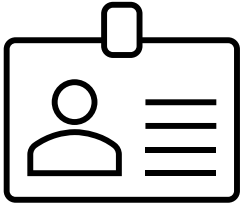


Context, approach, and progress

of the EOSC-A Long Term Data Preservation Task Force





- **Roxanne Wyns, innovation manager @ KU Leuven LIBIS**
 - Working with researchers in the design and development of scalable and sustainable RI
 - Part of the RDM steering comity @KU Leuven
 - FAIR + Time
 - Institutional Research Data Repository (RDR)
 - EOSC-A
 - Representative in the GA
 - Co-chair of the LTDP-TF together with Hervé L'Hours (UK Data Archive & UKDS)

Context of the LTDP-TF

LTDP in the Strategic Research Agenda

“Moreover, long-term open data archives and preservation services are required to enable a **sustainable EOSC and the sustainable access to data**. Data preservation not only refers to the long-term storage of data, but also includes ensuring the preservation and maintenance of data, as well as its context, understandability, interpretability, authenticity and integrity.”

“The interim findings of the FAIR Forever study, conducted by DPC on behalf of the Sustainability Working Group, noted that **digital preservation is not explicit in the context of EOSC** and the roles, responsibilities and accountability for digital preservation are currently not clearly defined. ”

“ **Clearer roles and responsibilities are needed**, including the assessment of capability as well as functions, salaries and funding streams for preservation.”

Source: SRIA v1.0, 21/06/2021

(https://www.eosc.eu/sites/default/files/SRIA_2022_01.pdf)

Context of the LTDP-TF

FAIR Forever recommendations with urgent priority

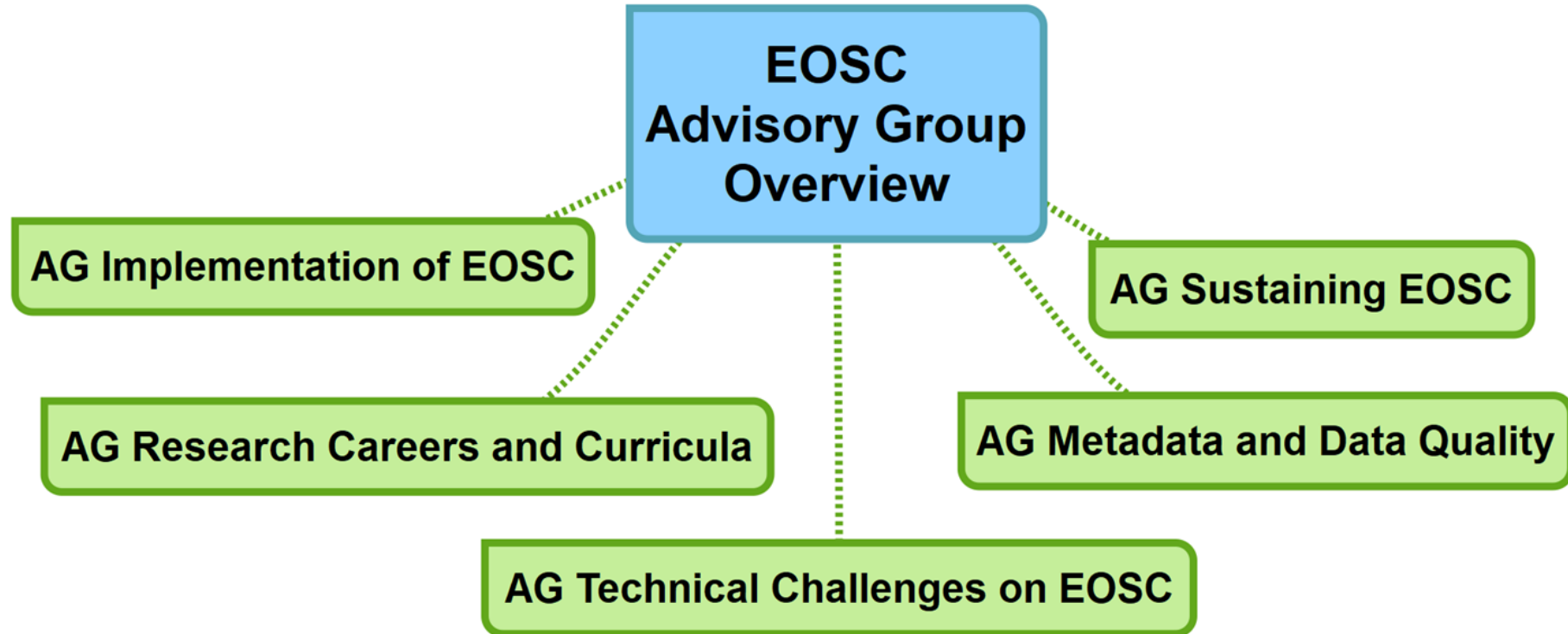
Recommendations for the EOSC Secretariat and EOSC-A with urgent priority status:

- Establish a working party or task group, reporting directly to the EOSC Association Board with respect to digital preservation
- Designate a Senior Digital Preservation Rapporteur on behalf of the Board to directly communicate and liaison with a Digital Preservation Task Group, to monitor and oversee EOSC's responses to digital preservation risks

Source: FAIR Forever study, conducted by DPC on behalf of the Sustainability Working Group, 17/02/2021
(<https://doi.org/10.5281/zenodo.4574234>)

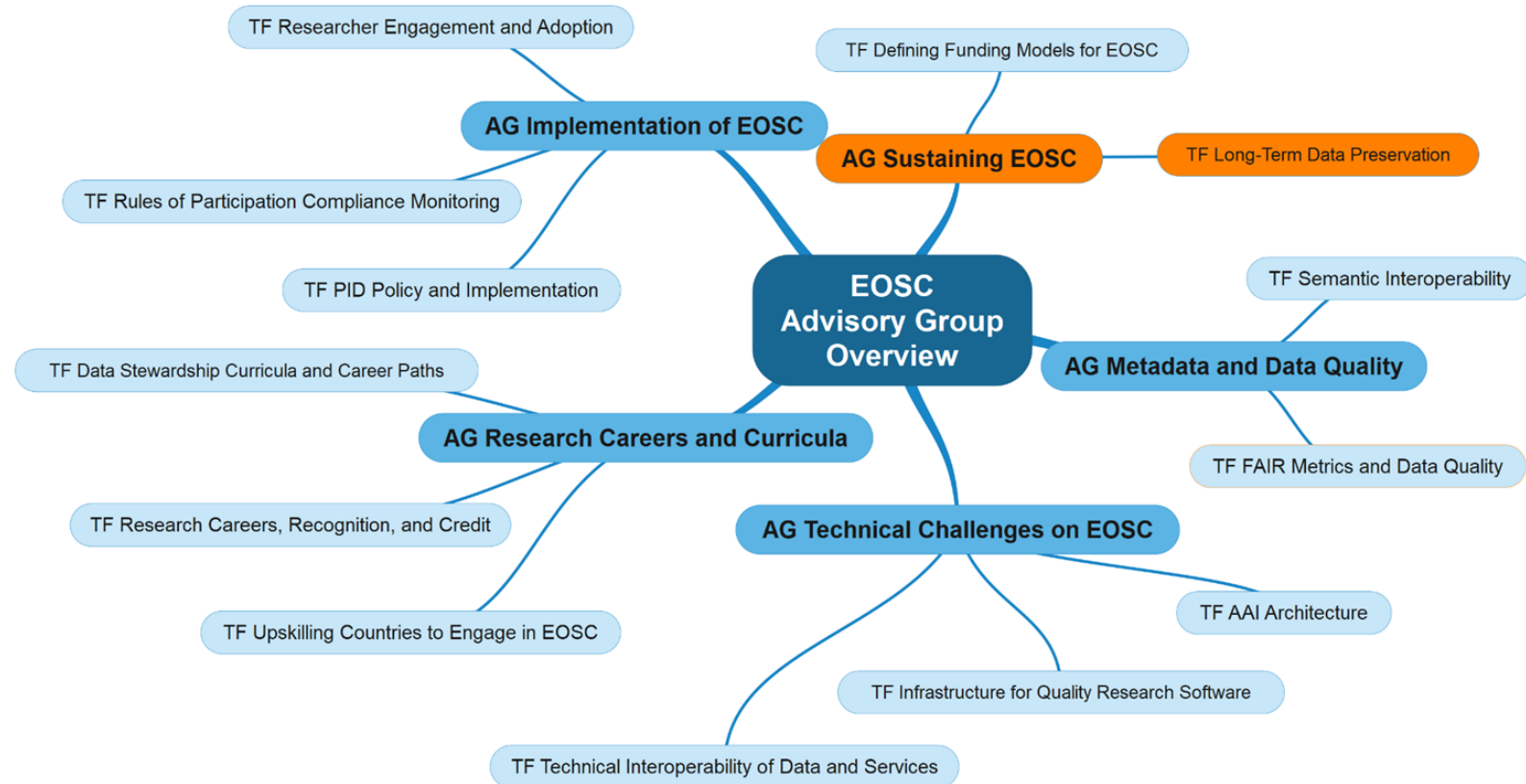
