CESSDA Trust Landscape

Webinar

November 10, 2022

cessda.eu   @CESSDA_Data

Program

15:00-15:15 Welcome and introduction
Trond Kvamme, Sikt

15:15-15:30 News from EOSC TF on Long-Term Data Preservation
Hervé L'Hours, UKDS, Task Force Co-chair

15:30-15:45 New CoreTrustSeal requirements
Mari Kleemola, FSD, CoreTrustSeal Board member

15:45-16:00 Discussion

cessda.eu

@CESSDA_Data

Background

Activity report on progress in Trust activities across the SSH domain

cessda

CESSDA ERIC Agenda 21-24, Tasks 21-22
Trust Pillar: Task 1 Trust Activities

D15 Activity report on progress in Trust activities across the SSH domain

Document info

Dissemination Level	PU
Due Date of Deliverable	31/10/2022
Actual Submission Date	31/10/2022
Type	Report
Approval Status	Approved by Working Group Leader Hervé L'Hours
Version	V1.0
Number of Pages	p.1 – p.29
DOI	10.5281/zenodo.5554445

www.cessda.eu

https://doi.org/10.5281/zenodo.5554445

Certification and FAIRness of repositories

- Revisions in the CoreTrustSeal requirements

- CoreTrustSeal curation & preservation levels discussion paper
https://doi.org/10.5281/zenodo.6908019

  Goal: distinguish between organisations offering *active preservation* of digital objects from *non-preservation* data and metadata services.

  Share comments / feedback: https://www.coretrustseal.org/contact/

cessda

Certification and FAIRness of repositories

Key outputs from the FAIRsFAIR project:

- FAIR-enabling Data Policy Checklist
- ACME-FAIR guide
- F-UJI tool
- FAIR-Aware online tool
- FAIRsFAIR Data Object Assessment Metrics
- CoreTrustSeal+FAIRenabling Capability Maturity Model
- Repository Discovery in DataCite Commons

cessda

Certification and FAIRness of repositories

Certification support

- SSHOC D8.3, "Trustworthy Digital Repository status update and certification solutions for SSHOC repositories"

- EOSC-Nordic D4.5 "Report on completed FAIR data standard adoption and certifications of data repositories in the region".

cessda

Other projects and initiatives

- FAIR-IMPACT (Expanding FAIR solutions across EOSC)
- IASSIST 2022
- Network of FAIR-enabling Trustworthy Digital Repositories
- Research Data Alliance
- European Research Data Landscape study
- EOSC Task Force on Long Term Data Preservation

cessda

Conclusions and recommendations

Main takeaway:

Trust and FAIR assessment continue to be useful in helping to clarify the added value of *trustworthy digital repositories offering preservation services*

in contrast to

more technically-driven repository systems that cannot guarantee the accessibility and usability of data in the long term.

cessda

Conclusions and recommendations

1	Closely follow the discussions accompanying the building of EOSC. Keep updated on work and outputs coming out of RDA. This will enable CESSDA to react adequately to relevant Trust-related developments within these networks.
2	Follow up on relevant outcomes from projects like SSHOC, FAIRsFAIR, FAIR-Impact and EOSC Nordic and integrate where appropriate to enable CESSDA to take advantage of synergies from common SSH practices.
3	Aim at playing a key role in providing certification and FAIR support service for SSH communities and also more widely.
4	Support work to identify different types of repositories and efforts to design selection/recommendation systems (like re3data.org or FAIRsharing).
5	Participate in shaping the (SSH) standards used in automated FAIR assessment tools.
6	Explore the development of routines and policies enabling the assignment of PIDs not only on dataset or study level, but also to authors, contributors and funders, and to parts of studies.
7	Endorse the TRUST principles.
8	Emphasise the need for domain/subject-based curation and deposition of data with a discipline specific TDR.
9	Include more machine-understandable metadata in the catalogues.

cessda

Program

15:00-15:15 Welcome and introduction
Trond Kvamme, Sikt

15:15-15:30 News from EOSC TF on Long-Term Data Preservation
Hervé L'Hours, UKDS, Task Force Co-chair

15:30-15:45 New CoreTrustSeal requirements
Mari Kleemola, FSD, CoreTrustSeal Board member

15:45-16:00 Discussion

cessda.eu

@CESSDA_Data

LTDP Task Force Overview

CESSDA Landscape Event

Hervé L'Hours & Roxanne Wyns: LTDP TF Co-Chairs

2022-11-09 Hervé L'Hours
Task Force Co-Chair,
UK Data Service,
Consortium of Social Science Data Archives (CESSDA)

Long Term Data Preservation Task Force

EOSC, FAIR and Preservation

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.

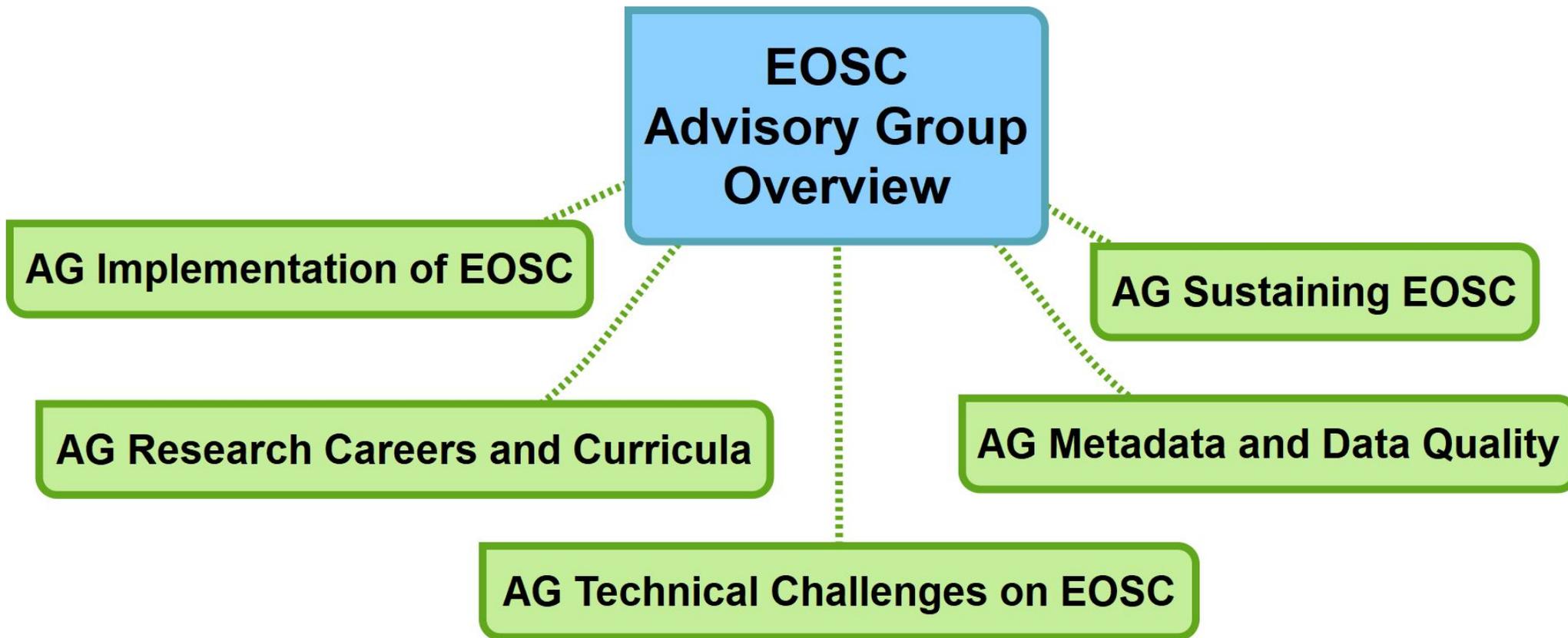Preservation an identified lack in EOSC development

