



iPRES 2022
GLASGOW

18th International Conference
On Digital Preservation

CoreTrustSeal 2023-2025 In a Preservation & Community Context

iPRES2022
Glasgow, Scotland
September 13, 2022

Ingrid Dillo
DANS
Data Archiving & Networked Services
Netherlands

Jonathan Crabtree
UNC Odum
Institute Data Archive
USA

Hervé L'Hours
UK Data Service,
Consortium of Social Science Data
Archives (CESSDA)
United Kingdom



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Overview

- New version of CoreTrustSeal for 2023-25
- CoreTrustSeal in a wider Context including
 - Long Term Digital Preservation Task Force (EOSC)
 - Network of Trustworthy Digital Repositories
- Calls to action to the iPres hive-mind



Certification standards



<https://www.coretrustseal.org/>



<http://www.iso16363.org/>



https://www.langzeitarchivierung.de/Subsites/nestor/EN/Siegel/siegel_node.html



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:

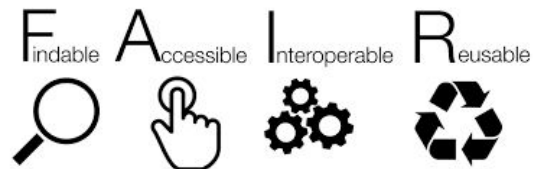
- Gives **data producers** the assurance that their data and associated materials will be stored in a reliable manner and can be reused;
- Provides **funding bodies** with the confidence that data will remain available for reuse;
- Enables **data consumers** to assess the repositories where data are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



Requirements: background

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

☐ Strong link with:



CoreTrustSeal; a brief history

Data Seal of Approval Certification of
Trusted Data
Repositories



WDS Certification of
Regular Members

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



Governance

CoreTrustSeal Board 2021–2024

Directors

1. Chair—**Jonathan Crabtree** (Odum Institute Data Archive, USA)
2. Vice-chair—**Hervé L'Hours** (UK Data Archive, United Kingdom)
3. Secretary—**Mari Kleemola** (Finnish Social Science Data Archive, Finland)
4. Treasurer—**Ingrid Dillo** (Data Archiving and Networked Services, The Netherlands)

Members

4. **Aude Chambodut** (International Service of Geomagnetic Indices, France)
5. **Reyna Jenkyns** (Ocean Networks Canada, Canada)
6. **Dawei Lin** (ImmPort Repository, DAIT-NIAID-NIH, USA)
7. **Jared Lyle** (Inter-university Consortium for Political and Social Research, University of Michigan, USA)
8. **Jonas Recker** (GESIS-Leibniz Institute for the Social Sciences, Germany)
9. **Olivier Rouchon** (Centre Informatique National de l'Enseignement Supérieur, France)

Ex officio

- **Lindsay Callaghan** (South African Environmental Observation Network, South Africa)
- **Lisa de Leeuw** (Data Archiving and Networked Services, The Netherlands)



Governance

Assembly of reviewers:

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



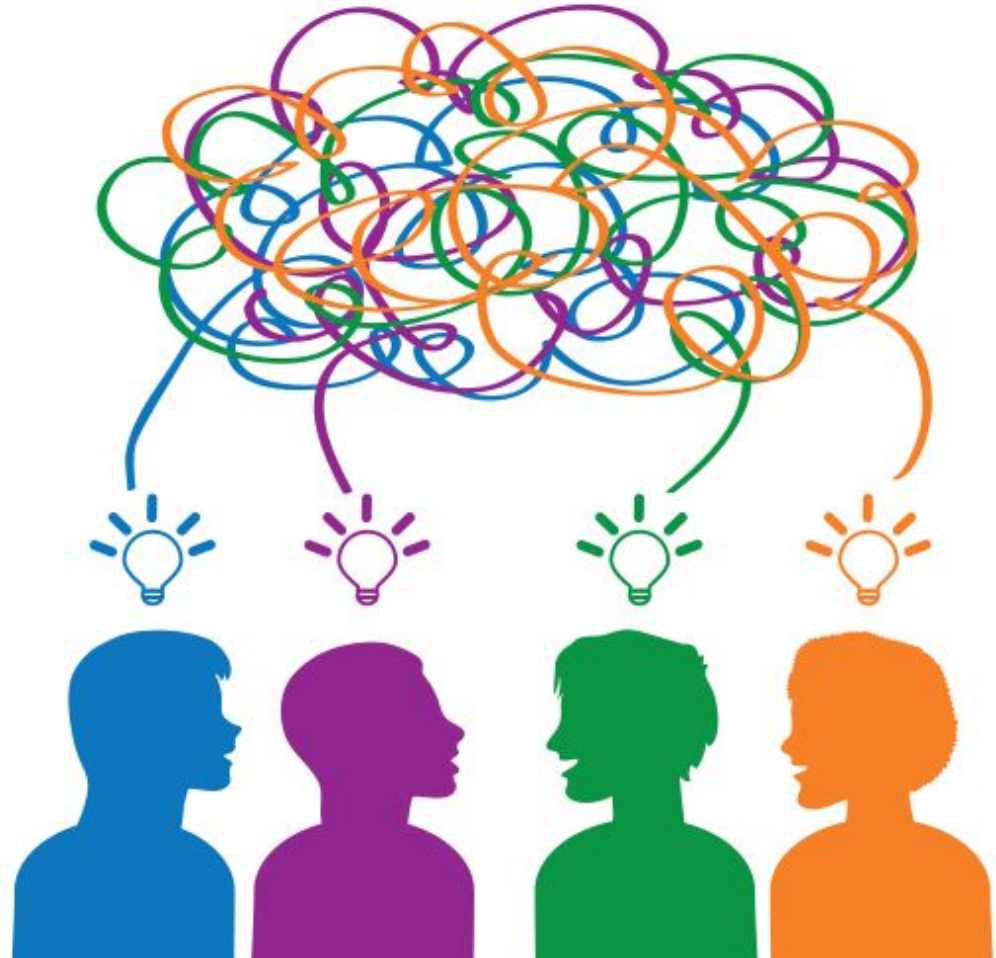
More information on governance in the Rules of

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



CoreTrustSeal initiative

- Not for profit
- Community based
- Strong ties with the Research Data Alliance (RDA)
- Global
- Domain agnostic



Two step certification process

Self assessment based on 16 Requirements (URLs of documented public evidence + compliance level) with extended guidance

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



Three year revision structure to continually meet..