Context of the LTDP-TF

EOSC-A Advisory groups



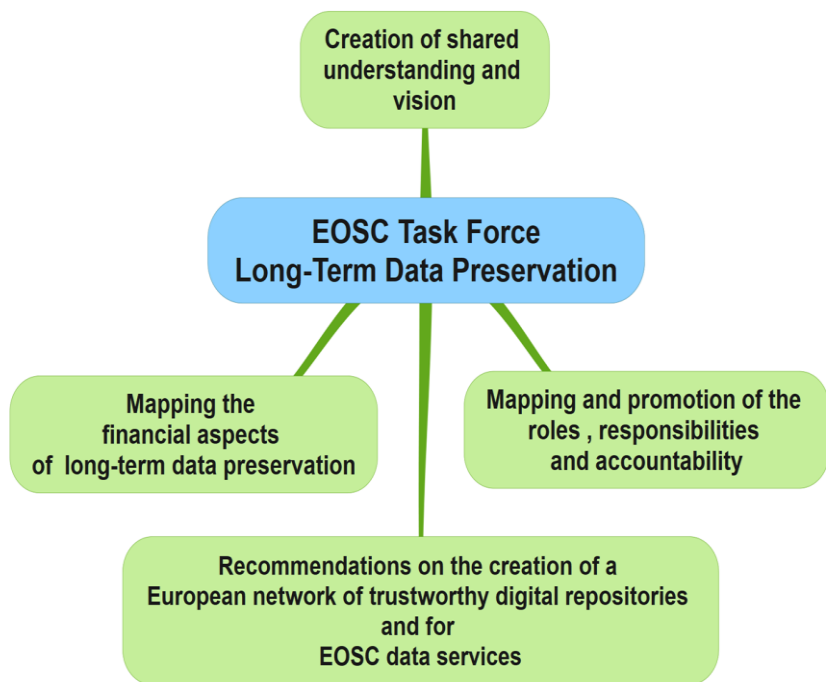
Context of the LTDP-TF





AG Sustaining EOSC – Long-Term Data Preservation TF



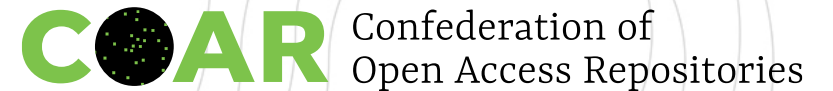
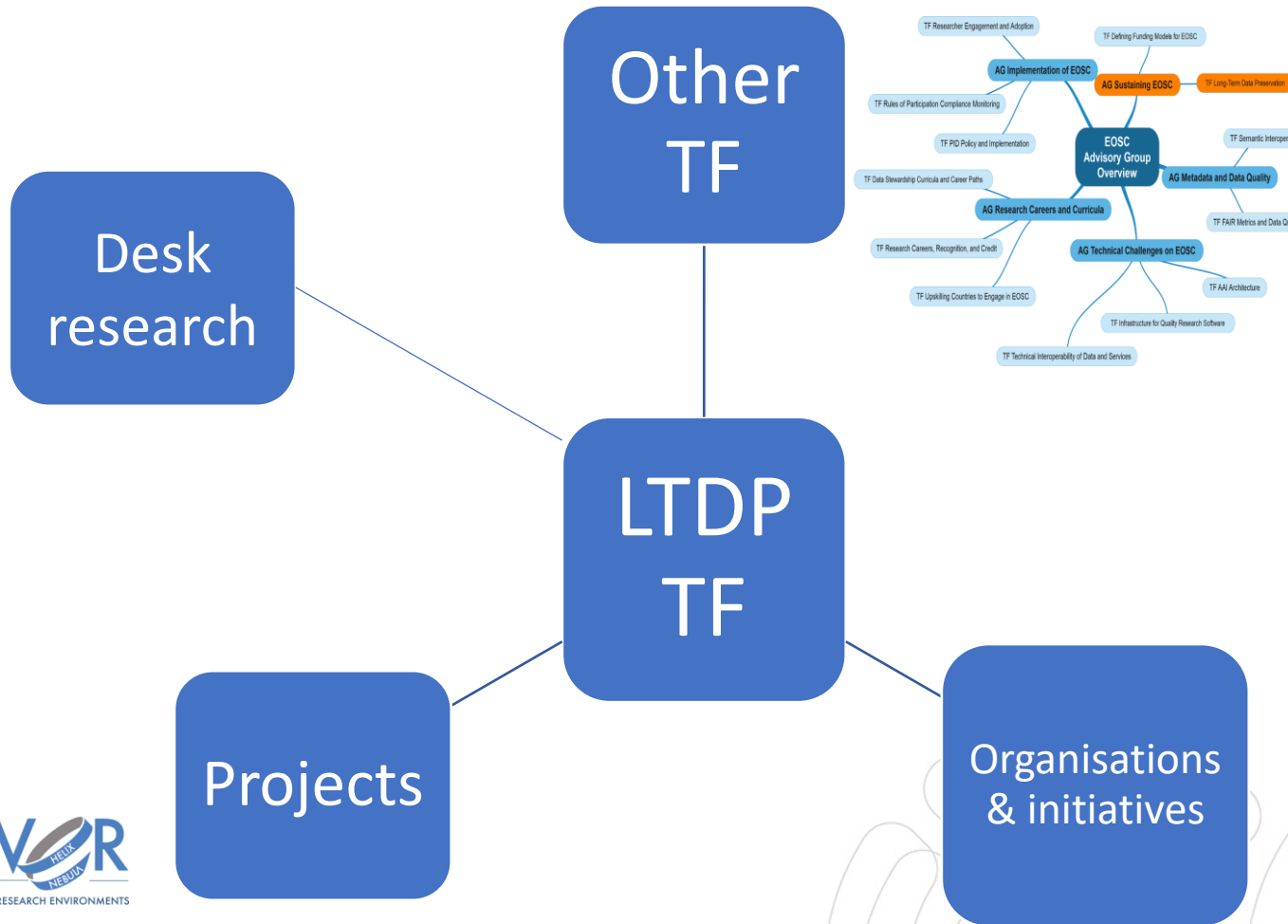
Context of the LTDP-TF