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
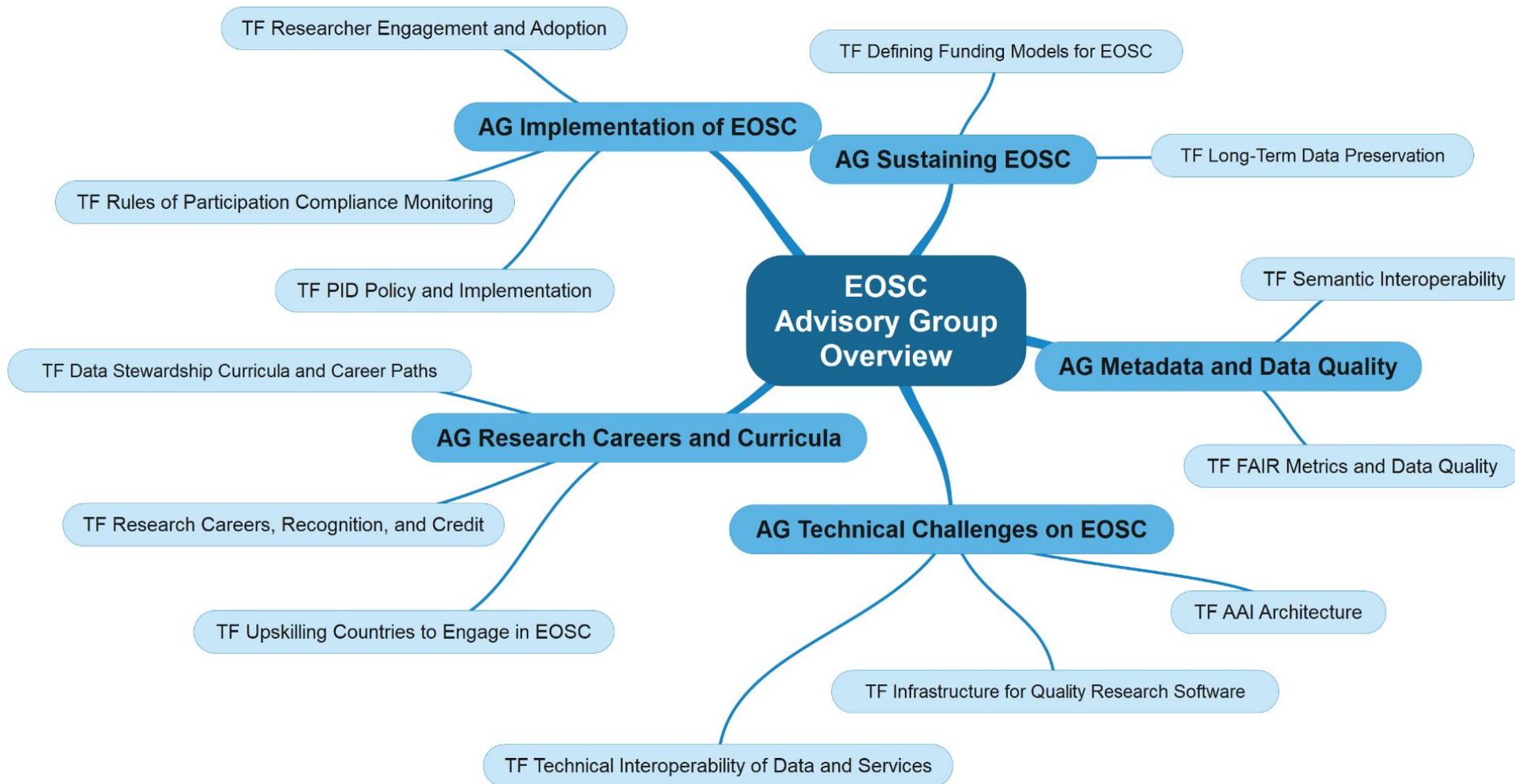
EOSC Task Forces

