Towards a network of European FAIR-enabling Trustworthy Digital Repositories

Co-organised by FAIRsFAIR, SSHOC and EOSC-Nordic projects

Online workshop - 13 January 2022

Trigger

- FAIRsFAIR, EOSC-Nordic and SSHOC repository support initiatives seeking to sustain the efforts.

Today's objective

- Explore ideas and needs of such a network.

Proposed outcome

- Identify steps to progress towards a European TDR network.

https://www.fairsfair.eu/events/fairsfair-event/workshop-network-fair-enabling-trustworthy-digital-repositories-tdrs

Programme

Introduction
11:00 - 11:10 Welcome and introduction of the initiative (Ingrid Dillo, DANS)
11:10 - 11:15 Horizon Europe, INFRAEOSC and a TDR network (Christian Cuciniello, European Commission)
11:15 - 11:20 EOSC Task Force Long Term Data Preservation (Hervé L'Hours, UKDA)

Overview of repository support approaches
11:20 - 11:25 FAIRsFAIR (Linas Cepinskas, DANS)
11:25 - 11:30 SSHOC (Henri Ala-Lahti, Finnish Social Science Data Archive)
11:30 - 11:35 EOSC-Nordic (Tuomas J. Alaterä, Finnish Social Science Data Archive)
11:35 - 11:40 Q&A

Towards a network of European FAIR-enabling TDRs
11:40-11:45 Towards a TDR network (Ilona von Stein, DANS)
11:45-12:25 Discussion
12:25-12:30 Wrap-up

*All times indicated above are in CET.

Registration

- Please add your name, affiliation and email address to the minutes:
  https://bit.ly/TDRnetwork

Zoom call

- Please mute your microphone if you are not speaking.
- Please feel free to ask questions via microphone or chat.

Recording

- This workshop will be recorded.
- Recording and slides will be made available on FAIRsFAIR website.
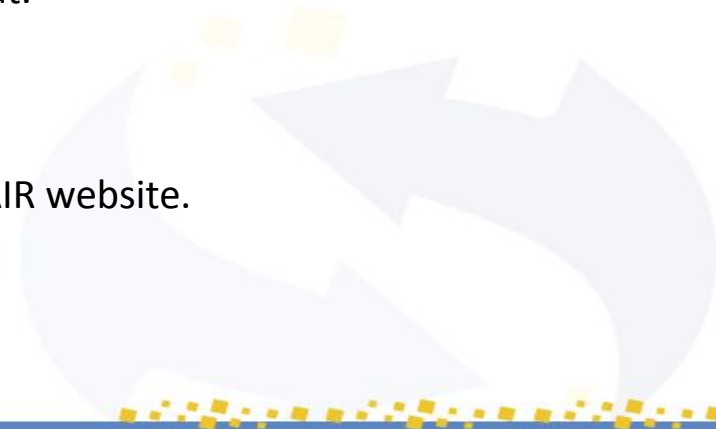
European network of FAIR-enabling TDRs
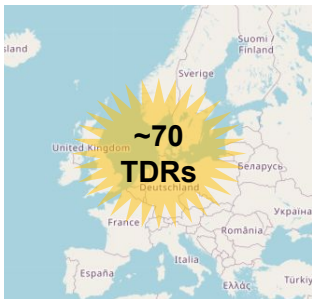
TRUST and FAIR certifications

European FAIR-enabling TDR Network

- *Cross-disciplinary/border coordination*
- *Trustworthy (LTP) + FAIR enabling*

~70 TDRs

EUROPEAN OPEN SCIENCE CLOUD

Geographical spread
European CTS repositories

European network of FAIR-enabling TDRs

TRUST and FAIR certifications

Disciplinary Infrastructures *e.g.*

~70 TDRs

Regional/National Networks *e.g.*

European FAIR-enabling TDR Network

- *Cross-disciplinary/border coordination*
- *Trustworthy (LTP) + FAIR enabling*

EUROPEAN OPEN SCIENCE CLOUD

European network of FAIR-enabling TDRs

TRUST and FAIR certifications

Disciplinary Infrastructures *e.g.*

Regional/National Networks *e.g.*

~70 TDRs

European FAIR-enabling TDR Network

- *Cross-disciplinary/border coordination*
- *Trustworthy (LTP) + FAIR enabling*

Technology providers *e.g.*

EUROPEAN OPEN SCIENCE CLOUD

Potential value of the network

For repositories:
- *Promote adoption of FAIR and TRUST practices*
- *Strengthen the voice of TDRs in Europe*

For researchers:
- *Facilitate researchers' access to Long Term Preservation*

For the EOSC:
- *Enhance repository compliance with EOSC RoP*
- *Help implement FAIR and TRUST assessments*