TF structure



<p>Co-chairs:</p> <ul style="list-style-type: none"> Hervé L'Hours (UK Data Archive) Roxanne Wyns (KU Leuven) 	<p>EOSC-A Liaison:</p> <ul style="list-style-type: none"> Bob Jones (BoD) René Buch (CTO)
<p> Shared Understanding and Vision Subgroup lead: Chris De Loof (Belnet) Pirjo-Leena Forsström (CSC), Rui Fernandes (C4G), Gerardo Ganis (CERN), Ville Tenhunen (EGI), Andras Holl (MTA), Maciej Brzeźniak (PSNC), Lara Lloret Iglesias (CSIC), Riccardo Smareglia (INAF), Eileen Gibney (UCD)</p>	<p> Mapping and promotion of roles Subgroup lead: Mariusz Majdański (IG PAS) Andrea Lammert (DKRZ), Sangeetha Shankar (DLR), Christian Cuciniello (EC), Marcello Maggi (INFN), Florina Piroi (TU WIEN), Mojib Wali (TU GRAZ), Sabine Crépé-Renaudin (CNRS), Bregt Saenen (Science Europe)</p>
<p> Mapping of financial aspects Subgroup lead: Paul Stokes (JISC) Jean-Yves Nief (CC-IN2P3), Jiri Novacek (CEITEC), Toni Andreu (EATRIS), Martina Stockhause (IPCC DDC)</p>	<p> Recommendations TDR Subgroup lead: Didi Lamers (Radboud university) Olesia Dubois (Sciences Po Paris), S. Venkataraman (OpenAire), Pierre-Yves Burgi (OLOS.swiss), Lluís Anglada (CSUC), Cécile Cavet (Univ. Paris), Ingrid Dillo (DANS), David Antos (CESNET), Draženko Celjak (SRCE), Matthew Viljoen (EGI)</p>

Approach



August 30, 2022

Towards a European Trustworthy Working Paper

Philippe Conze, Ingrid Dillo, Kurtz, Pedro Principe, Oliver

This working paper is a bottom-up purpose is to outline an aspirational (TDRs). This initiative originates paper was created in close connection with the first draft of the paper was shared on Long Term Digital Preservation recommendations on the creation of TDRs.

The working paper puts together community's needs and its most engagement, and coordination and development. The specific activities hosted under these umbrella functions could address the wide range of topics that are important to TDRs. Beyond these functions and the challenges they address, the paper presents a framework to highlight aspects of the Network to further explore in the next steps of its development.

This Working Paper was created with input from the European repository community (17th June - 4th July 2022)

Preview

Page 1 of 21 | 110%

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

A Working Paper

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

Approach and Progress

Overview discussion paper

- Working on an overview and discussion paper that presents the context, activities and ongoing issues surrounding the EOSC-A LTDP-TF
- In review by the EOSC-A Quality and review committee
- To be released for public consultation, seeking comments from a wide range of stakeholders across the digital object management lifecycle
- Feedback to be integrated into an updated overview
- Consultation on the recommendations addressing digital objects' preservation through EOSC at the European, national, and institutional level

Preservation in the context of EOSC and FAIR

Overview discussion paper

Defining what we mean:

- Preservation outcomes: these concern digital objects that, having been curated for FAIRness and other desirable characteristics, are maintained to retain those characteristics for as long as necessary ...
- Preservation systems: accept the deposit of digital objects for storage, access, and also curate them with a long term FAIR-enabling perspective on the objects and their designated community ...
- Preservation actions: are the changes to digital objects that are intended to keep them FAIR over time ...

Preservation *outcomes* depend on *actions* by *systems* with a sustainable, long term perspective

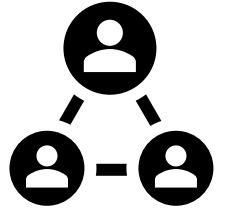
Importance of:

- FAIR (Findable Accessible, Interoperable and ReUsable) digital objects
- Trusted Repositories
- Transparency over current levels of FAIRness and trustworthiness (current status, plans for improvement)