Long Term Data Preservation Task Force

Sub-tasks

Creation of shared understanding and vision

EOSC Task Force
Long-Term Data Preservation

Mapping the
financial aspects
of long-term data preservation

Mapping and promotion of the roles , responsibilities and accountability

Recommendations on the creation of a
European network of trustworthy digital repositories
and for
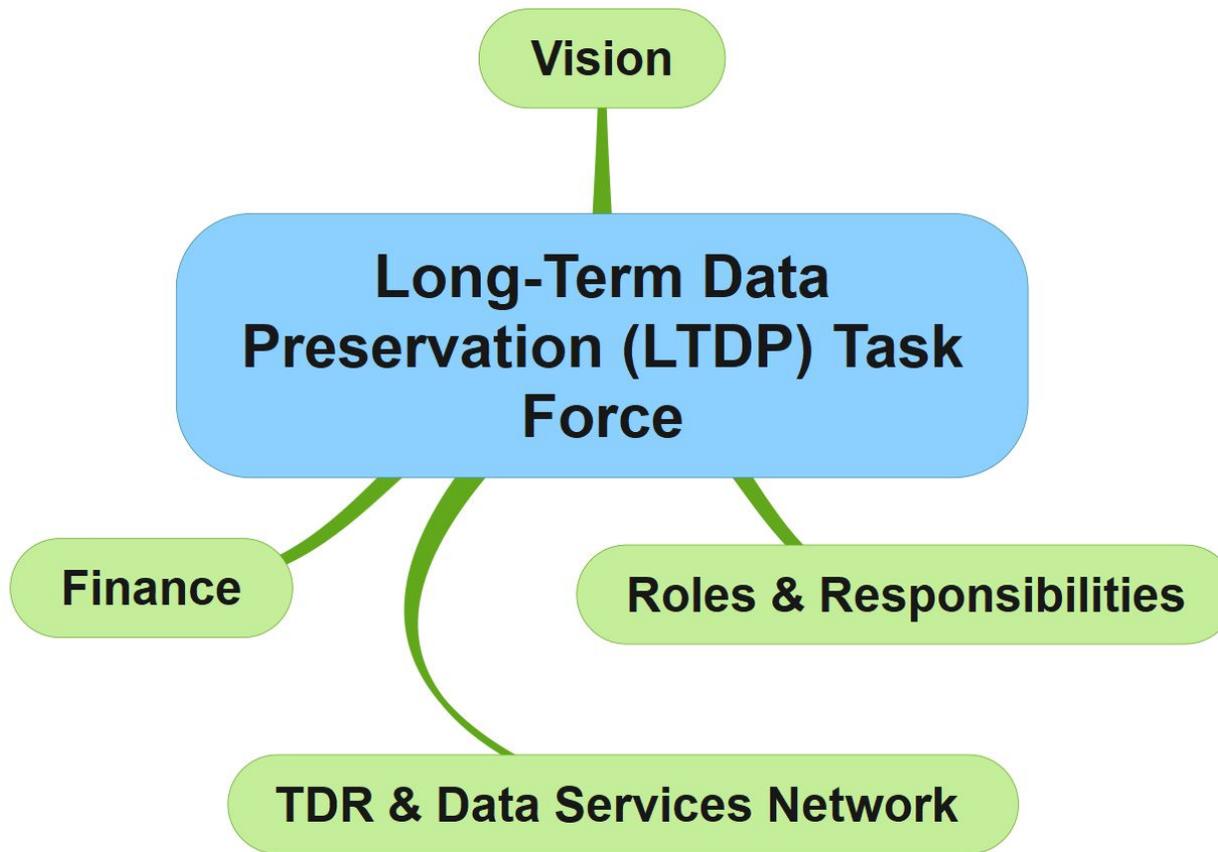EOSC data services
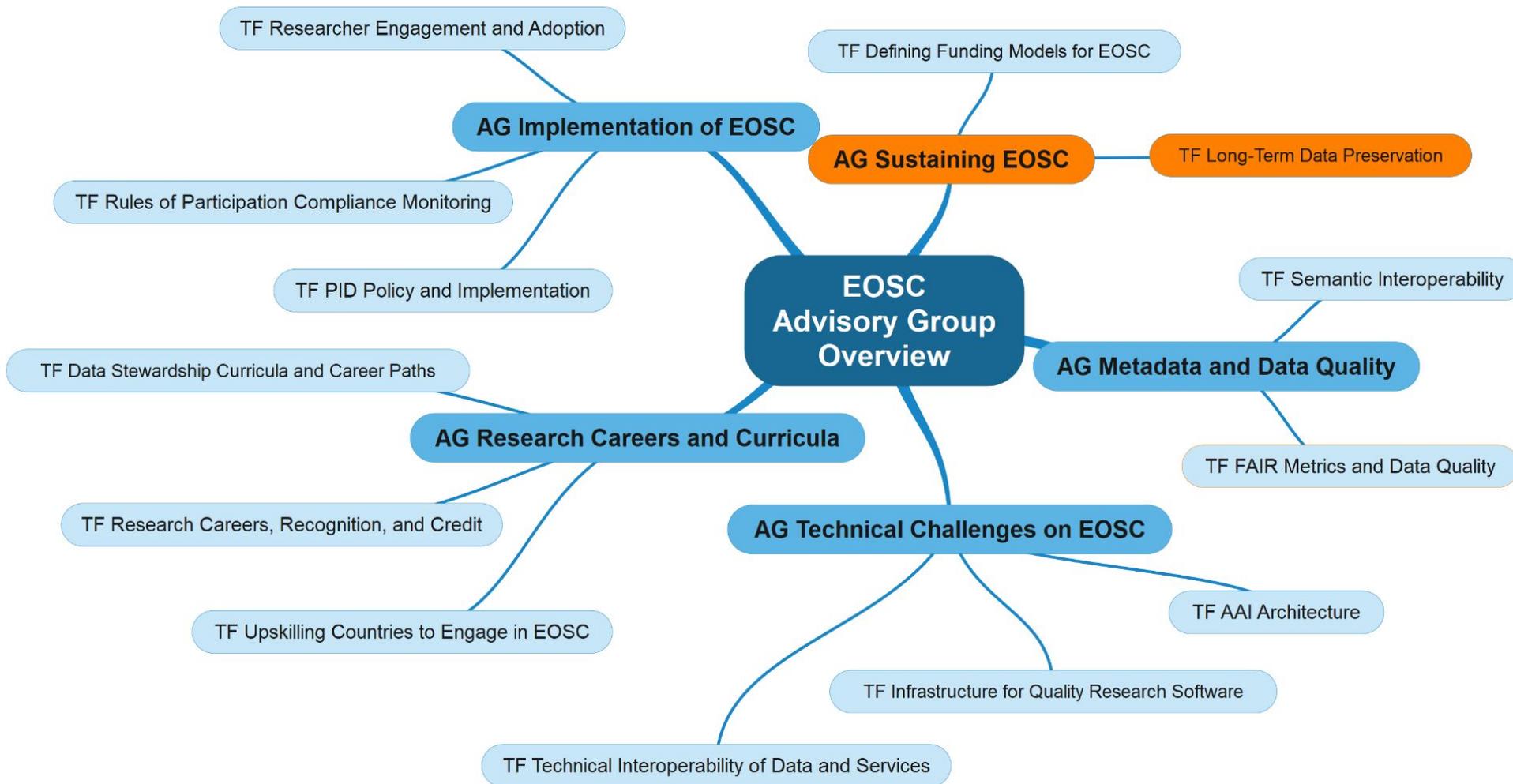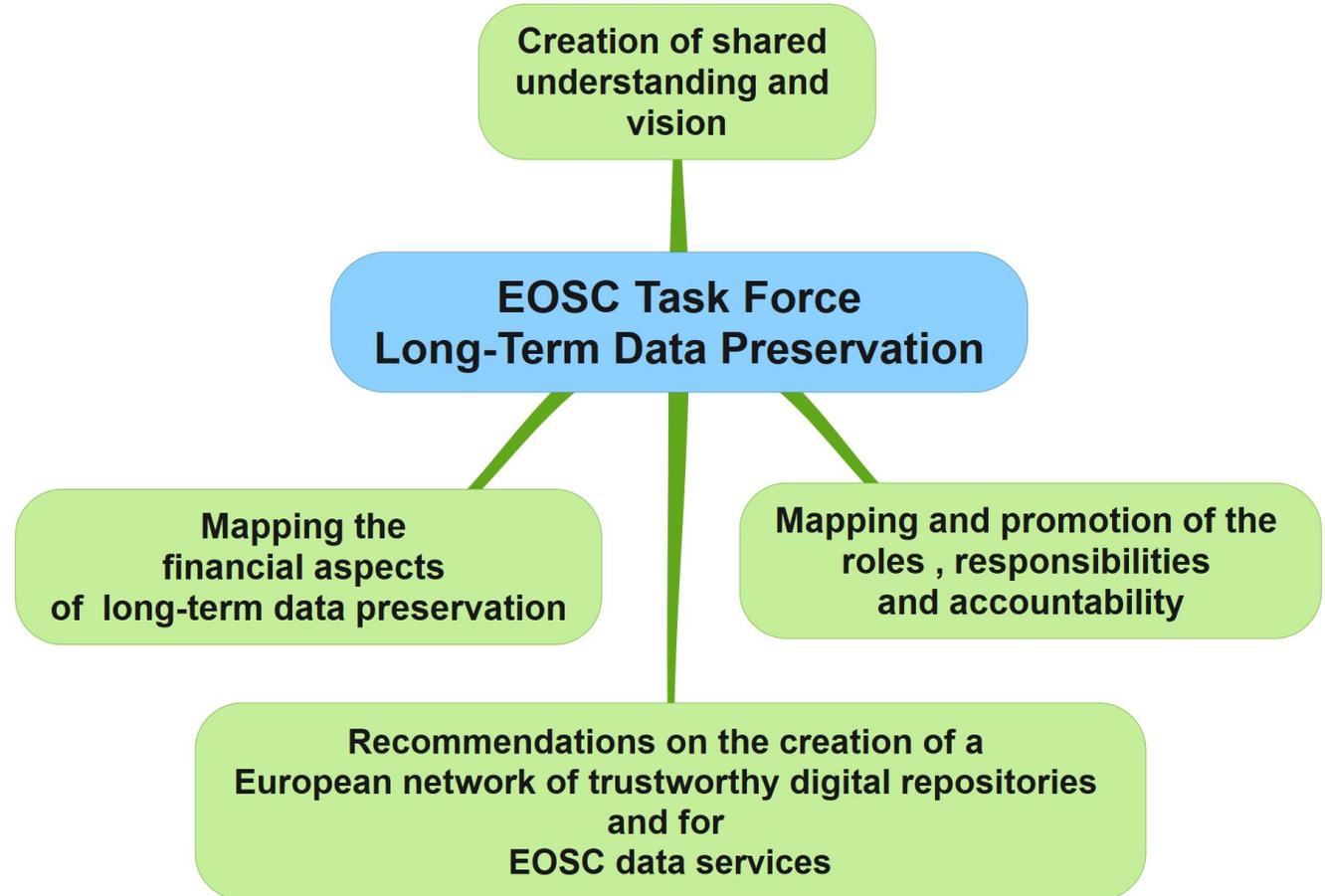
Long Term Data Preservation Task Force

