Data sharing and trust
- CoreTrustSeal catalogue and procedures
- Benefits of certification
- Future developments

CoreTrustSeal: a brief history

Data Seal of Approval

WORLD DATA SYSTEM

RDA
RESEARCH DATA ALLIANCE

CORE TRUST SEAL

RDA output

https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:72520
https://zenodo.org/record/1406133#.XNqvmC-Q3OQ

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:

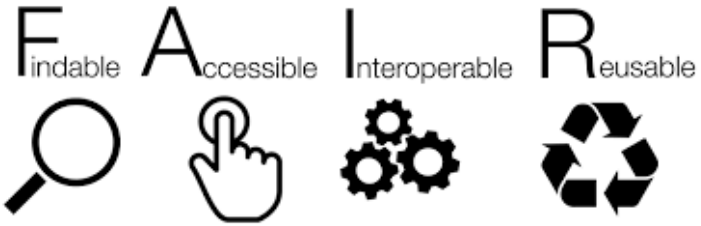
- Gives **data producers** the assurance that their data and associated materials will be stored in a reliable manner and can be reused;
- Provides **funding bodies** with the confidence that data will remain available for reuse;
- Enables **data consumers** to assess the repositories where data are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

Requirements: background

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:

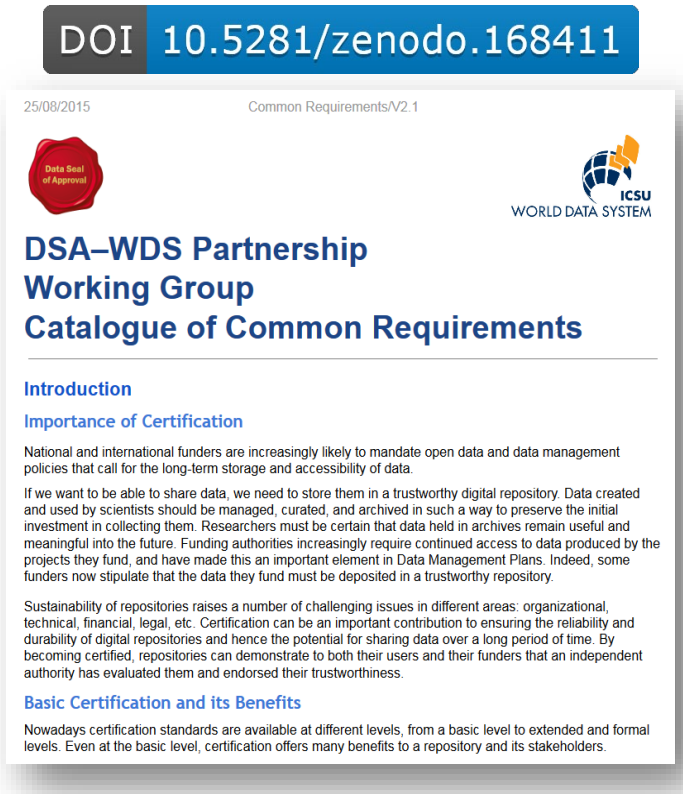


CORE TRUST SEAL

16 Requirements

Categories:

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



CORE TRUST SEAL

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 TRUST SEAL

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

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).