Potential value of the network

For all stakeholders involved:
***Keep** data FAIR over time*

FAIRSFAIR
Fostering Fair Data Practices in Europe

SRIA of the EOSC 2021:
EOSC will ensure long-term preservation and long-term availability of these research data and tools.
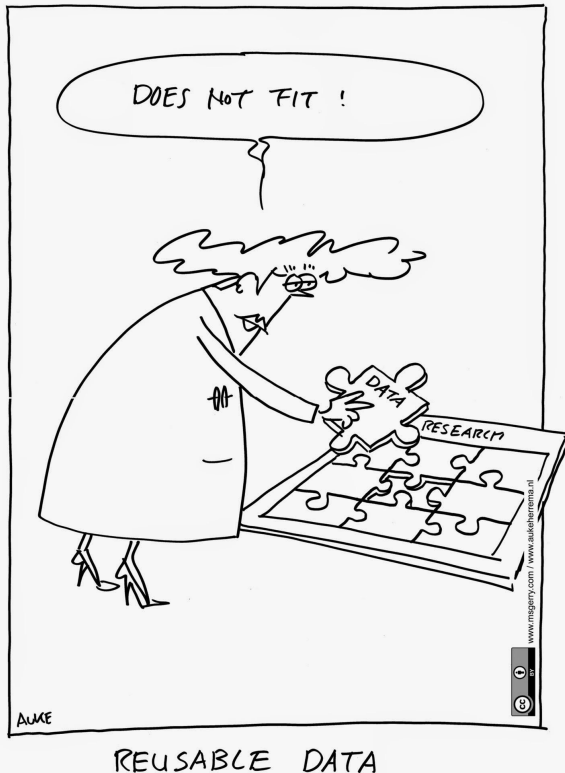
Long-term access to and preservation of data

Objective

Support preservation actions and long-term access for EOSC content.

Approach

- Long-term preservation identified as a gap in the Turning FAIR into Reality report
- Trustworthy digital repositories or a network of such repositories that
  1) follow FAIR principles,
  2) ensure long-term access and
  3) enable long-term preservation of data.

Task Forces: AAI Architecture, PID Policy and Implementation; Long-Term Data Preservation

Quality of Research Repositories

Objective

Define indicators for quality research repositories by developing standards for FAIR certification of repositories and the data within them.

Approach

- Fostering FAIR data principles throughout the research data life cycle is a key goal of EOSC.
- Define certification schema to embed and roll them out, supporting repositories to achieve such status and be a recognised authoritative source of quality data which is preserved for the long-term.

Task Forces: Infrastructure for Quality Software, Long-Term Data Preservation

HE Work Programme 2023-24: draft priorities of the EOSC-A

EOSC TF Long Term Data Preservation

- Recommendations on the creation of a European network of trustworthy digital repositories following FAIR-enabling principles with disciplinary and geographical spread.

- Recommendations for EOSC data services to connect to this network and a roadmap to further mature the LTP aspects of these repository services.

https://www.eosc.eu/advisory-groups/long-term-data-preservation

Thank you for listening!

European Commission - Christian Cuciniello, EC

Trusted repositories and the concept of network of TDR in Horizon Europe

Priorities for Open Science

Improve the practice of research and innovation

- Openly accessible scholarly publications

- Early sharing of research outputs; open methods and resources

- **All digital outputs FAIR, RDM**

- Reproducible results

- Societal engagement & responsibility

Develop proper enablers

- Rewards and incentives to adopt Open Science practices, with responsible use of metrics

- Appropriate skills and education, including for research integrity

- **Open Research Infrastructures** including the European Open Science Cloud (EOSC)

- Legal & regulatory environment for data and copyright

How the EC supports the operationalisation of Open Science

With projects under WIDERA WP:

- open access publishing; reforms of research assessment reproducibility of research results; citizen science

With projects under INFRA WP:

- EOSC implementation à Horizon Europe co-programmed partnership
  - EOSC federation layer; EOSC core and advanced services; FAIR data services and support to adoption and implementation; governance; skills and training; etc. data trusted, virtual, federated environment

With HE programme provisions on OS

With programme services à to "walk the talk" and "lead by example"

1. Open access to scientific publications

Beneficiaries must <u>ensure</u> **OA to peer-reviewed scientific publications** relating to their results. In particular, **they must ensure**:

- **deposition** in a <mark>trusted repository</mark> + **immediate open access via the repository** under **CC BY** or equivalent (CC BY-NC/CC BY-ND are allowed for long-text formats)
- **information** via the repository about any research output/tools/instruments needed to **validate the conclusions of the scientific publication**

Metadata must be open under CC 0 or equivalent, **in line with the FAIR principles** and provide information about the licensing terms and persistent identifiers, amongst others.

Beneficiaries (or authors) must retain **sufficient intellectual property rights** to comply with the OA requirements

Publication fees are reimbursable only if publishing venue is full open access.

Requirements in Horizon Europe
2. Research data management

Beneficiaries **must manage the digital research data in line with the FAIR** principles and:

- establish + regularly update a **data management plan** ('DMP') for generated (and/or collected) data;
- **deposit** the data in a <mark>trusted repository</mark> (federated in the EOSC if required in the call conditions) **+ ensure OA under CC BY, CC 0 or equivalent, following the principle 'as open as possible as closed as necessary'**
- provide information via the repository about any research output/tools/instruments needed to **re-use or validate the data**

Metadata must be open under CC 0 or equivalent (to the extent legitimate interests or constraints are safeguarded), **in line with the FAIR principles** and provide information about the licensing terms and persistent identifiers, amongst others.