Sub-tasks

Long Term Data Preservation Task Force

Workflow

HLH-LTDP-TF-Subtask-Overview_v00_01.vsdx

Understand the Present; Vision for the Future

Overview and discussion paper for broad feedback

Iterative consultation on recommendations:

- European,
- National,
- Institutional Level.

LTDP-TF Finance & Roles

Can roles and finances 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

LTDP TF Network of TDRs

Towards a European network of FAIR-enabling Trustworthy Digital Repositories

- 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

LTDP Trust Network Subtask in progress.

- Supporting, contributing, aligning
- FAIR IMPACT Synchronisation Force as reference point for aligned actors and implementation
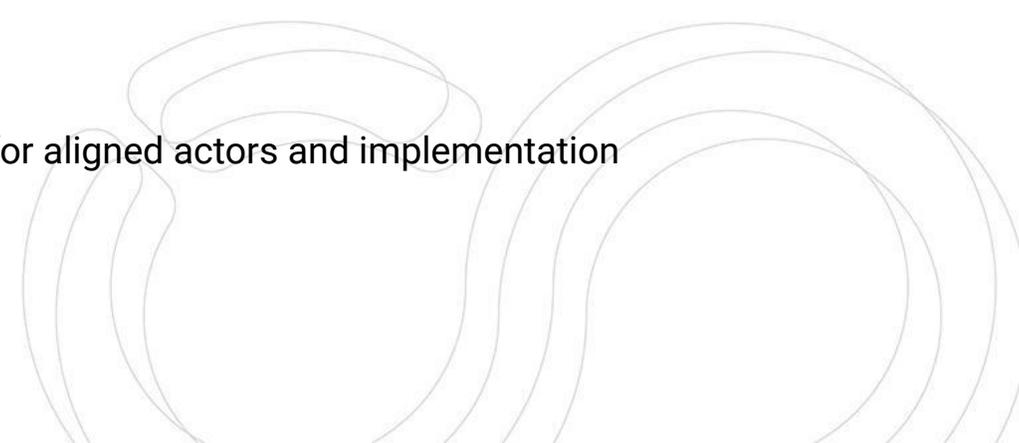
LTDP TF Network of TDRs

Towards a European network of FAIR-enabling Trustworthy Digital Repositories

- 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"

LTDP TF Network of TDRs

Towards a European network of FAIR-enabling Trustworthy Digital Repositories

- 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

Long-term Data Preservation
Task Force

eosc

Towards a European network of FAIR-enabling Trustworthy Digital Repositories

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

LTDP TF Network of TDRs

Towards a European network of FAIR-enabling Trustworthy Digital Repositories

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

Outcomes, Systems and Actions

Preservation Outcomes

Objects are maintained for the long term with characteristics desirable to the designated community.

Preservation Systems

Provide services for deposit, storage, access and long term preservation.The designated community's knowledge base and technological needs are understood and monitored over time. These systems have sufficient resources, including personnel and financial resources, to be sustainable in terms of their organisational, technology and security infrastructure.

Preservation Actions

Changes to digital objects' (metadata and data) to keep them usable, understandable and accessible by their designated community.

These actions include technical steps such as emulation or transformations to modern, in-demand, file formats and metadata schemas; but also actions to ensure that the conceptual content of data and metadata, including semantic artefacts such as ontologies, continues to be understood and re-usable.