CORE TRUST SEAL

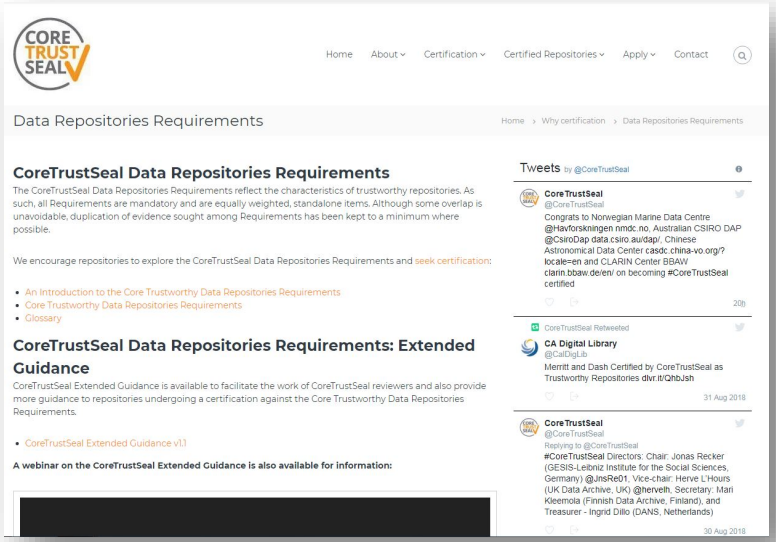
Two step certification process

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

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

Resources

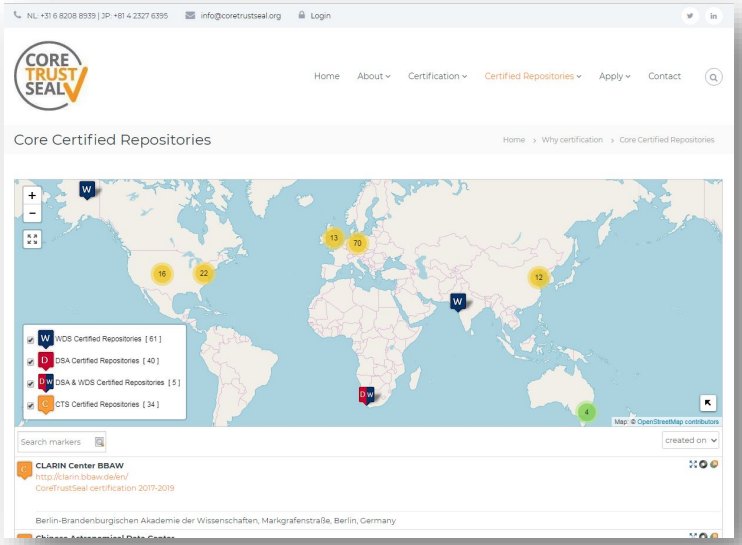


www.coretrustseal.org/why-certification/requirements/

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

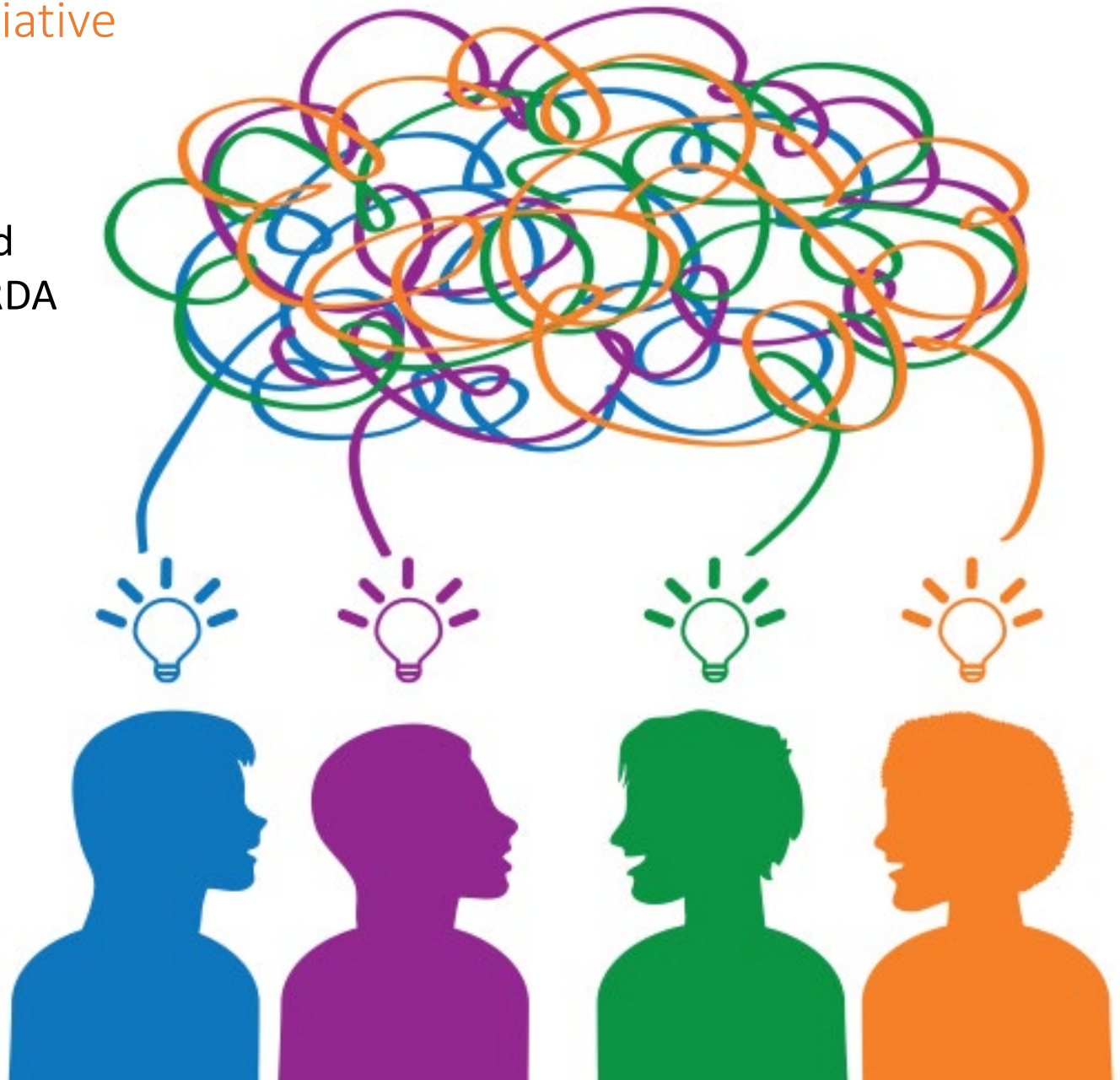
https://www.coretrustseal.org/why-certification/certified-repositories/