- The rapidly changing nature of data and information management infrastructure.
- The presence of clear community expectations against which assessments can be made.
- The need to deliver a low-barrier to entry and 'core' set of requirements.



Current Requirements 2020-2022

- Active for 2020-2022
- Revisions integrated and improved through an open community working group under the auspices of the RDA Certification of Repositories Interest Group.
- The 2020-22 version of the CoreTrustSeal Requirements sought to support the transition from prior certifications while integrating the experience of initial applicants, reviewers and Board members.



Revision Process for 2023-2025

- Reviewed application over last three years
- Drafted suggested changes and additions
- Surveyed the community
- Reviewed feedback
- Incorporated feedback into draft
- Board reviewed and approved new draft
- Release of final draft
- Final error checking by community and board
- Release of 2023-2025 requirements

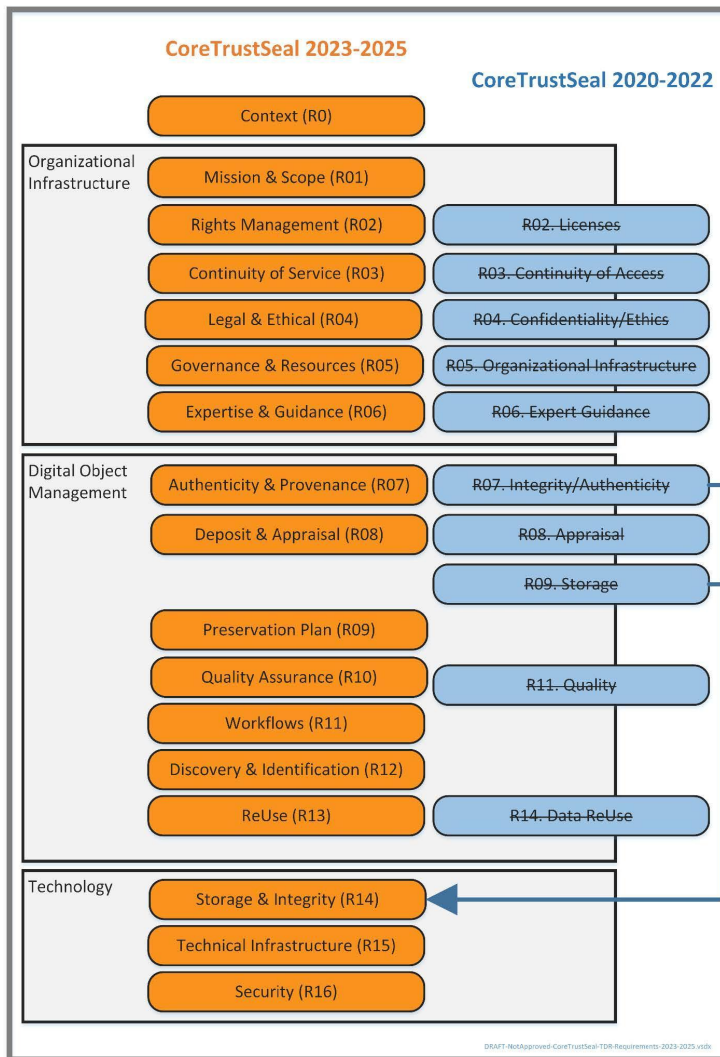


Average responses across all proposed changes

- Suggested changes to the requirements listed in the community survey were overwhelmingly supported by the survey participants with 78% agreeing with proposed changes on average across the survey.



Overview of Changes



Compliance levels

The applicant must indicate a compliance level for each of the Requirements:

- *In Progress: the repository is in the implementation phase.*
- *Implemented: the requirement has been fully implemented by the repository.*

Compliance levels are an indicator of the applicant's self-assessed progress, but reviewers judge compliance against response statements and supporting evidence.



Additional Triage Process

- Expedite Questions of In Scope/Out of Scope
- Identify Active vs Passive preservation differences
- Assist community discussions on levels of preservation with discussion paper
 - <https://doi.org/10.5281/zenodo.6908018>



Clarifications Added

- In cases where the scope was too narrowly focused on 'data', the use of the term 'digital object' has been expanded and used, alongside "data and metadata".

Updates to Repository Background Information

- Context of the repository typology has been replaced by a free text option and a request to select either 'specialist' or 'generalist'. Specialist repositories are asked to clarify their specialist scope.

Typical Changes

- 2023-25 addresses
 - structural changes
 - textual changes
 - updates to improve clarity
 - updates to maintain alignment with the repository and data infrastructure landscape.

Change examples

Continuity of Service (R03)

R03. The Repository has a plan to ensure ongoing access to and preservation of its data and metadata.

Was: 3. Continuity of access. R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.

This change more accurately reflects the scope of the requirement as covering ongoing services offered by the repository including access but also measures to ensure ongoing preservation. Avoids possible confusion with Access in the sense used by the FAIR Principles.

Legal & Ethical (R04)

R04. The repository ensures to the extent possible that data and metadata are created, curated, preserved, accessed and used in compliance with legal and ethical norms.

Was: 4. Confidentiality/Ethics. R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.

This change highlights that many data protection measures are legally as well as ethically governed. This was already covered in the guidance text but is made more explicit. There is a stronger focus on evidence that demonstrates the applicants understanding of the legal and ethical framework they work within. References to 'discipline' have been adjusted to support a wider range of applicants. There is clearer separation of general guidance from that related to digital objects with a disclosure risk.

.