Requirements in Horizon Europe
Definition of trusted repository

What is required **to be TRUSTED**

- display specific characteristics of organisational, technical and procedural quality (e.g. secure integrity and authenticity facilitating their use and re-use short- and long-term
- offer explicit information online about their policies, services (e.g. acquisition, access, security of content, long-term sustainability of service including funding etc.).
- provide ideally open access to content free at the point of use
- respect applicable legal and ethical limitations
- assign persistent unique identifiers to contents (e.g. DOIs, handles, etc.)
- ensure metadata sufficiently detailed and of sufficiently high quality following the standards of the respective community
- facilitate mid- and long-term preservation of the deposited material
- have mechanisms or provisions for expert curation and quality assurance for the accuracy and integrity of datasets and metadata
- meet generally accepted criteria for security

HE - RI - INFRAEOSC tentative list of objectives 2023-2024

- **Further consolidation & development of the EOSC Core**, as well as, **extension of the resource offering via EOSC Exchange**
- **Competence centers**, supporting research communities
- Expansion of EOSC's **access to resources provided by public authorities**
- **Cross-articulation** with **sectoral data spaces** & other relevant EU **cloud** & **data** infrastructures
- **Federated network of trusted repositories** for data & other research digital objects
- Development of community-based approaches for ensuring and **improving the quality of scientific software and code**
- Quality of data and long-term **access and preservation** within EOSC
- **FAIR and open data sharing** in support of the mission: **Adaptation to Climate Change**;
- **Planning, tracking and assessing scientific knowledge production** (e.g., science knowledge graphs, data management plans, extension of metrics);
- Support to the **coordination and consolidation of the EOSC Partnership**;
- Build on the science **cluster approach to ensure the uptake of EOSC by all RIs** and research communities.

Next planned service
Multidisciplinary cloud based repository
for research data and other research digital objects

Horizontal Expenditure for HE 2021-2022
- provide a centrally-managed, cloud-based repository platform to offer services at no cost to beneficiaries during and after the end of their grants
- accept research data and other outputs from all scientific disciplines (easily adaptable to new provisions and practices)
- implement a reach **research data management cycle,** including a framework for data curation
- ensure the implementation of the FAIR principles in line with the specificity of the different communities
- be included among the service providers of the EOSC federation and in the network of trusted repositories

→ From the new ERA perspective, an EU repository for research data could play a catalytic role towards a wider, interconnected network of general-purpose research repositories operated by research funders at European and national levels

Thank you for listening!

EOSC LTDP TF Overview

EOSC Long Term Digital Preservation Task Force Overview

Workshop: European network of Trustworthy Digital Repositories (TDRs)

Hervé L'Hours/Roxanne Wyns (LTDP TF Co-Chairs)

2022-01-13

EUROPEAN OPEN SCIENCE CLOUD

EOSC Advisory Groups / Task Forces

Community Consultation on Topics has lead to:

- 5 Advisory Groups containing 13 Task Forces

- Inform Association/others to help steer the implementation of EOSC

- Identify strategic gaps and areas for investment as input to SRIA

- Liaise with EOSC projects and offer feedback and advice

EUROPEAN OPEN SCIENCE CLOUD

Task Force topics

Implementation of EOSC

- Rules of Participation compliance monitoring

- PID policy and implementation

- Researcher engagement and adoption

Technical challenges on EOSC

- Technical interoperability of data and services

- Infrastructure for quality research software

- AAI Architecture

Metadata and data quality

- Semantic interoperability

- FAIR metrics and data quality

Research careers and curricula

- Data stewardship curricula and career paths

- Research careers, recognition and credit

- Upskilling countries to engage in EOSC

Sustaining EOSC

- Defining funding models for EOSC

- Long-term data preservation

Liaisons for questions / approvals

- Implementation of EOSC – Suzanne Dumouchel

- Technical challenges on EOSC – Ignacio Blanquer

- Metadata and data quality – Sarah Jones

- Research careers & curricula – Wilhelm Widmark

- Sustaining EOSC – Bob Jones

Suzanne DUMOUCHEL
Director, three-year mandate

Ignacio BLANQUER
Director, two-year mandate

Sarah JONES
Director, two-year mandate

Wilhelm WIDMARK
Director, one-year mandate

Bob JONES
Director, one-year mandate

EUROPEAN OPEN
SCIENCE CLOUD

Horizon Europe Work Programme: TF Input

- ## Quality of Research Repositories

  Task Forces: Infrastructure for Quality Software, Long-Term Data Preservation

- ## Long-term access to and preservation of data

  Task Forces: AAI Architecture, PID Policy and Implementation; Long-Term Data Preservation