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



CoreTrustSeal initiative

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



Current uptake

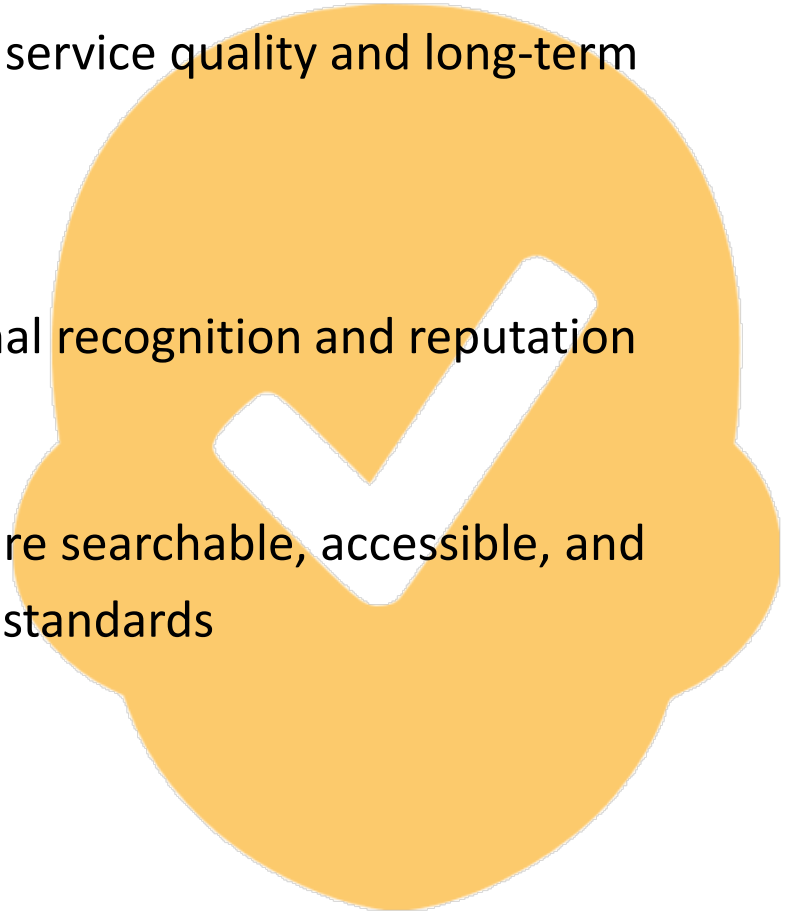
https://www.coretrustseal.org/why-certification/certified-repositories/

Topics

- Data sharing and trust
- CoreTrustSeal catalogue and procedures
- Benefits of 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



CORE TRUST SEAL

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

Topics

- Data sharing and trust
- CoreTrustSeal catalogue and procedures
- Benefits of certification
- Future developments