Next Steps and Timeline

- October 31st pause submissions to transition over
- November- December work on existing submissions
- Early 2023 begin accepting application under new requirements

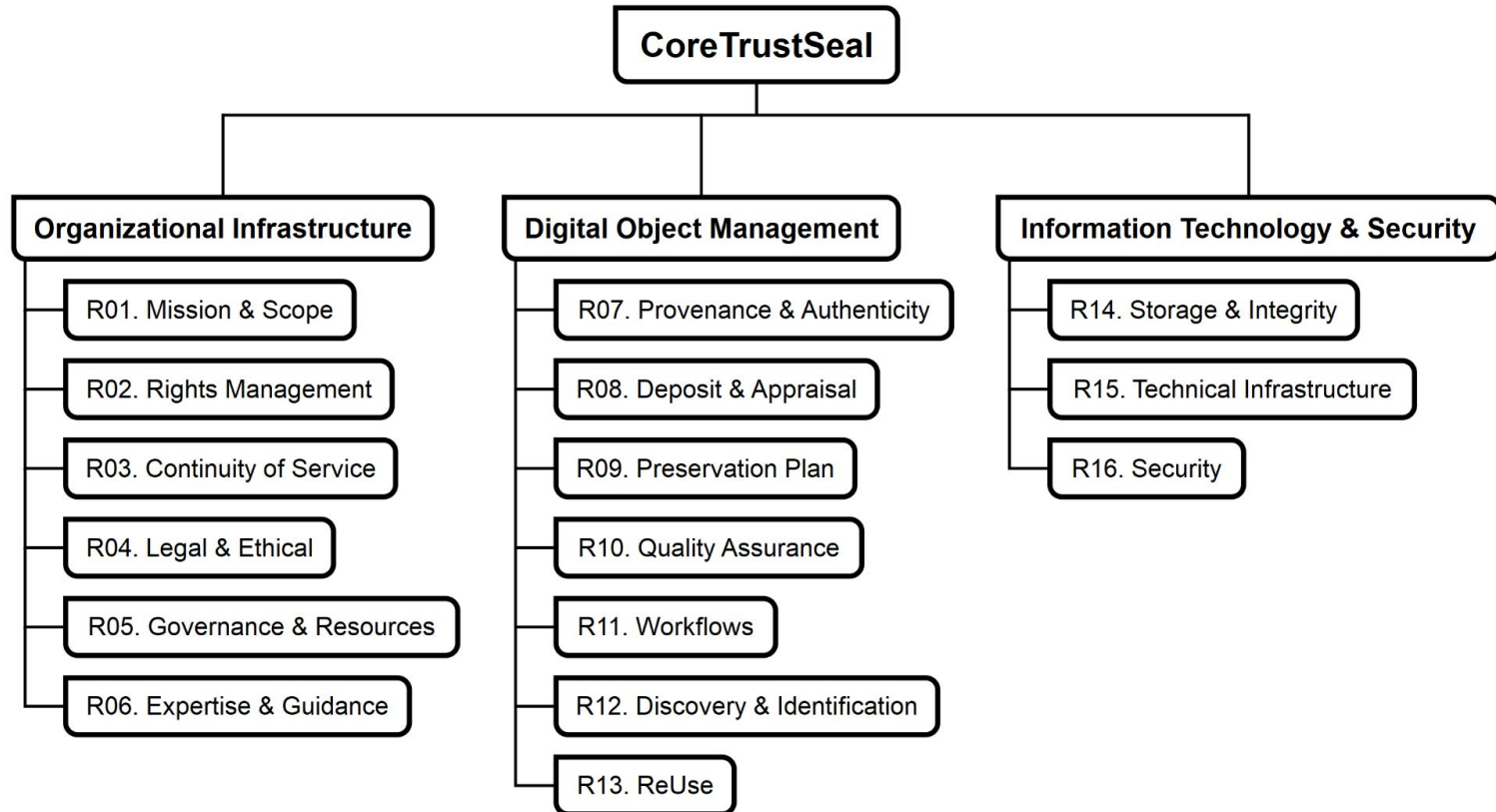


References

- **CoreTrustSeal Requirements v03.00
2023-2025** (<https://doi.org/10.5281/zenodo.7051011>)
- **CoreTrustSeal Extended Guidance v03.00
2023-2025** (<https://doi.org/10.5281/zenodo.7051095>)
- **CoreTrustSeal Glossary v03.00
2023-2025** (<https://doi.org/10.5281/zenodo.7051124>)
- <https://www.coretrustseal.org>



CoreTrustSeal in Context



CoreTrustSeal in Context

- Community Consensus on Curation & Preservation Levels
- Supporting Generalist *and* Specialist Repositories
- Data Services (non-TDR) - interactions and outsourcing
- Human-mediated vs Machine actionable:
 - repositories, objects and assessments
- Elaborating around the 'core' of CoreTrustSeal:
 - CoreTrustSeal+FAIR and more

Why is this still so hard?



CoreTrustSeal in Context: Why is This Still So Hard???

So.. why is this still so hard?

Consensus > Implementation > Assessment > Outcomes (not just certification!)

- Since OAIS and TDR became a thing, storage has moved from a challenge for every repository to more of an off the shelf commodity (sort of)
- But we still have no community agreed minima of how many copies, on how many media, in how many locations, with what back up frequency and integrity measures.
- If we can't define the minimum acceptable criteria for storage, then clearly the TDR and trust in other Data Services is an evolving challenge to define and then journey towards.
- A lack of agreed minimal goals (standards) and a means to evaluate them (assessment) weakens the foundations we have to build on

But nobody likes to be judged...

But the attention repositories and repository standards are receiving is also an opportunity.



CoreTrustSeal in Context: EOSC Association

EOSC: European Open Science Cloud

Representative of a global trend towards: consolidation, defragmentation, centralised services and networks of partnership and outsourced services.

Focus on:

- FAIR (Findable Accessible, **Interoperable** and ReUsable) digital objects
- Trusted Repositories

A 'journey': where we are, where we need to be, and how to get there.