Assessing Repositories vs Objects

CoreTrustSeal Repository Assessment

FAIR digital object assessment. E.g. F-UJI tool

Both the subject of discipline-specific use cases in FAIR IMPACT Project

Repository

Contains   Is Contained By
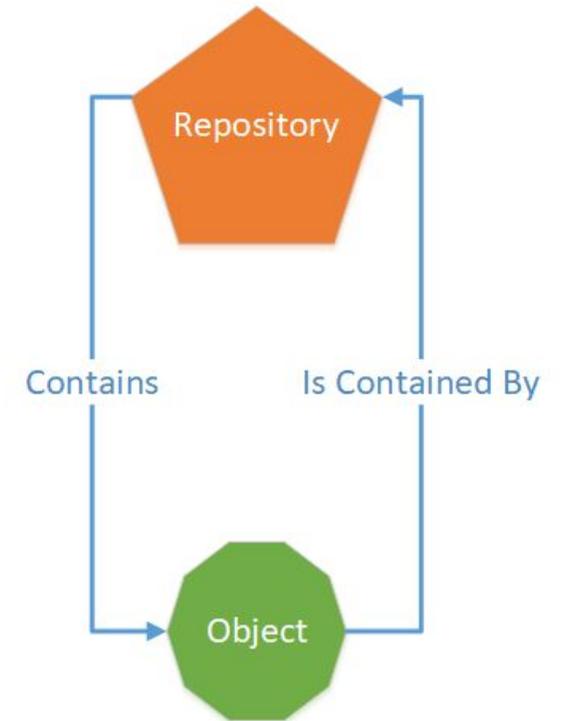
Object

Repository ObjectMetadata A_00_01.vsdx

Evidence

Artefacts that support practice…

provide evidence…

to demonstrate compliance.

Artefacts:

- Policy
- Standard Operating Procedure
- Metadata

eosc | Long-term Data Preservation Task Force
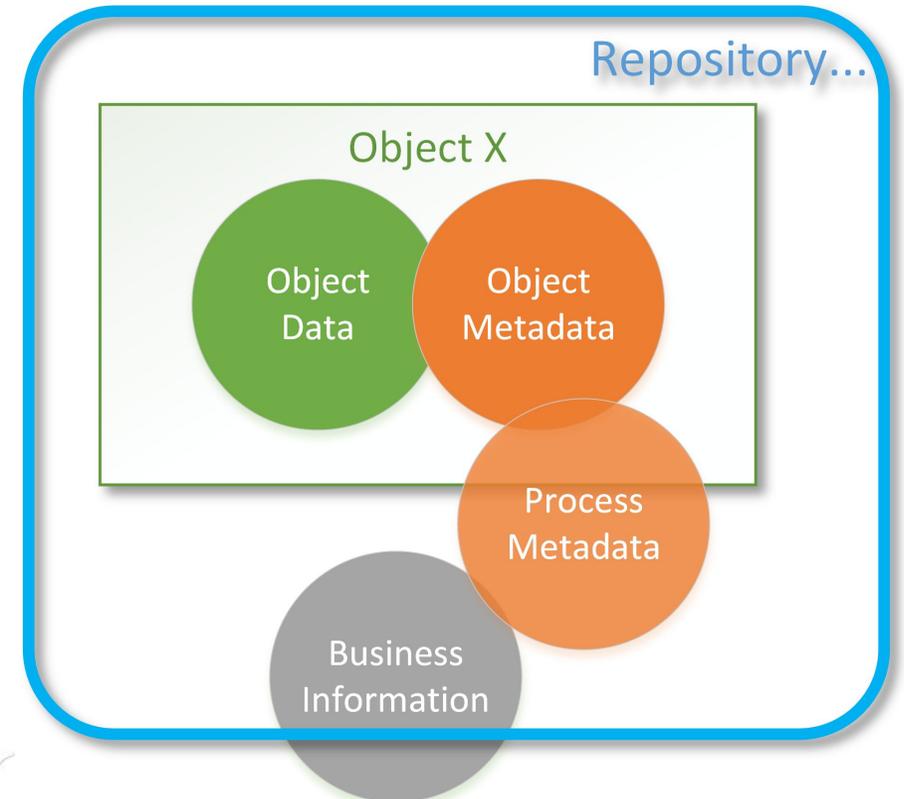
CoreTrustSeal 2020-2022
Mapped to FAIR

FAIR + Time
Requires
Preservation

Revised during
FAIR IMPACT
Project

F R13

F1. (meta)data are assigned a globally unique and persistent identifier.
F2. data are described with rich metadata (defined by R1 below).
F3. metadata clearly and explicitly include the identifier of the data it describes.
F4. (meta)data are registered or indexed in a searchable resource.
R13. Data discovery and identification

A R10 R15 R16

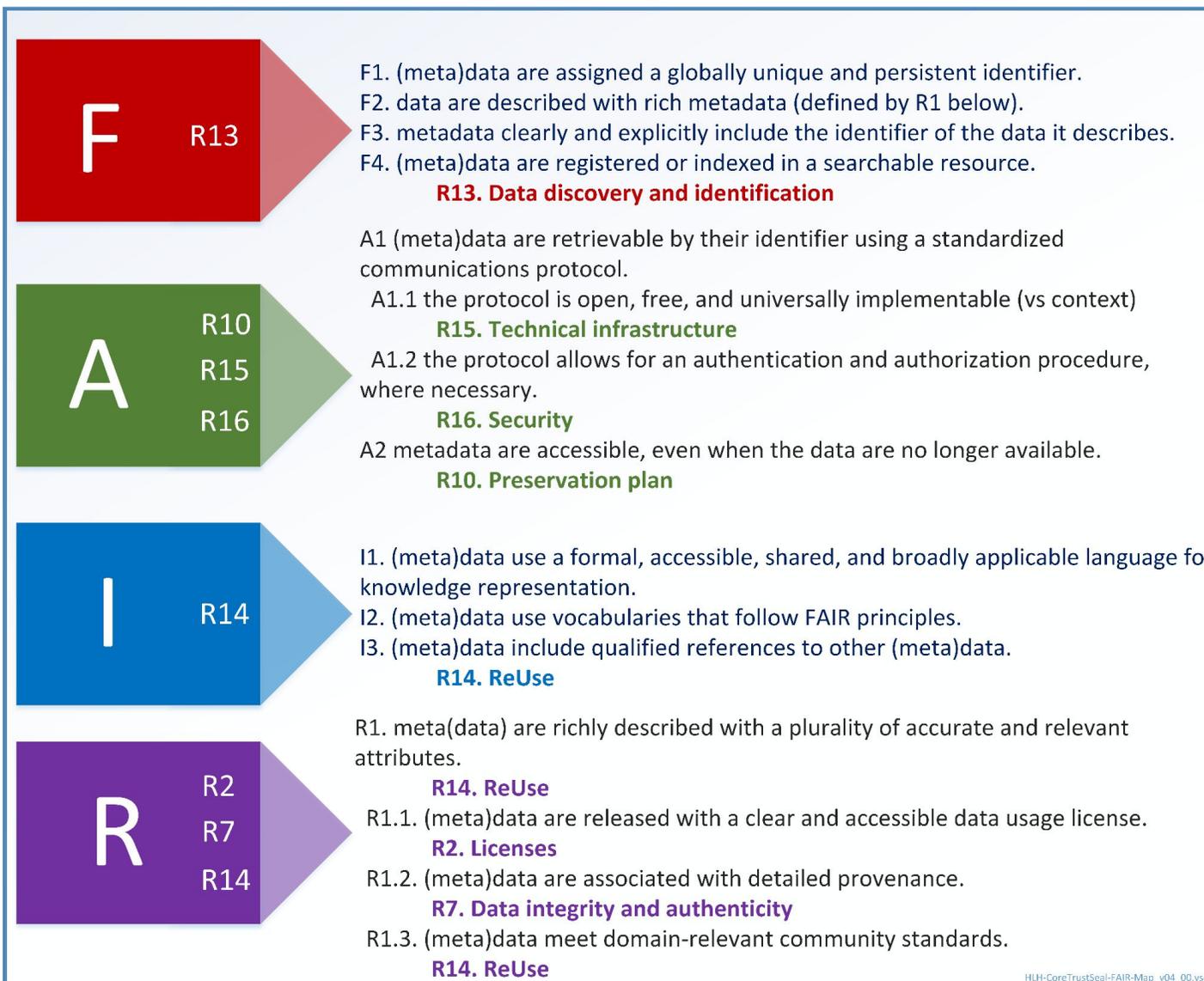A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
A1.1 the protocol is open, free, and universally implementable (vs context)
R15. Technical infrastructure
A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
R16. Security
A2 metadata are accessible, even when the data are no longer available.
R10. Preservation plan