- ## Other Topics

  Quality of Software
  Linking to sensitive data from public authorities
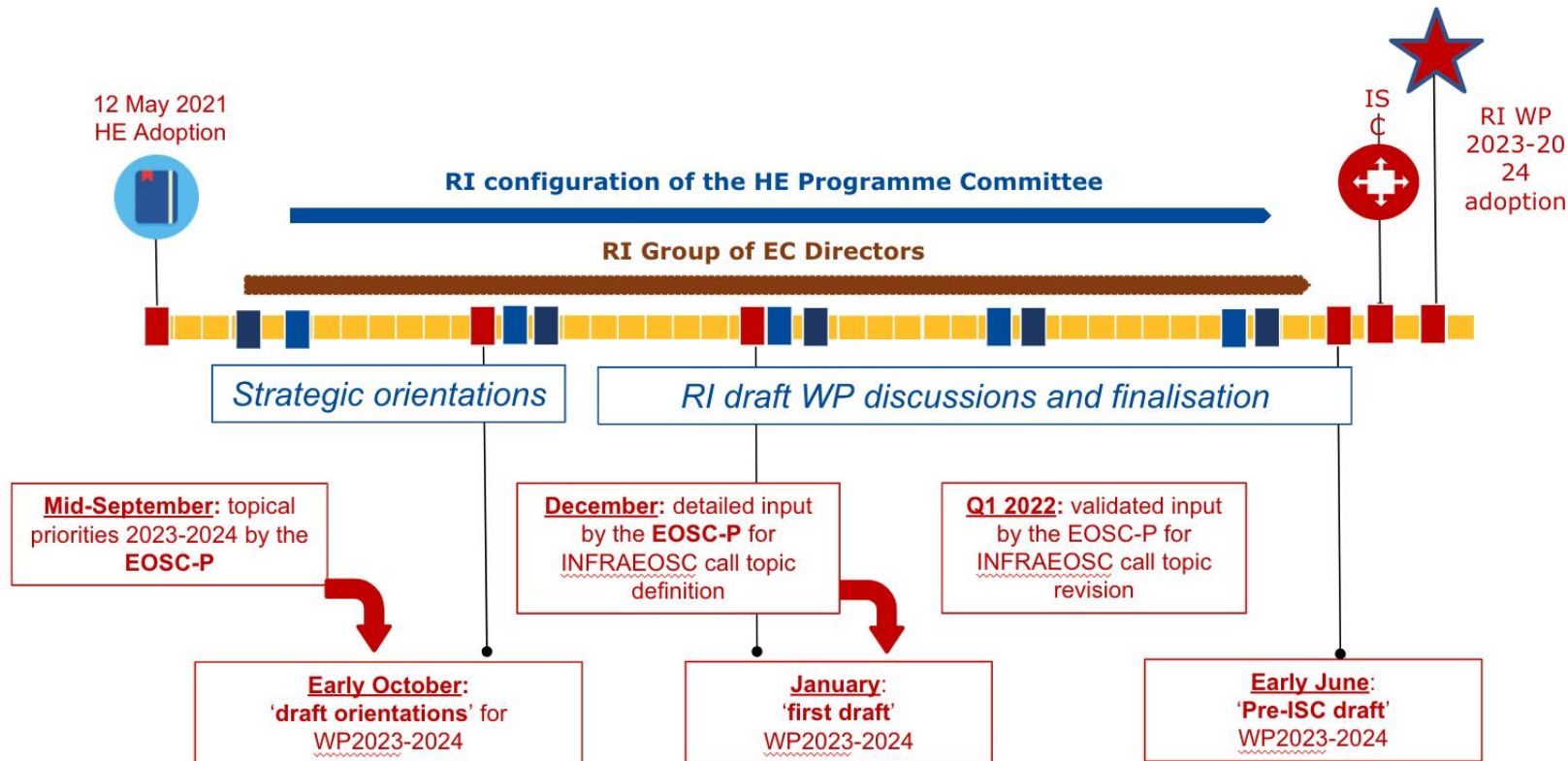 AI and access to un-structured and non-standard data
  Multilingualism in EOSC
  Science Clusters and the long-tail
  Quality of Research Data

EUROPEAN OPEN
SCIENCE CLOUD

Schedule for inputs to the Work Programme

12 May 2021
HE Adoption

IS
C

RI WP 2023-2024 adoption

RI configuration of the HE Programme Committee

RI Group of EC Directors

Strategic orientations

RI draft WP discussions and finalisation

Mid-September: topical priorities 2023-2024 by the **EOSC-P**

December: detailed input by the **EOSC-P** for INFRAEOSC call topic definition

Q1 2022: validated input by the EOSC-P for INFRAEOSC call topic revision

Early October: 'draft orientations' for WP2023-2024

January: 'first draft' WP2023-2024

Early June: 'Pre-ISC draft' WP2023-2024

EUROPEAN OPEN SCIENCE CLOUD

Horizon Europe Work Programme: Next steps

- LTDP EC meetings initiated

- LTDP TF inputs to WP 23/24 provided

- EOSC-Association summarised and shared with the Commission

- Consultation on inputs to the Work Programme in Q1 2022

- Inputs to update SRIA, specifically Multi-annual roadmap in Spring 2022

EUROPEAN OPEN SCIENCE CLOUD

Initial activities to establish LTDP TF

- Arrange initial meeting

- Start process to identify chairs

- Approve the charter with full membership

- Define a comprehensive workplan for the first year

LTDP TF Sub-groups

- Creation of shared understanding and vision

- 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**

  Plus shared activities

Creation of a shared understanding and vision

- Starting from what we mean by digital preservation in the context of EOSC and a mapping of the existing landscape

- Resulting in a suggestion for a strategy, where horizontal EOSC preservation policy enables the connection and collaboration on national, community and local level

- This TF will provide recommendations for the EOSC board on the vision and sustainable implementation of long-term data preservation policies and practices, as well as suggestions to later strategy execution

Mapping of the financial aspects of LTDP

- Mapping the financial aspects of long term data preservation

- The cost of digital preservation, financial responsibilities of the different stakeholders, as well as possible business models for repositories

Mapping and promotion of the roles, responsibilities and accountability

- Mapping and promotion of the roles, responsibilities and accountability of the actors within the EOSC ecosystem with respect to long-term data preservation and responsibility levels defined in the SRIA

- Complementing this by identifying the stakeholders in the different stages of the research data life cycle and their respective roles and responsibilities, including recommendations for awareness raising and training, especially on comprehensive and accountable DMPs

Recommendations trustworthy digital repositories

- Recommendations on the creation of a European network of trustworthy digital repositories following FAIR-enabling principles with disciplinary and geographical spread

- Recommendations for EOSC data services to connect to this network and a roadmap to further mature the LTP aspects of these repository services

Engage

- Projects
- Wider community consultation
- Identify other initiatives and align

EUROPEAN OPEN
SCIENCE CLOUD

Validating / endorsing outputs

- Consensus within Task Force - discussion across sub-groups
- Consultation with projects and wider community
- Solicit inputs from EOSC Association Membership
- Consultation across Task Forces via Chairs forum
- Proposal to Board via liaison