Vision in progress

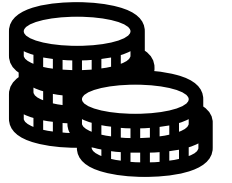


- Ideally, all data of interest to EOSC researchers retain their value and remain FAIR for a defined community for as long as necessary (FAIR+Time)
- But resources are not infinite so a nuanced approach in the vision and direction of the TF needed
- Vision concepts:
 - FAIR
 - Appraisal and reappraisal
 - Retention, curation, long-term preservation
 - Transparency on level of care provided by repositories



- Gaps in understanding the roles of various actors
- Problems in identifying and financing those with the skills to take responsibility for curation and preservation
- A need for support from specialised training roles that provide general and specialist guidance
- Different roles, responsibilities and accountability on institutional, national and European level
- Additional responsibilities for data creators, technicians, data stewards etc, all incur costs → Finance subgroup

FAIR Forever study as the basis for the further identification of roles

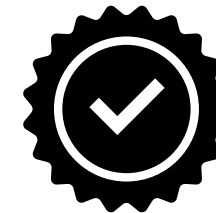


- Examine and collate existing cost models (e.g. curationexchange.org, 4C Project)
- Look at preservation-specific costs vs. other data services costs and define subset of preservation costs
- What are the costs for a community if a digital object of value for reuse is not cared for
- Problem of short term funding
- Resources are not infinite, so appraisal and reappraisal is critical over time → Also incurs costs
- Practices will differ between generalist and specialist domain repositories → Cooperation for sharing expertise, services and costs
- Balance between potential of automation and maintaining enough resources of human expertise
- Identifying the likely costs of preservation at a sufficiently granular level is challenging. The same goes for calculating the costs of inaction with risks of data loss

Network of Trust



- Vision can only be achieved through an ongoing engagement with curation and preservation professionals
- In need of a body that integrates the expertise of, and provides a voice for trustworthy digital repositories
- Reviewing relevant stakeholders and initiatives



Providing feedback on existing initiatives

Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDRs) - A Working Paper

(30/08/2022, <https://doi.org/10.5281/zenodo.7034315>)

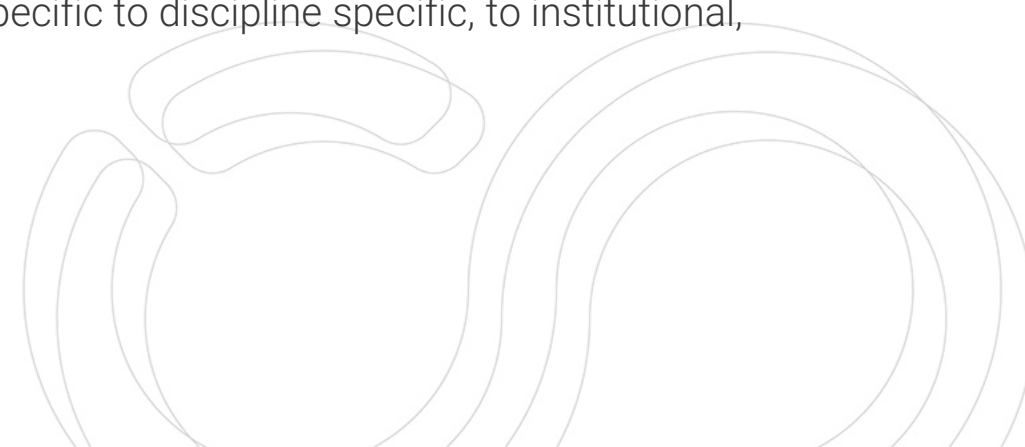
- **Functions of the network**

- Networking and knowledge exchange function
- Stakeholder advocacy and engagement function
- Coordination and development function

- **Membership**

- The goal is to coordinate existing FAIR-enabling TDRs and increase the number of such repositories in the European research landscape
- Therefore an inclusive network is envisioned, open for TDR as well as repositories aspiring trustworthy certification
- Open range of repository types, from domain specific to discipline specific, to institutional, regional/national, generic

Preparing a response from the TF





Thank you!
Questions?

Email: Roxanne.Wyns@kuleuven.be