European ICT Technical Specification

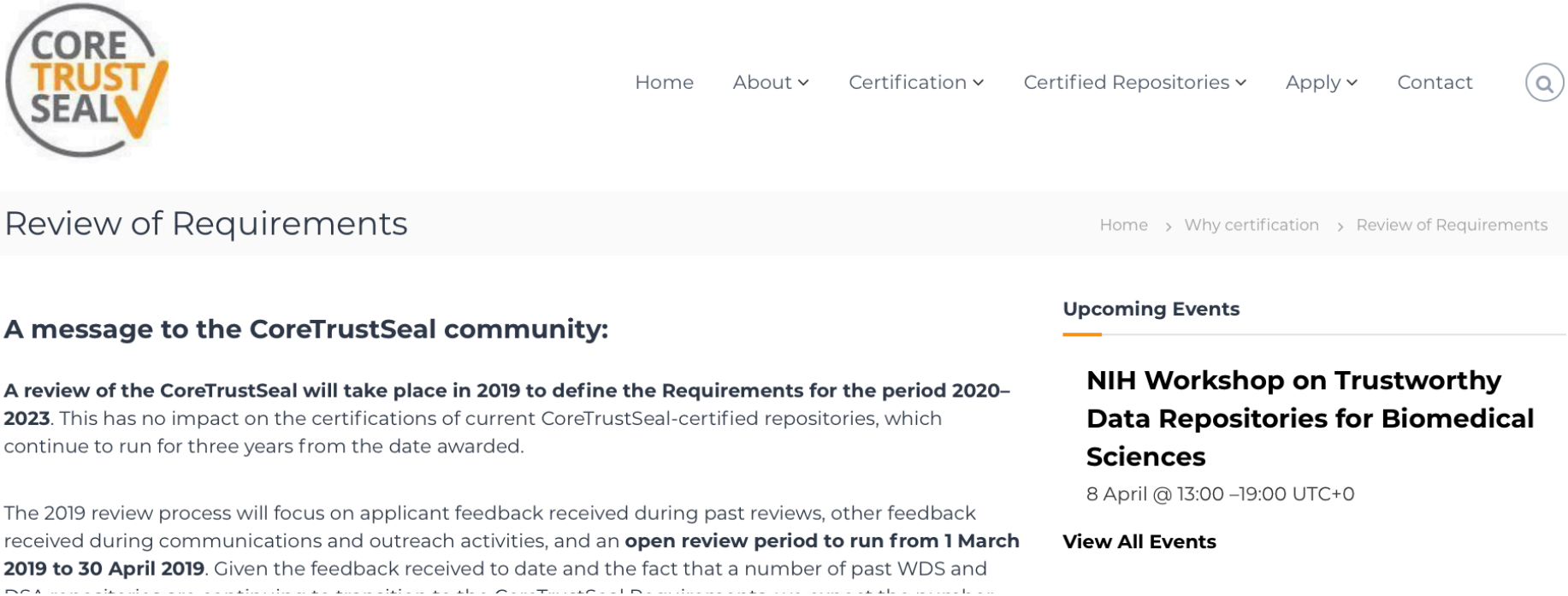
- The rules on European standardisation allow the European Commission to identify ICT technical specifications - that are not national, European or international standards - to be eligible for referencing in public procurement.

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

CORE TRUST SEAL

Review of TDR Requirements

- 3 year cycle of review (2017-2019)