Preservation an identified lack in EOSC development

cf: the DPC-led FAIR Forever Report. <https://doi.org/10.5281/zenodo.4574234>

Strategic Research and Innovation Agenda (SRIA)

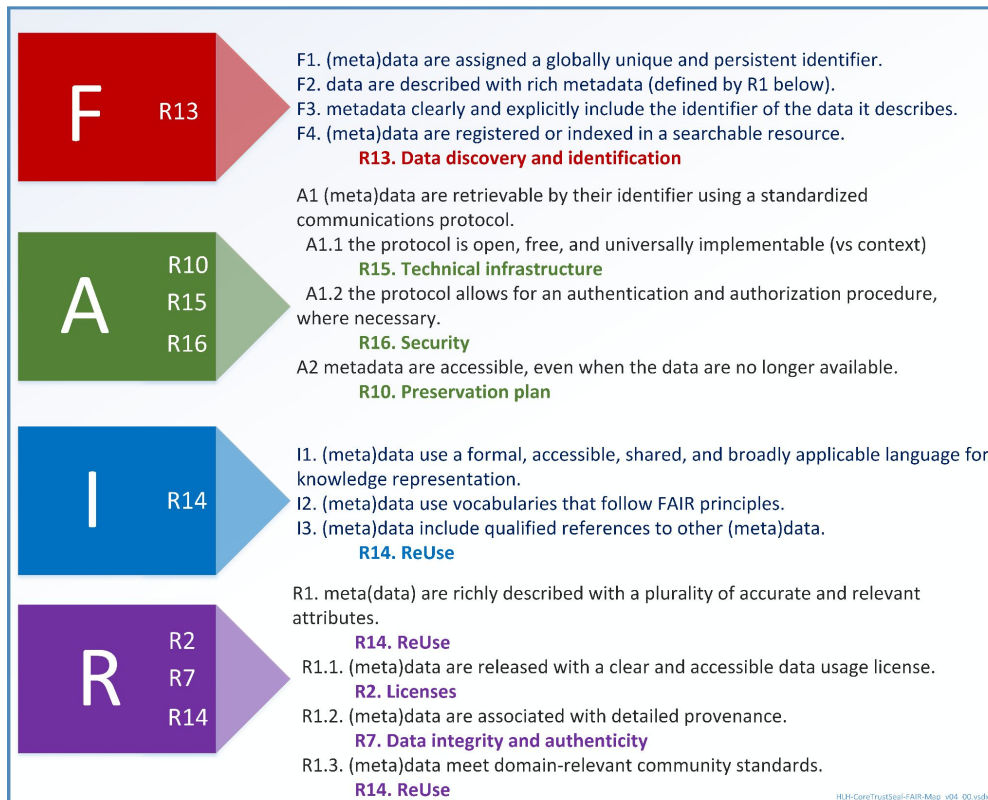
<https://data.europa.eu/doi/10.2777/935288>