I R14

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles.
I3. (meta)data include qualified references to other (meta)data.
R14. ReUse

R R2 R7 R14

R1. meta(data) are richly described with a plurality of accurate and relevant attributes.
R14. ReUse
R1.1. (meta)data are released with a clear and accessible data usage license.
R2. Licenses
R1.2. (meta)data are associated with detailed provenance.
R7. Data integrity and authenticity
R1.3. (meta)data meet domain-relevant community standards.
R14. ReUse

HLH-CoreTrustSeal-FAIR-Map_v04_00.vsdx

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

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

CoreTrustSeal Curation & Preservation Levels Discussion Paper

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

Storage, Curation & Preservation Levels

Community discussions on levels of preservation https://doi.org/10.5281/zenodo.6908018

Z. Level Zero. Content distributed as deposited. Unattended deposit-storage-access.

Distributed as deposited

C. Basic Compliance and/or curation

Compliance checks and/or curation to meet compliance at deposit. No long term preservation

B. Logical-Technical Curation

In addition to C. Long term responsibility for updates to newer standards and formats in response to:

i. technical risks (e.g. file format obsolescence) and/or

Ii. the changing needs of the designated community (e.g. newer alternate formats become necessary for reuse).

A. Conceptual preservation for understanding and reuse

In addition to B and C.

• Monitors changes to the definition and demands of their designated community, including their knowledge base,

• responsibility for the preservation actions that ensure digital objects can be understood and re-used.

Usually this will involve updates to the content of metadata elements and other semantic artifacts such as controlled vocabularies and ontologies. For some repositories it may include responsibility for editing the structure and content of deposited data.

Long-term Data Preservation
Task Force

LTDP Task Force Overview

CESSDA Landscape Event

Hervé L'Hours & Roxanne Wyns: LTDP TF Co-Chairs

2022-11-09 Hervé L'Hours
Task Force Co-Chair,
UK Data Service,
Consortium of Social Science Data Archives (CESSDA)

New CoreTrustSeal requirements

Mari Kleemola, Tampere University/FSD
CoreTrustSeal Board member

CESSDA Landscape event, 11 November 2022 (online webinar)