Draft TF Roadmap

ROADMAP FOR EOSC LTP TF

	4Q2021	1Q2022	2Q2022	3Q2022	4Q2022	1Q2023	2Q2023
Shared activities	-definition of main concepts - compilation of related work - Landscape of repositories (from existing information)	Communicating with relevant TFs and initiatives	Communicating with relevant TFs and initiatives	Communicating with relevant TFs and initiatives	Communicating with audience	Communicating with relevant TFs and initiatives	Shared understanding and vision
Creation of shared understanding and vision		Compiling existing EU level, national and sectoral LTP policies and legal constraints	Drafting first version of EOSC EOSC LTP network strategy and policy (including appraisal, curation, retention periods etc.)		Second version of strategy and policy, including a roadmap to mature repositories in the network	Strategy, policy and roadmap to comments	Final strategy, policy, roadmap
Mapping the financial aspects of long-term data preservation	Liaison with financial sustainability TF	Compilation of existing work on financial aspects	First compilation of costs and financial responsibilities I EOSC context	Suggestions for EOSC LTDP business models	Recommendations to comments	Second compilation of costs and financial responsibilities I EOSC context	Incorporating financial aspects to vision, strategy, policy roadmap
Mapping and promotion of the roles, responsibilities and accountability	Liaison with relevant TFs and stakeholders	Identifying the actors within the EOSC ecosystem with respect to long-term data preservation and responsibility levels	Identifying the stakeholders in the different stages of the research data life cycle and their respective roles and responsibilities	Cross-walk on DMPs and digital preservation plans	Recommendations for awareness raising and training, especially on comprehensive and accountable DMPs	Recommendations to comments	Incorporating recommendations to vision, strategy, policy roadmap
Recommendations on the creation of a European network of trustworthy digital repositories and for EOSC data services	Liaison with relevant stakeholders	- drafting use cases EOSC LTDP network (a la DICE) - assessing FAIR-enabling repository maturity/certification	- analysing gaps and missing functionalities	- First version of recommendations on the creation of a European network of FAIR-enabling trustworthy digital repositories	- Recommendations on connecting this network to data services within the EOSC ecosystem and a roadmap to mature these repositories.	Recommendations to comments	Incorporating recommendations to vision, strategy, policy roadmap

EUROPEAN OPEN SCIENCE CLOUD

Thank You

Recommendations trustworthy digital repositories

- Recommendations on the creation of a European network of trustworthy digital repositories following FAIR-enabling principles with disciplinary and geographical spread

- Recommendations for EOSC data services to connect to this network and a roadmap to further mature the LTP aspects of these repository services

Overview of repository support approaches

Towards making your repository more trustworthy and FAIR-enabling

Linas Čepinskas, Policy Officer, DANS (The Netherlands)

13 January 2022

FAIRsFAIR - Fostering FAIR data practices in Europe

- Goal: Practical solutions for the use of FAIR principles throughout the research data life cycle

- Budget: €10 million

- Duration: 2019-2022

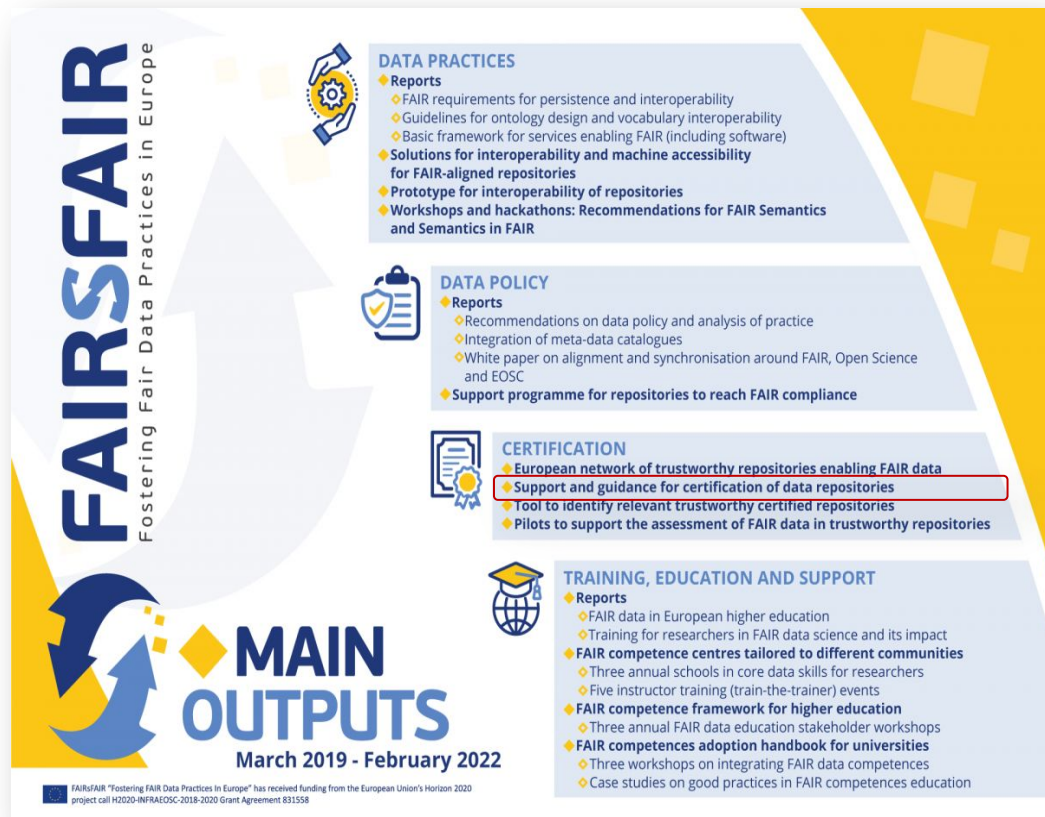- 22 partners from 8 EU member states

- Project coordinator: DANS

Image source: https://www.fairsfair.eu

A journey towards becoming trustworthy and FAIR-enabling

Support with CoreTrustSeal certification

Research data repositories

Trustworthy repository practices and lessons learnt

CoreTrustSeal repository certification programme

Repository support: services

5 support workshops

Test reviews

Practical tools

1:1 support and advice

Who we work with

→ a country with a selected repository

Growing together

Trust and FAIR

Global repository support initiatives, e.g.