CoreTrustSeal in Context: CoreTrustSeal + FAIR

CoreTrustSeal 2020-2022
Mapped to FAIR

FAIR + Time
Requires
Preservation



CoreTrustSeal in Context: Evolving FAIR

4 Letters
F.A.I.R

15
Principles

40+
Indicators

??
Metrics

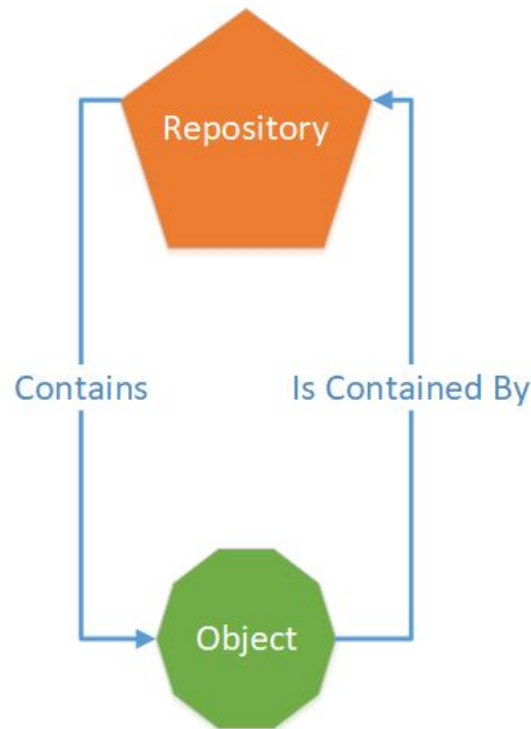
??
Tests

HLH-FAIR-AcronymPrinciplesIndicatorsMetricsTests_02_00.vsd



CoreTrustSeal in FAIR-enabling Context

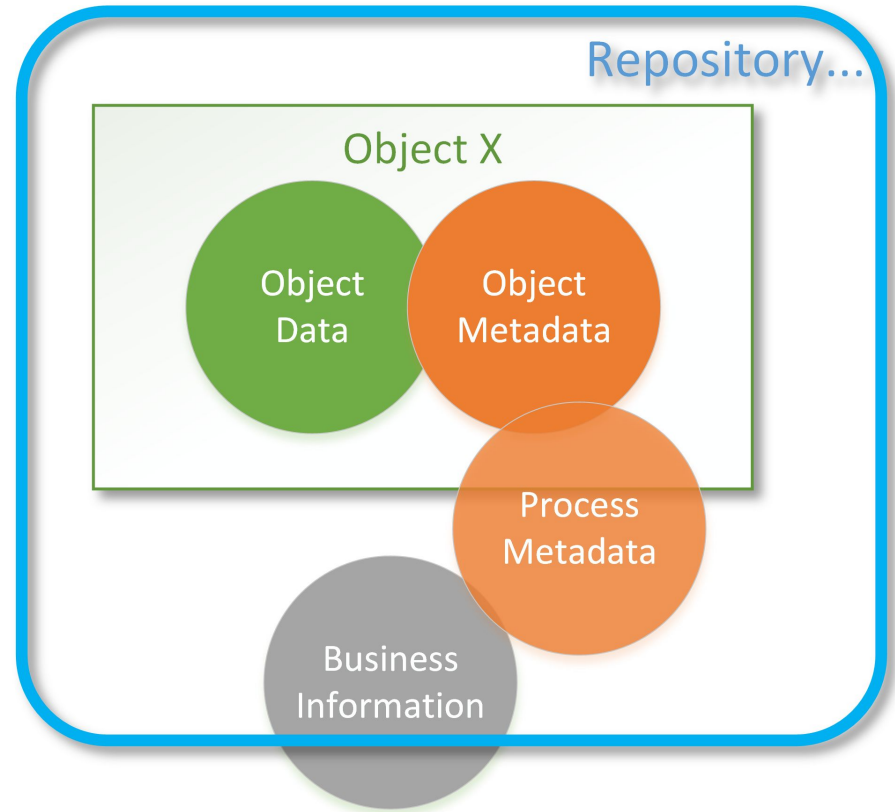
Assessing Repositories & Objects



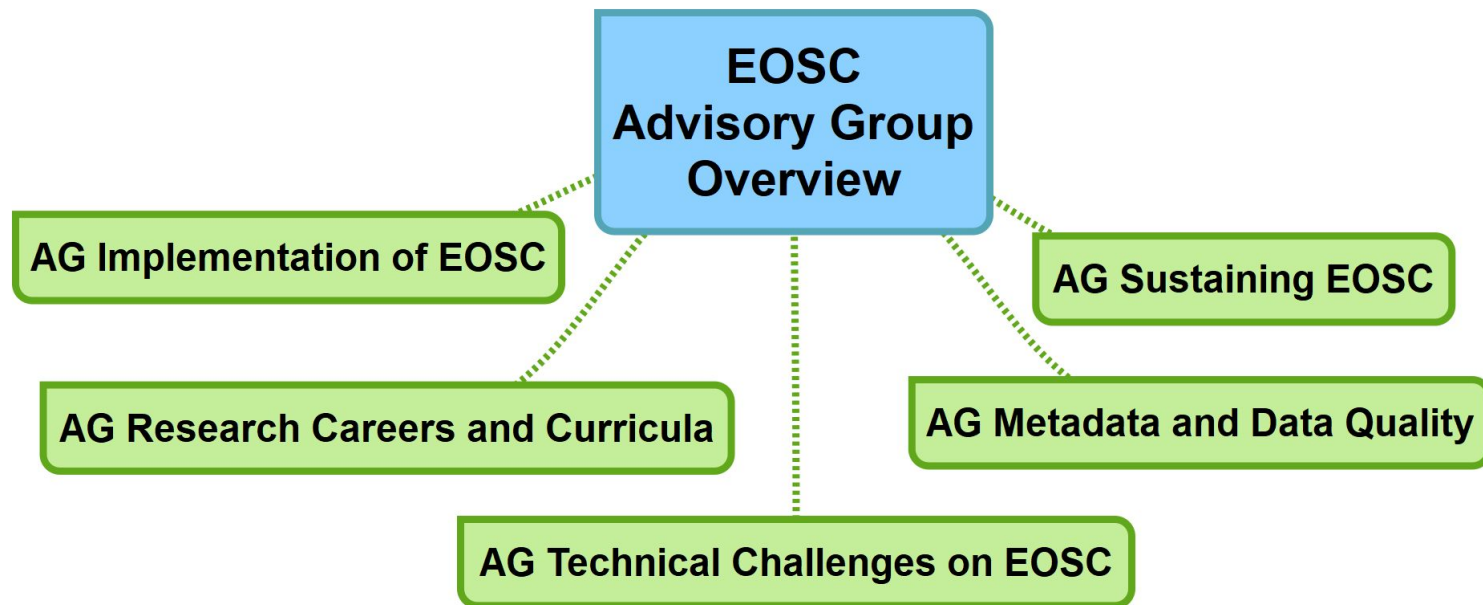
Repository ObjectMetadata A_00_01.vsd

CoreTrustSeal in Context

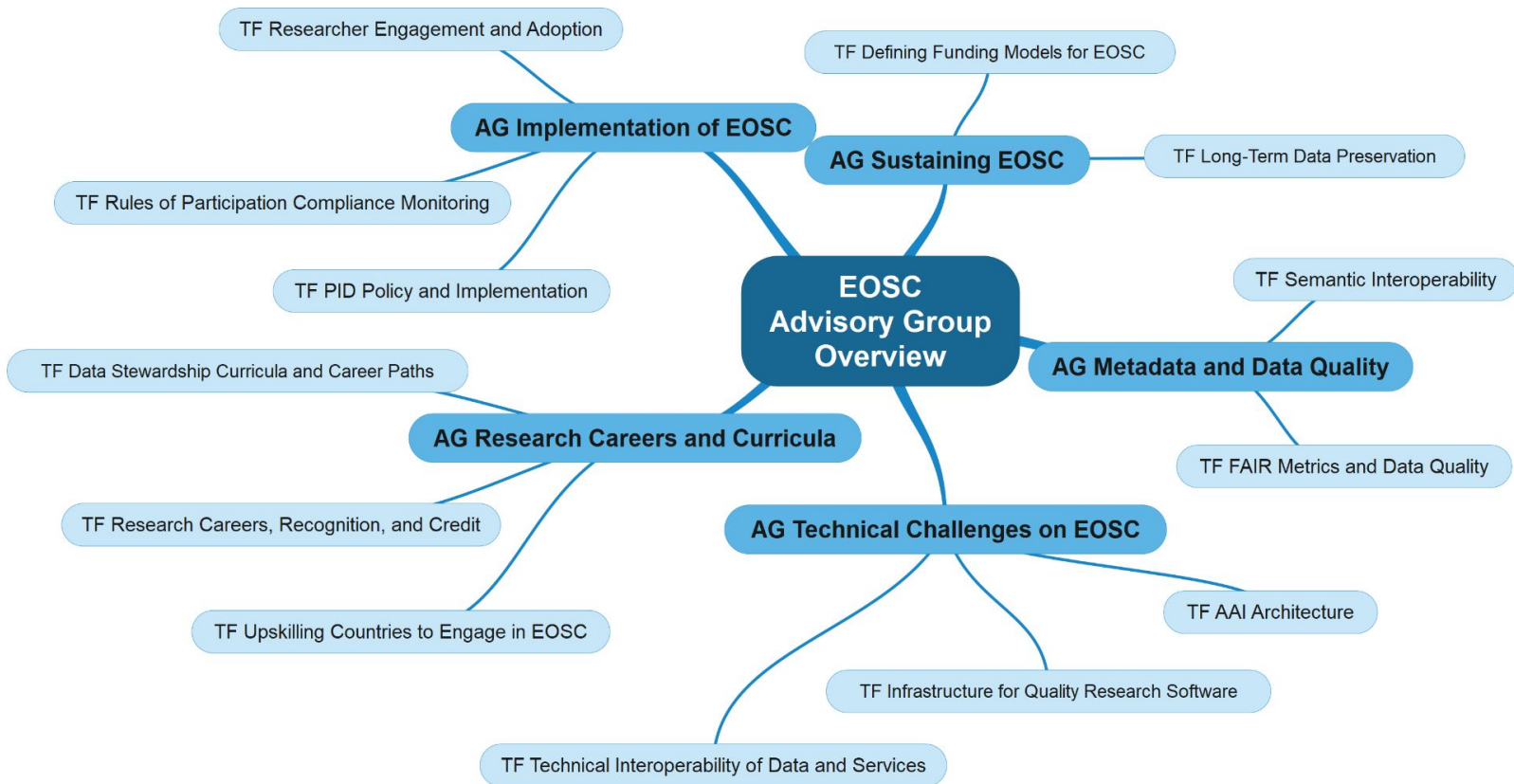
Artefacts that support practice
provide evidence
to demonstrate compliance