cessda.eu @CESSDA_Data

About CESSDA

Cross-European **resource discovery**

Improved **quality** of data and metadata

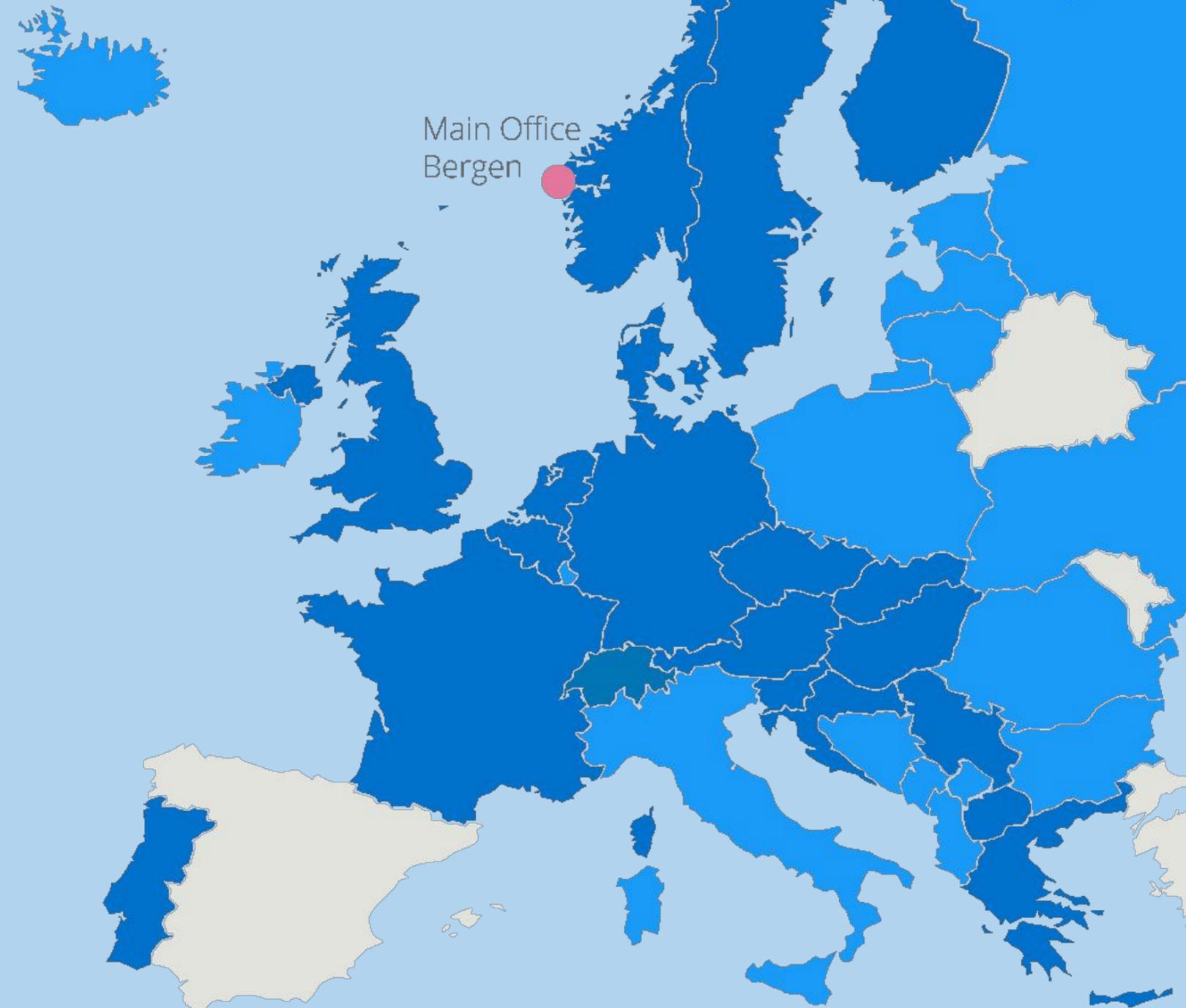A wider selection of **comparable data**

Certification of data archiving organisations

Professional **training** for data archivists and scientific community

Improved mechanisms for **data access and analysis**

Strong involvement of organisations **outside Europe**

20 members

16 partners

Main Office Bergen

CoreTrustSeal Trustworthy Digital Repositories
Requirements 2023-2025

V01.00

- Revision of Requirements every three years
- Community effort – lot of comments and feedback from the community
- Global
- Domain agnostic

cessda

CoreTrustSeal Governance and Community

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

5. **Aude Chambodut** (International Service of Geomagnetic Indices, France
6. **Reyna Jenkyns** (Ocean Networks Canada, Canada)
7. **Dawei Lin** (ImmPort Repository, DAIT-NIAID-NIH, USA)
8. **Jared Lyle** (Inter-university Consortium for Political and Social Research, University of Michigan, USA)
9. **Jonas Recker** (GESIS-Leibniz Institute for the Social Sciences, Germany)
10. **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)

Assembly of reviewers:
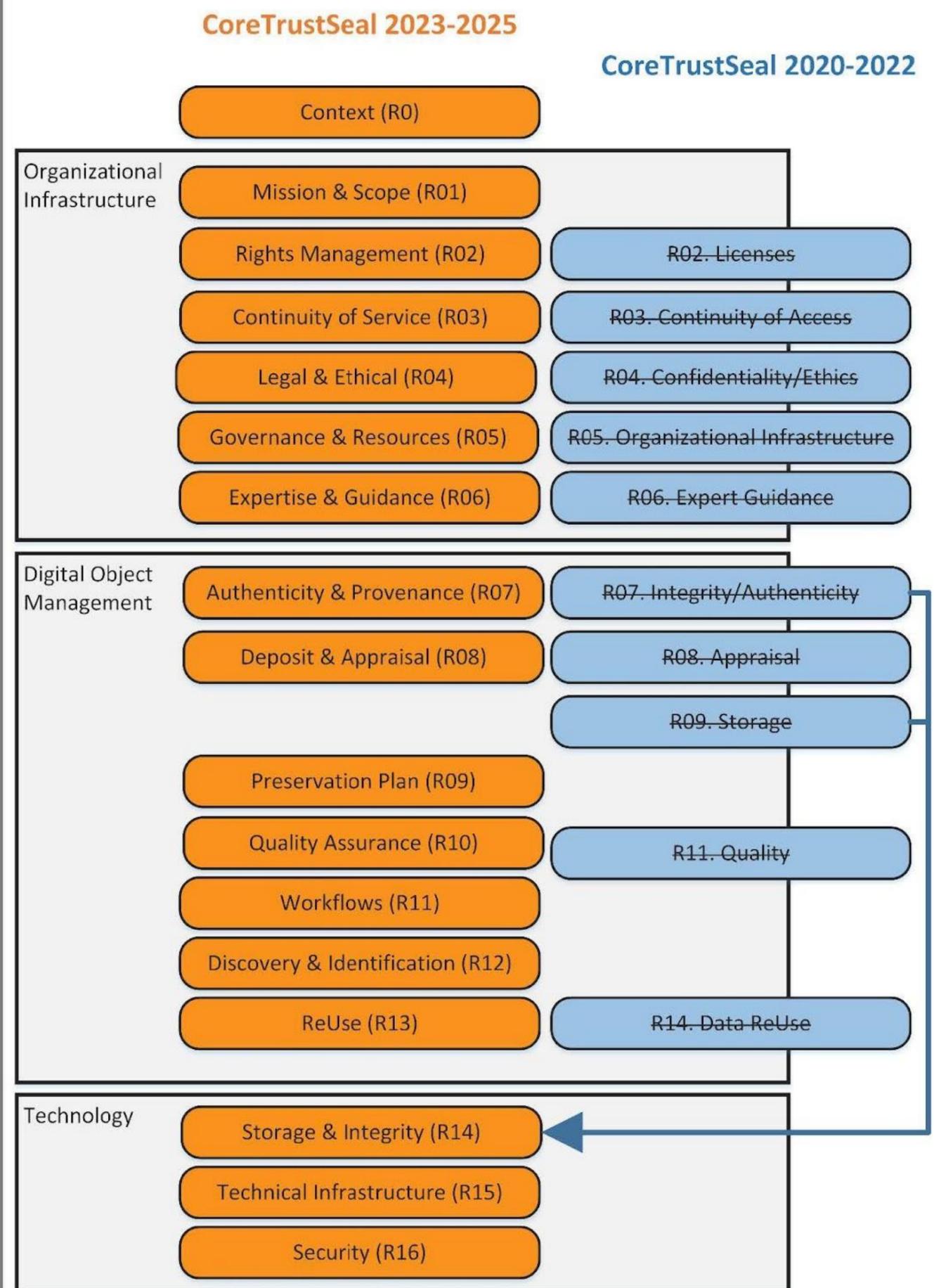https://www.coretrustseal.org/about/assembly-of-reviewers/
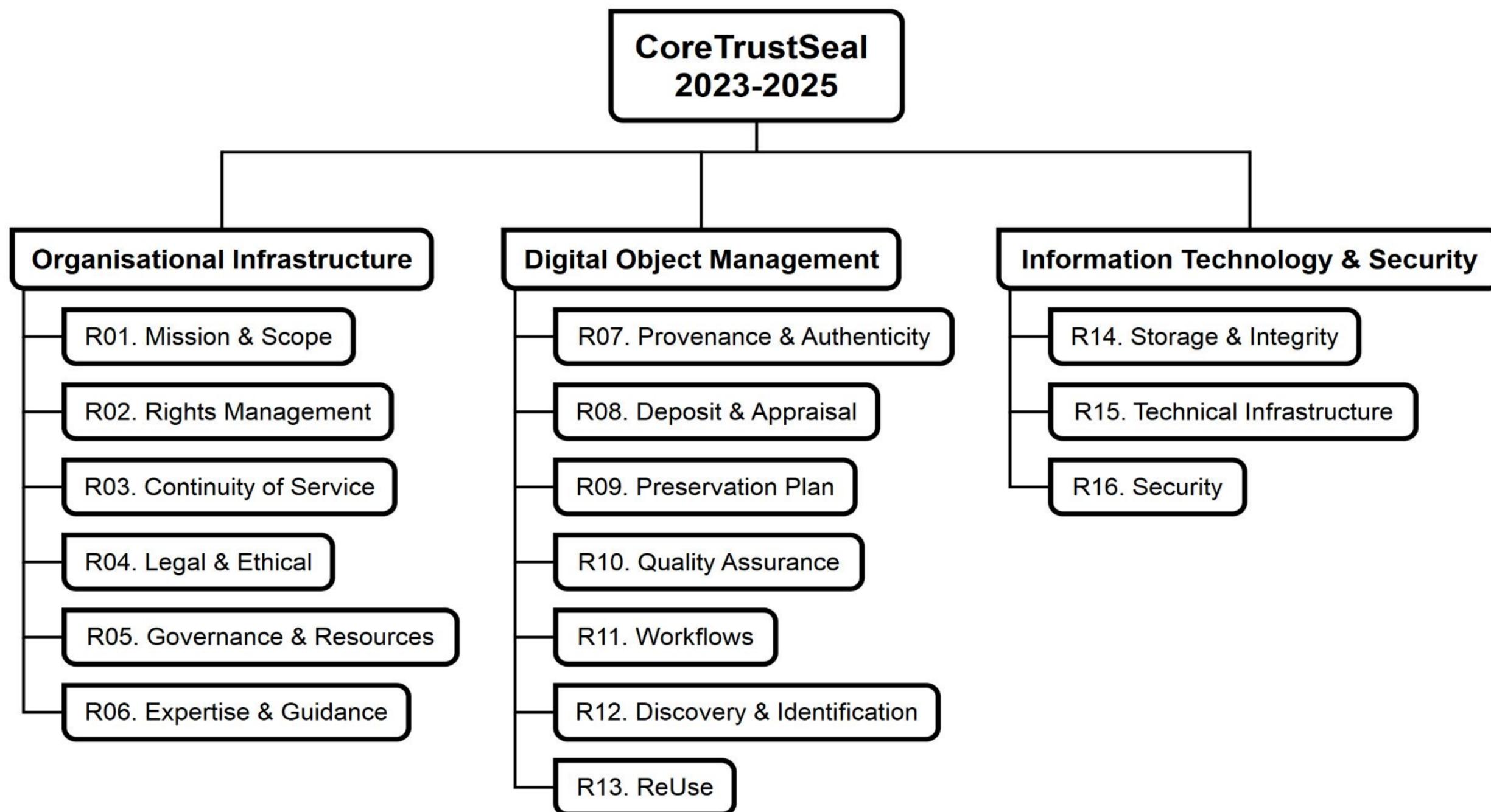
cessda