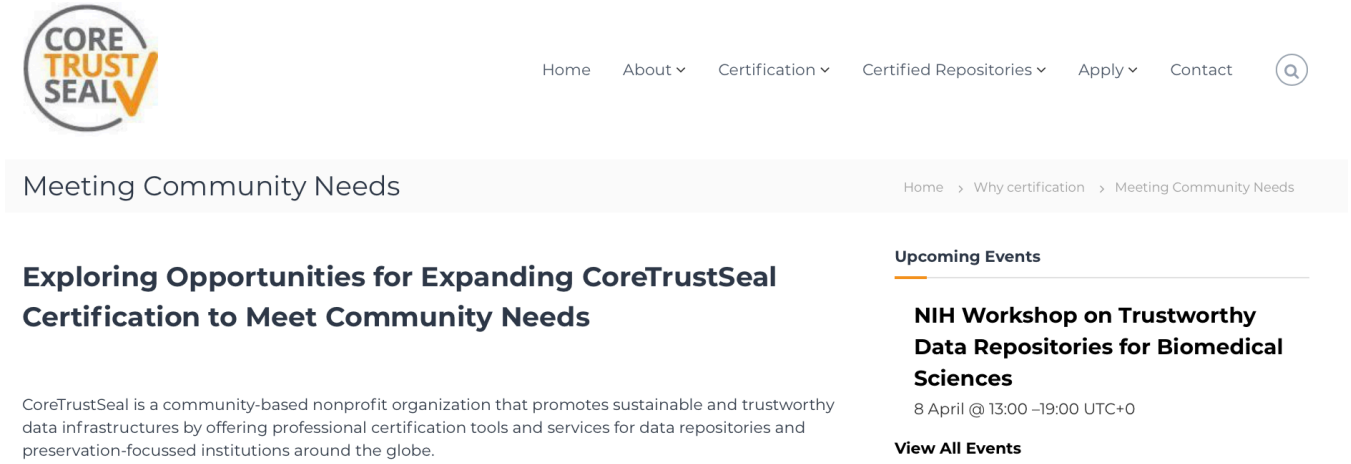
https://www.coretrustseal.org/why-certification/meeting-community-needs/

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/

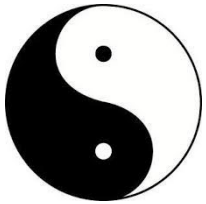


FAIRytale? FAIR and CTS complementarity

"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

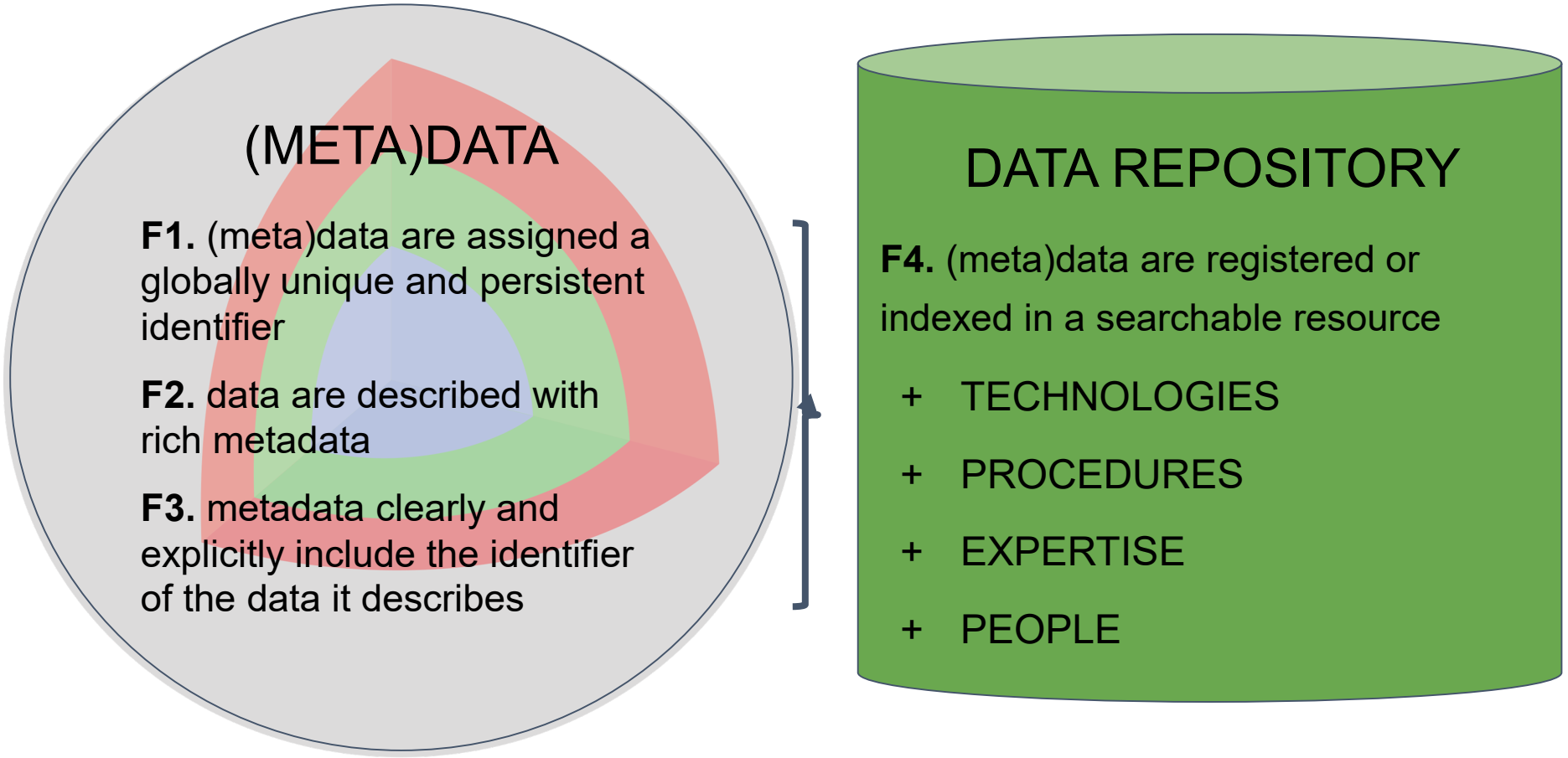
https://tietoarkistoblogi.blogspot.com/2018/11/being-trustworthy-and-fair.html