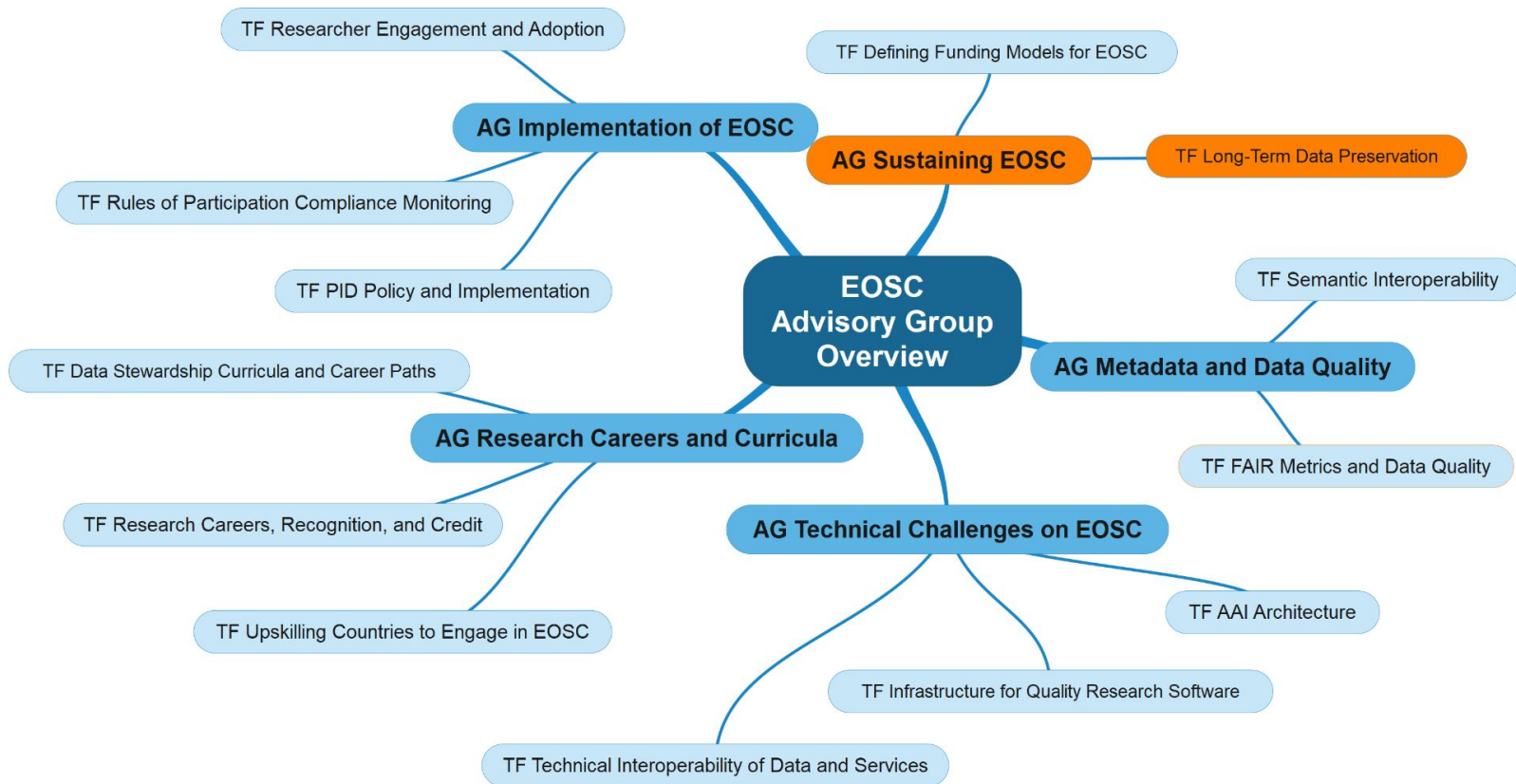
CoreTrustSeal in Context: EOSC A Advisory Groups



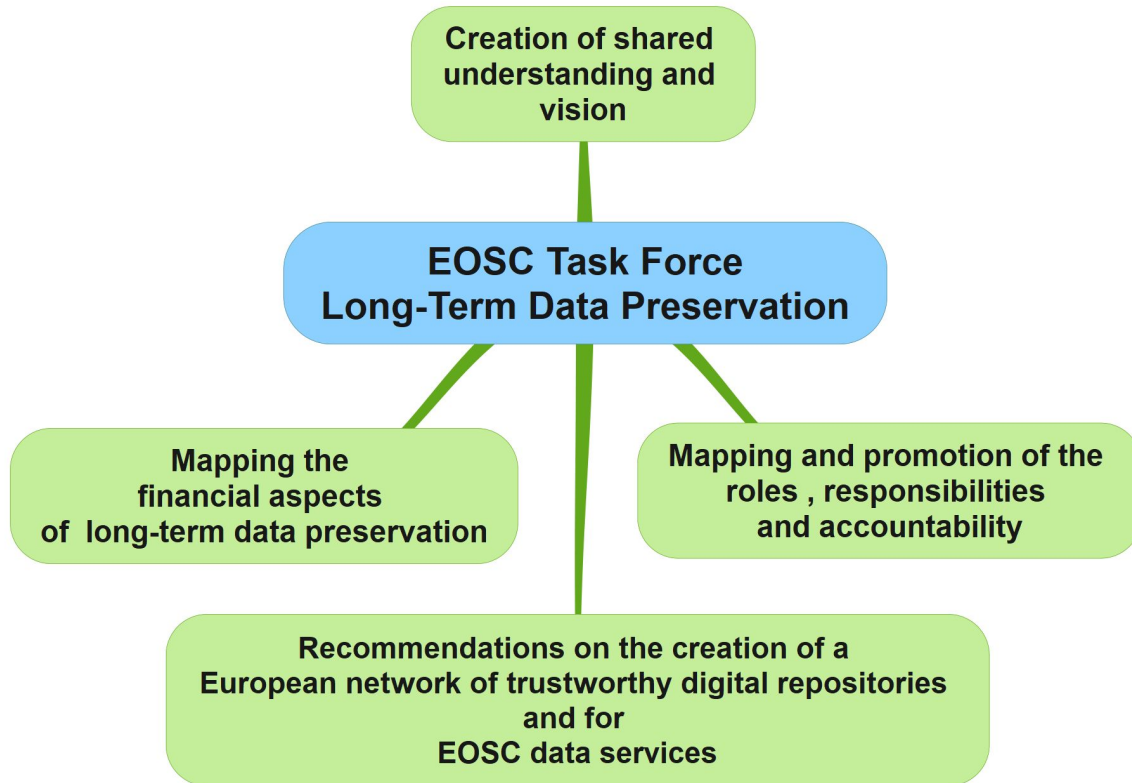
CoreTrustSeal in Context: EOSC Task Forces