Overview of changes

Changes include:

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

Change log available!

Image by Hervé L'Hours

CoreTrustSeal 2023-2025

CoreTrustSeal 2020-2022

Context (R0)

Organizational Infrastructure

CoreTrustSeal 2023-2025	CoreTrustSeal 2020-2022
Mission & Scope (R01)	
Rights Management (R02)	R02. Licenses
Continuity of Service (R03)	R03. Continuity of Access
Legal & Ethical (R04)	R04. Confidentiality/Ethics
Governance & Resources (R05)	R05. Organizational Infrastructure
Expertise & Guidance (R06)	R06. Expert Guidance

Digital Object Management

CoreTrustSeal 2023-2025	CoreTrustSeal 2020-2022
Authenticity & Provenance (R07)	R07. Integrity/Authenticity
Deposit & Appraisal (R08)	R08. Appraisal
	R09. Storage
Preservation Plan (R09)	
Quality Assurance (R10)	R11. Quality
Workflows (R11)	
Discovery & Identification (R12)	
ReUse (R13)	R14. Data ReUse

Technology

CoreTrustSeal 2023-2025
Storage & Integrity (R14)
Technical Infrastructure (R15)
Security (R16)

cessda

The new 2023-2025 Requirements

Compliance levels

The applicant must indicate a compliance level for each Requirement

- 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.

cessda

Clarifications

- 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".

- 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.

cessda

Examples of changes to Requirements

Rights Management (R02)

R02. The repository maintains all applicable rights and monitors compliance.

Was: 2. Licenses. R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.

Updated to reflect the fact that rights management goes beyond the traditional signing of a license agreement at the point of deposit or access and includes all the measures necessary to manage the permission, prohibitions and obligations of all actors involved in managing data and metadata. Many digital objects have some rights attached even if there is no license artifact as traditionally understood.

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.

cessda

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.

Quality Assurance (R10)

R10. The repository addresses technical quality and standards compliance, and ensures that sufficient information is available for end users to make quality-related evaluations.

Was: 11. Data quality. 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.

Repository quality assurance is often related to 'standards compliance'. The requirement is intended to demonstrate that the repository provides data and metadata of sufficient 'technical quality'. This should be sufficient to allow users to make assessments about their 'scientific quality'. References to 'expertise' are removed to avoid overlap with Expertise & Guidance (R06)

cessda

Workflows (R11)

R11. Digital object management takes place according to defined workflows from deposit to access.

Was: 12. Workflows. R12. Archiving takes place according to defined workflows from ingest to dissemination.

The language of the Requirement has been updated to reflect that most commonly used within the applicant community.

cessda

Storage & Integrity (R14)

R14. The repository applies documented processes to ensure data and metadata storage and integrity.

Was: 9. Documented storage procedures R9. The repository applies documented processes and procedures in managing archival storage of the data.

It was clear from previous applications that the topic of storage was addressed primarily in technical terms. This change moves Storage into the Information Technology and Security sub-section and unites it with integrity (previously included under R07) to cover the avoidance of unintended changes to data and metadata.

cessda

Next steps

- CoreTrustSeal Submissions are paused to allow transition
- Nov-Dec: work on existing submissions
- New applications will be accepted starting mid-January 2023
- The Application Management Tool (AMT) renewed, too
- Training for the Assembly of Reviewers in December
  - Training session on Tuesday Dec 6 & Open Coffee Hour on Thursday Dec 8. The Assembly of Reviewers have been invited separately. Please note that this will be on invitation only but parts will be published on CoreTrustSeal Youtube channel afterwards

cessda

Sources and links

- CoreTrustSeal: https://www.coretrustseal.org/

- CoreTrustSeal Requirements 2023-2025: https://doi.org/10.5281/zenodo.7051011

- Extended Guidance: https://doi.org/10.5281/zenodo.7051095

- CoreTrustSeal Revision Working Group Change Log and Associated Materials: https://doi.org/10.5281/zenodo.7035111

cessda

Thank you!

cessda.eu

@CESSDA_Data