Two way complementarity:

1. Long-term preservation, accessibility and assessibility of FAIR data
2. Baseline of FAIRness

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.

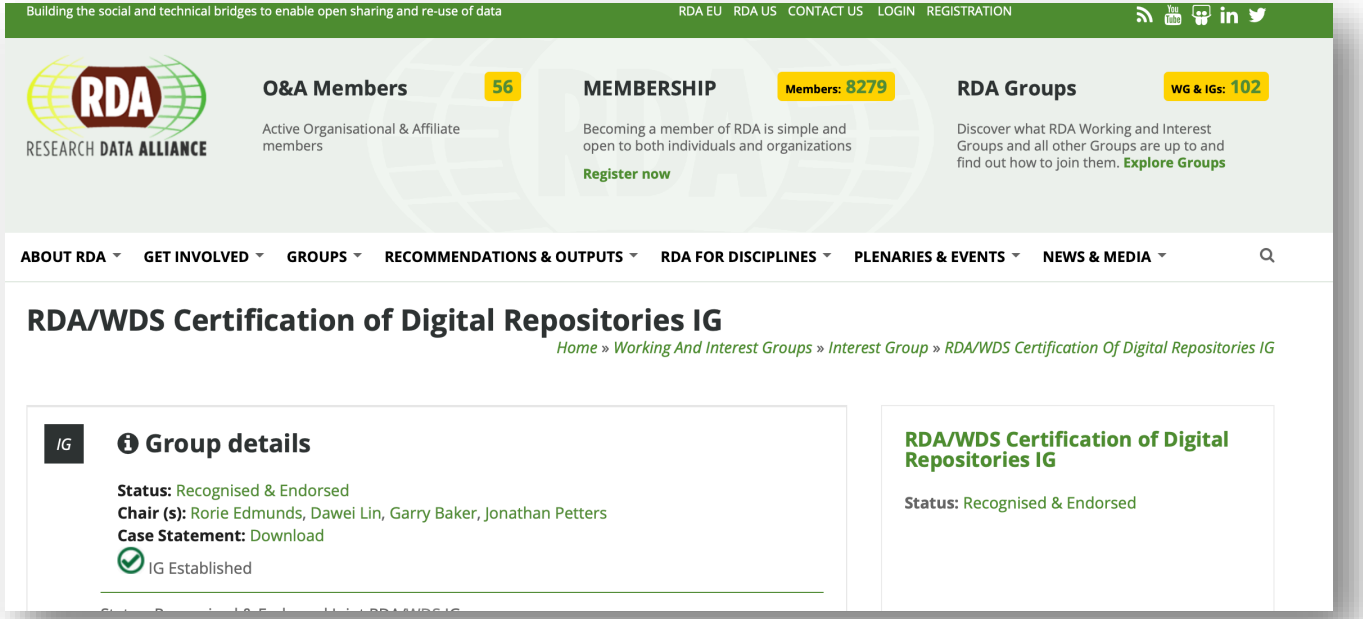
S - Sustainability is the capability to support long-term data preservation and use.