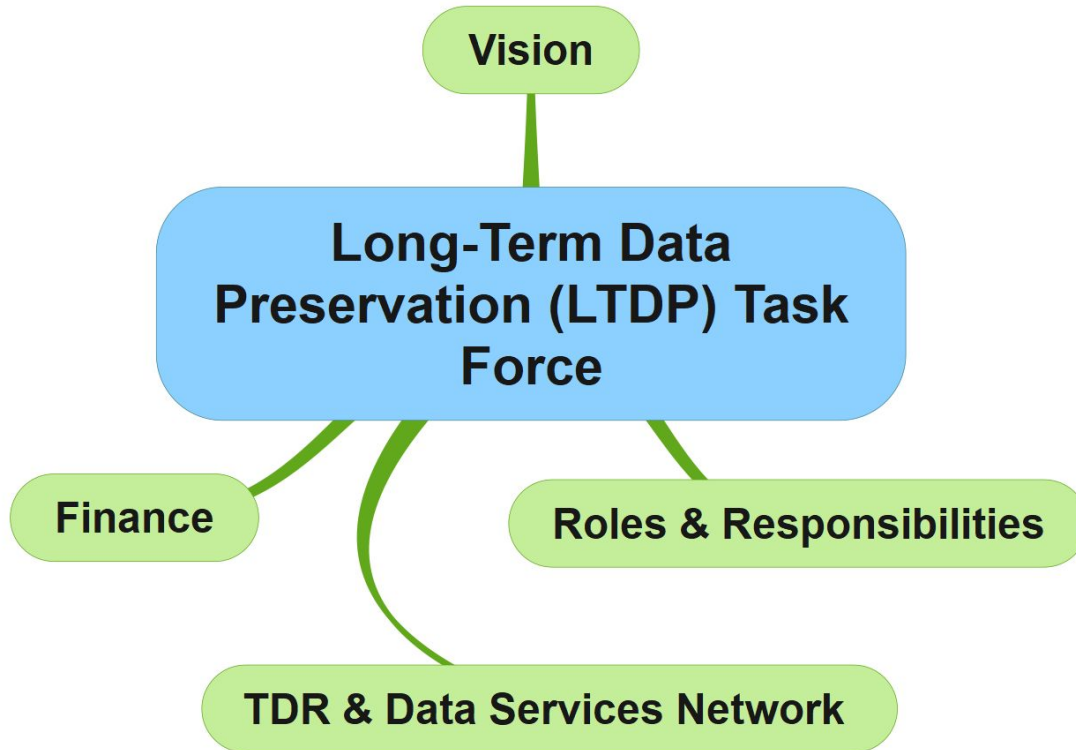
CoreTrustSeal in Context: EOSC Task Forces



CoreTrustSeal in Context: LTDP TF



CoreTrustSeal in Context: LTDP TF



CoreTrustSeal in Context: LTDP TF

Vision for the Future

Overview and discussion paper for broad feedback

Iterative consultation on recommendations:

- European,
- National,
- Institutional Level.



CoreTrustSeal in Context: LTDP TF

Finance & Roles

Can these be aligned with a common set of?

- cost centres?
- (responsibilities for) actions?

Definitions: archive, archiving, repository, preservation, curation etc

Insufficient, internally facing (**we** know what **we** mean)-in group/out group



CoreTrustSeal in Context: LTDP TF

Finance & Roles

Can these be aligned with a common set of

- cost centres
- (responsibilities for) actions

Beyond storage and (snapshot) curation to e.g. FAIRness

**What are the *unique* costs and responsibilities of
'preservation'.**



CoreTrustSeal in Context: Calls to Action

CODATA Research Data Management Terminology
(2022 edition) Deadline end of September

https://docs.google.com/document/d/1d7q4XcFkOFdwLfpdTFT9AfxiTiYRwyvqrY5-g_YdzuY/edit#heading=h.awik3akmznw

CoreTrustSeal Curation & Preservation Levels
Discussion Paper

<https://doi.org/10.5281/zenodo.6908019>

Review of CoreTrustSeal Applicability to
non-Preservation (Trustworthy) Data Services

<https://doi.org/10.5281/zenodo.6865991>



Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDR)

Ingrid Dillo DANS Data Archiving & Network Services

- FAIRsFAIR <https://www.fairsfair.eu/>
- SSHOC <https://sshopencloud.eu/>
- EOSCNordic <https://www.eosc-nordic.eu/>
- Workshop <https://doi.org/10.5281/zenodo.5849658>

Working paper for public feedback

<https://doi.org/10.5281/zenodo.7034314>



Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDR)

Philipp Konzett | The Arctic University of Norway (UiT)

Ingrid Dillo | Data Archiving and Networked Services (DANS)

Francoise Genova | Strasbourg astronomical Data Center (CDS)

Natalie Harrower | Digital Repository of Ireland (DRI)

Vasso Kalaitzi | Data Archiving and Networked Services (DANS)

Mari Kleemola | Finnish Social Science Data Archive (FSD)

Amela Kurta | Data Archive for Social Sciences in Bosnia and Herzegovina (DASS-BIH)

Pedro Principe | University of Minho (UMinho)

Olivier Rouchon | National Computing Center for Higher Education (CINES)

Hannes Thiemann | World Data Centre for Climate (WDCC)

Maaïke Verburg | Data Archiving and Networked Services (DANS)



Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDR)

- Vision and feedback to identify community needs
- Framework for next steps of development

Range of maturity and practices, certified and aspiring TDRs.