- 'Enabling FAIR Data CoreTrustSeal cohort'
- ARDC, Australia
- NDRIO-Portage Network, Canada

Share experiences and lessons learned

Explore collaboration
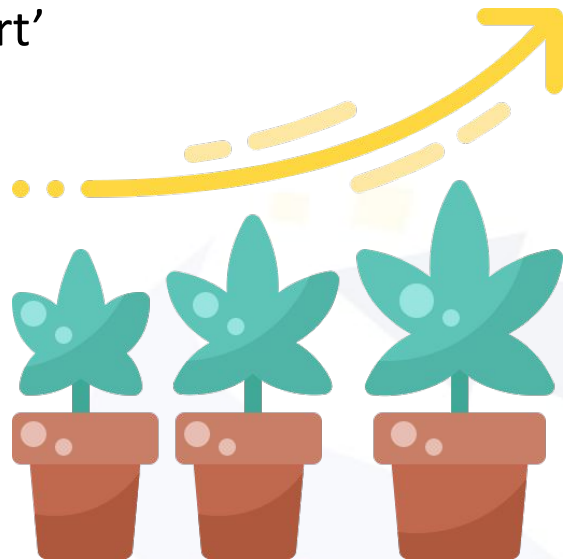
Key takeaways

- Great diversity in terms of **geography**, **domains** and **maturity** of participating repositories.
- **9** out of **10** repositories have submitted their **applications to CoreTrustSeal** and the other repository is on its way to do so.
- **1-on-1 support** and a **hybrid approach** are appreciated the most.
- Taking into account the **starting point** of each **repository** is key.
- Initiatives should take note of which CoreTrustSeal aspects are found **most difficult** and dedicate **more support** to those areas.
- The **support** is **most useful** when it **mimics the actual certification process** as closely as possible.

Recommendations for developing a support programme

Broad engagement within applicant organisations can minimise the effort of preparing self-assessments while maximising internal communication and knowledge exchange.

In setting up a support team, include members with a **variety of backgrounds and disciplinary experience** to provide high-quality support.

After the introductory information stage, consider **dividing support work** across the supported repositories based on their **experience** of CoreTrustSeal, their **maturity** as an organisation and the **outcomes** they have identified as key goals of the process.

Design the support programme with **different support routes** and **engagement types**. Identify and divide content into specific areas of focus. This enables targeted learning and the revision and re-use of support content so that it is optimised for different approaches.

Further information

Practical tools

- Hervé L'Hours, & Linas Cepinskas. (2021). Policy and Evidence Planning for Data Services: Business Information Management. Zenodo. https://doi.org/10.5281/zenodo.5779791
- Herterich, Patricia. (2020, April 6). FAIRsFAIR support towards achieving CoreTrustSeal certification - roadmapping exercise. Zenodo. http://doi.org/10.5281/zenodo.3741693
- FAIR-Aware tool to assess knowledge of FAIR (en français aussi)
- F-UJI tool to assess FAIRness of datasets
- FAIR Data Forum. A gateway to FAIR data practices

Certification support programme

- Ilona von Stein, Olivier Rouchon, Maaike Verburg, Serenella Muradore Gallas, Sara Pittonet Gaiarin, & Andrea Greco. (2021). Making your repository more trustworthy and FAIR-enabling a support programme. Zenodo. https://doi.org/10.5281/zenodo.5106165
- Hervé L'Hours, Ilona von Stein, Jerry deVries, Linas Cepinskas, Joy Davidson, Patricia Herterich, Robert Huber, & Benjamin Jacob Mathers. (2021). M4.3 CoreTrustSeal+FAIRenabling, Capability and Maturity (1.0). Zenodo. https://doi.org/10.5281/zenodo.5346822
- Maaike Verburg, Ilona von Stein, Linas Cepinskas, Hervé L'Hours, Patricia Herterich, Joy Davidson, Kevin Ashley, Olivier Rouchon, Andrea Greco, Serenella Muradore Gallas, & Sara Pittonet Gaiarin. (2021). D4.3 Report on the certification support and guidance for repositories and reviewers. Zenodo. https://doi.org/10.5281/zenodo.5137552
- FAIRsFAIR support programme for repositories https://www.fairsfair.eu/application-results-open-call-data-repositories
- Frans Huigen and Ilona von Stein (8 October 2020). How is FAIRsFAIR supporting the repositories selected for CoreTrustSeal Certification?https://www.fairsfair.eu/articles-publications/how-fairsfair-supporting-repositories-selected-coretrustseal-certification

Promoting trust through repository support

A brief look at SSHOC Certification support

Henri Ala-Lahti, Information Specialist,
Finnish Social Science Data Archive

**Towards a network of
European FAIR-enabling
Trustworthy Digital Repositories**

13 January 2022
Online workshop

SSHOC
social sciences & humanities open cloud

Project:

SSHOC
social sciences & humanities open cloud

Horizon 2020
European Union Funding
for Research & Innovation

Type of action & funding:
Research and Innovation action
(INFRAEOSC-04-2018)

Partners: 45
(20 beneficiaries + 25 LTPs)

SSH ESFRI Landmarks and Projects
& international SSH data infrastructures

Project budget:
€ 14,455,594.08

Project website:
www.SSHopencloud.eu

Duration: 40 months
(January 2019 – 30 April 2022)

Objectives:

• creating the social sciences and humanities **(SSH)** part of European Open Science Cloud **(EOSC)**

• maximising **re-use** through **Open Science** and **FAIR** principles (standards, common catalogue, access control, semantic techniques, training)

• interconnecting existing and new infrastructures (clustered cloud infrastructure)

• establishing appropriate **governance model** for SSH-EOSC

SSHOC Task 8.2 Trust & Quality Assurance

Purpose: Bolstering and improving trust in and quality of the SSHOC repositories by supporting them and by exploring the trust landscape and providing feedback to certification bodies.

Team:
- CESSDA/FSD
- CESSDA/SND
- CESSDA/UKDS
- KNAW/DANS
- CLARIN ERIC
- CNR/E-RIHS
- DARIAH/PNSC

Supported repositories

- CLARIN-LV
- Corpus OVI dell'italiano antico
- Croatian Social Science Data Archive (CROSSDA)
- DAIS - Digital Archive of the Serbian Academy of Sciences and Arts
- DARIAH-DE Repository
- Digital Library of University of Maribor
- Digital Repository of Scientific Institutes
- Historic Graves
- Lithuanian Data Archive for Social Sciences and Humanities (LiDA)
- mdw Repository
- NAKALA
- SB/CLARIN Repository (Språkbanken Text CLARIN Repository)
- Sciences Po, Center for Socio-Political Data (CDSP)
- Slovak Archive of Social Data (SASD)

SSHOC
social sciences & humanities open cloud

Support activities

- Online webinars and workshops on certification
- Regular one-on-one meetings with repository representatives and reviewers assigned to them from T8.2
- Comments and feedback on self-assessments
- Webpage for repositories seeking certification

SSHOC
social sciences & humanities open cloud

Key takeaways

- Result: facilitation of the certification process, increased understanding of what is required for certification

- Result: improved repository practices and documentation through self-assessments

- Result: feedback from repositories to improve certification

- Varying repository types, practices and levels of maturity

- Notable differences in resources available

- Need for increased cooperation and peer-to-peer networks not only in terms of certification but also other aspects of repository practices

SSHOC
social sciences & humanities open cloud

Further information

- Mari Kleemola, Tuomas J. Alaterä, Niko Koski, Henri Ala-Lahti, Birger Jerlehag, Hervé L'Hours, Franciska De Jong, Dieter Van Uytvanck, Tomasz Parkoła, Emiliano Degl'Innocenti, Roberta Giacomi, Maurizio Sanesi & René Van Horik. (2020). *SSHOC D8.2 Certification plan for SSHOC repositories* (v1.0). Zenodo. https://doi.org/10.5281/zenodo.4558303

- SSHOC Trusted Repositories: https://sshopencloud.eu/sshoc-trusted-repositories

- Mari Kleemola et al. (forthcoming) *SSHOC D8.3 Trustworthy Digital Repository status update and certification solutions for SSHOC repositories*

SSHOC
social sciences & humanities open cloud

Thank you for your attention!

Join our community

https://www.sshopencloud.eu

@SSHOpenCloud

info@sshopencloud.eu

/in/sshopencloud

SSHOC
social sciences & humanities open cloud
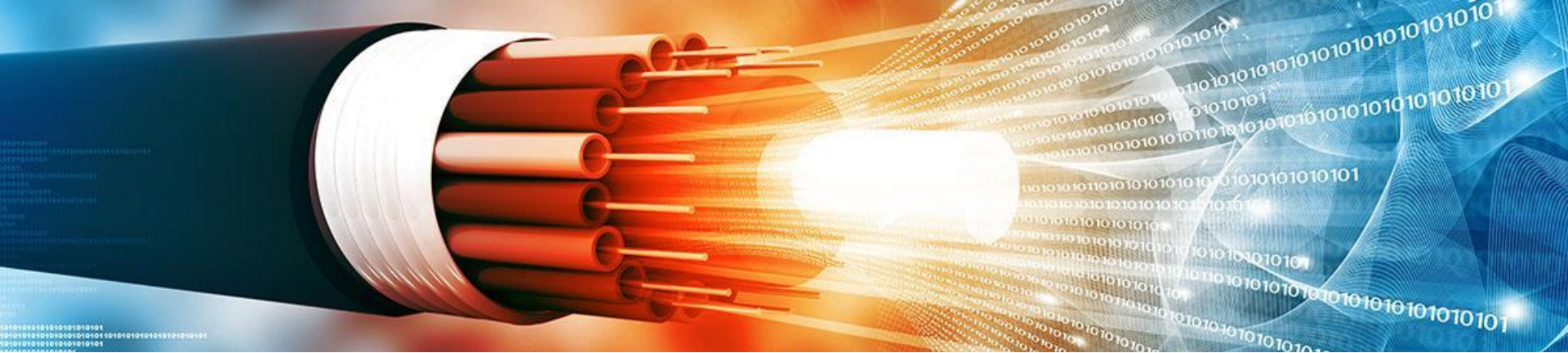
Supporting the adoption of FAIR data standards and certification: Case EOSC-Nordic