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

CORE TRUST SEAL

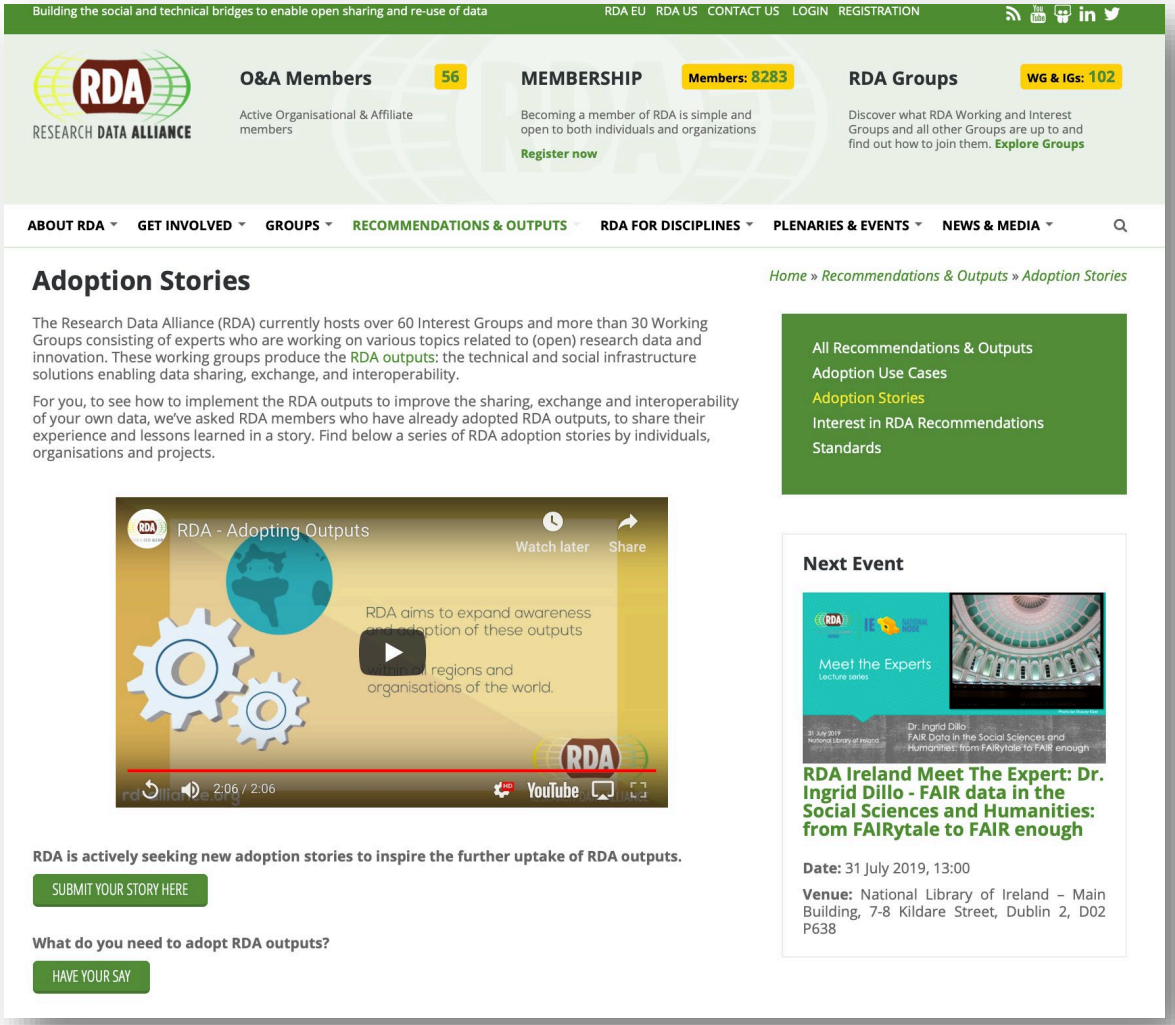
TRUST Principles White Paper

- Version 0.01
- Dawei Lin, Jonathon Crabtree, Ingrid Dillo, Robert R. Downs, Rorie Edmunds, Wim Hugo, and Mustapha Mokrane, ..
- Link: https://bit.ly/2Ih7g8F



CORE TRUST SEAL

Finally: RDA Adoption Stories



CORE TRUST SEAL

Thank you for listening

ingrid.dillo@dans.knaw.nl

www.dans.knaw.nl

www.coretrustseal.org