“not one size fits all”



Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDR)

- Beyond limited project timeframes and perspectives
- Empower trustworthy and FAIR-enabling repositories in the EOSC ecosystem.
- A 'voice' for European TDRs within EOSC
- Community led, knowledge exchange, peer to peer support, network spaces
- Continuous validation of standards and practices



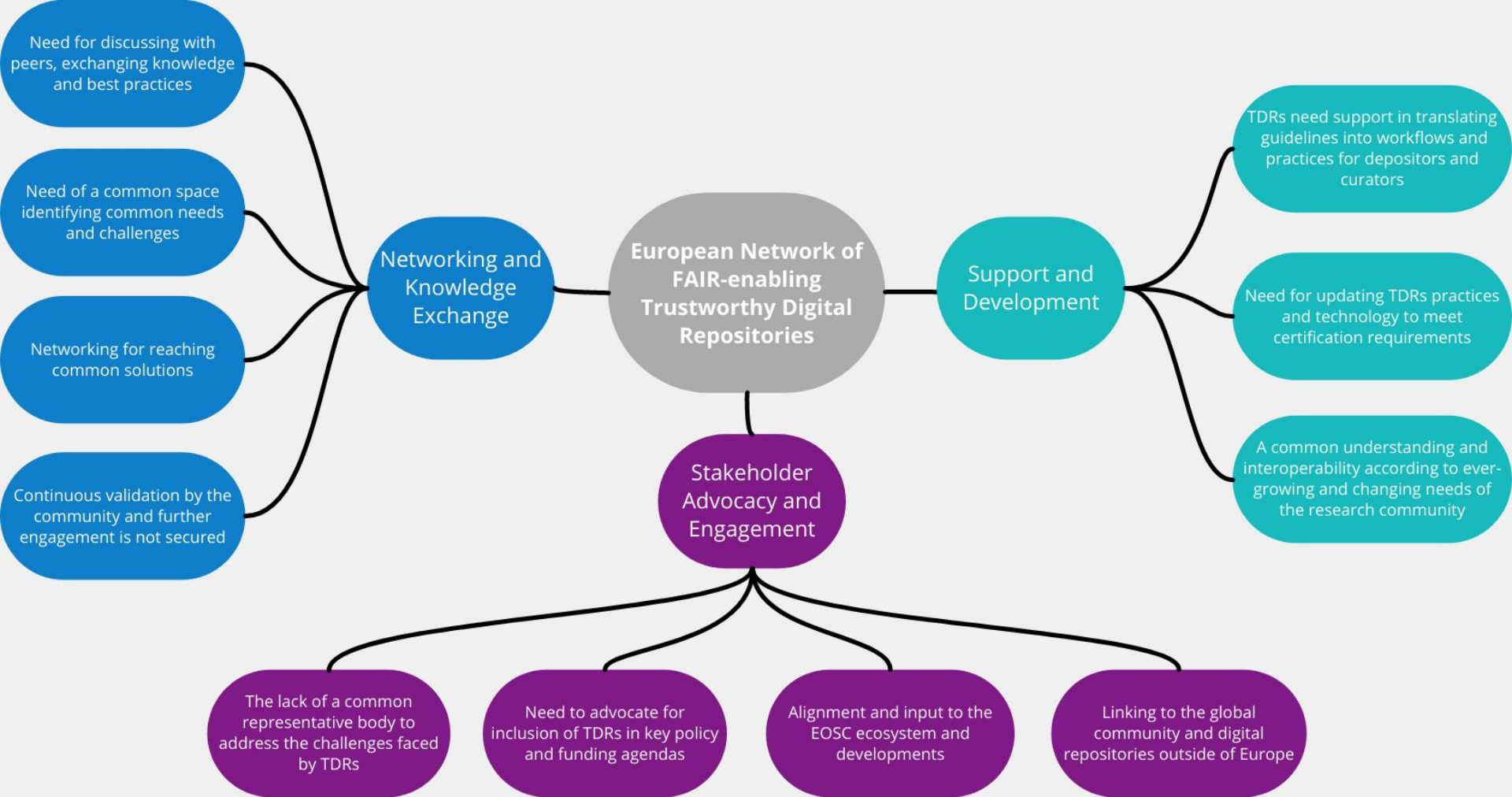
Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDR)

Repositories at the Core of the Network: domain or discipline-specific, to institutional, to regional/national, to generic.

Engage with national and international initiatives, wide range of stakeholders, such as researchers, funders, the EOSC, the larger community of research data management experts.

Addressing: principles (FAIR, TRUST, CARE) common frameworks for best practices, licensing, standardisation, metadata, controlled vocabularies, legal and ethical compliance, interoperability





Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDR)

- Networking & Knowledge Exchange
- Support & Development
- Stakeholder Advocacy & Engagement



Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDR)

Framework- Next Steps:

Membership

- Repositories and other initiatives that contribute to increasing the number of FAIR-enabling TDRs
- Benefits, conditions, structure, application process, business model

Governance

Key characteristics, expertise, existing bodies and strategies, roles



CoreTrustSeal in Context: Calls to Action

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

CoreTrustSeal: always open for feedback: info@coretrustseal.org

CoreTrustSeal Curation & Preservation Levels Discussion Paper

<https://doi.org/10.5281/zenodo.6908019>

Review of CoreTrustSeal Applicability to non-Preservation (Trustworthy) Data Services

<https://doi.org/10.5281/zenodo.6865991>

Network of Trustworthy Digital Repositories

Draft paper open for feedback <https://doi.org/10.5281/zenodo.7034314>

Long Term Digital Preservation Task Force (EOSC)

Overview and Recommendations forthcoming for public feedback

CODATA Research Data Management Terminology (2022 edition) Deadline end of September

https://docs.google.com/document/d/1d7q4XcFkOFdwLfpdTFT9AfxiTiYRwyvqrY5-g_YdzuY/edit#heading=h.awik3akmznw



Thank you for listening



Email:

Website: <https://www.coretrustseal.org/>

Questions?