Tuomas J. Alaterä, FSD / Tampere University

Towards a network of European FAIR-enabling Trustworthy Digital Repositories

January 13, 2022

EOSC-Nordic 9/2019 - 8/2022

SWEDEN
- SWEDISH NATIONAL INFRASTRUCTURE FOR COMPUTING (UNIVERSITY OF UPPSALA)
- SWEDISH RESEARCH COUNCIL
- UNIVERSITY OF GOTHENBURG

ICELAND
- UNIVERSITY OF ICELAND

NORWAY
- NEIC (NORDFORSK)
- UNINETT SIGMA2
- NORWEGIAN CENTRE FOR RESEARCH DATA

FINLAND
- CSC - IT CENTER FOR SCIENCE
- UNIVERSITY OF HELSINKI
- TAMPERE UNIVERSITY
- UNIVERSITY OF EASTERN FINLAND
- FINNISH METEOROLOGICAL INSTITUTE

DENMARK
- DEIC (DTU)
- COMPUTEROME (DTU)
- THE UNIVERSITY OF SOUTHERN DENMARK
- DANISH NATIONAL ARCHIVES
- UNIVERSITY OF COPENHAGEN
- CAPITOL REGION OF DENMARK
- NORDUNET AS

ESTONIA
- UNIVERSITY OF TARTU
- NATIONAL INSTITUTE OF CHEMICAL PHYSICS AND BIOPHYSICS

LATVIA
- RIGA TECHNICAL UNIVERSITY

LITHUANIA
- UNIVERSITY OF VILNIUS

THE NETHERLANDS
- GO FAIR FOUNDATION

GERMANY
- GERMAN CLIMATE COMPUTING CENTRE

- Regional EOSC implementation project
  - Collaboration with FsF
  - Nordic-Baltic testbed

- Cross-disciplinary

- 6 work-packages, 24 partners from 10 countries

- WP4 on FAIR Data and certification support

EOSC-Nordic WP4 Goal

FAIR digital objects in Trustworthy digital repositories

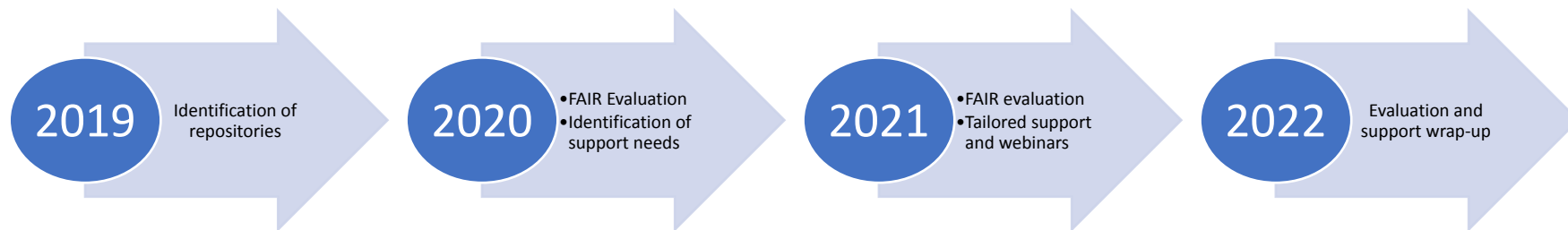Automated FAIR evaluation, FAIRicifation support, identifying incentives for FAIR uptake

CoreTrustSeal support and repository specific guidance for certification, FAIR certification feedback

Support timeline

2019 — Identification of repositories

2020
- FAIR Evaluation
- Identification of support needs

2021
- FAIR evaluation
- Tailored support and webinars

2022 — Evaluation and support wrap-up

- Initially identified 98 evaluation / support candidates
- Automated FAIR evaluation of the most continues
- Based on invitation and interest, 13 repositories selected as support recipients

Bolin Centre Database, Sweden
DataDOI, Estonia
DTU, Denmark
Finnish Biodiversity Information Facility (FinBIF)
Icelandic Social Science Data Service (DATICE)
Lithuanian Data Archive for Social Sciences and Humanities (LiDA)

Music Archive Finland
NIRD Research Data Archive, Norway
QSAR DataBase, Estonia
SU, Sweden
AIDA Data Hub, Sweden
Metabolic Atlas, Sweden
USN, Norway

Support activities provided

- Conversational webinars on CTS & "FAIRification Steps" webinar series
- Goal setting for each repository (certification, self-assessment, FAIRification, or both)
- Regular/themed support meetings of dedicated WP4 experts and repository representatives
- Feedback on self-assessments (google docs, email)
- Knowledge hub for participants

Key takeaways (with 8 months to go)

- Supporting both FAIRification and certification not always simple and mutually productive
  - Technology dependence, knowledge base, requirement
- Interest often exceeds resources available
  - Need for support within the infrastructures / countries
- Notable progress in documentation, cooperation and understanding of certification concepts via self-assessments and support calls
- CTS and FAIR applicable to various disciplines - easier to see and implement when networking with various actors

- What is a "repository" still a bit of a mystery!

Further information

- Tuomas J. Alaterä: *Takeaways from the EOSC-Nordic WP4 Workshop "From Self-Assessment to Certification with FAIR Results"* (2021). https://www.eosc-nordic.eu/takeaways-from-the-eosc-nordic-wp4-workshop-from-self-assessment-to-certification-with-fair-results/
- Andreas Ortmann Jaunsen, Mari Kleemola, Tuomas J. Alaterä, Heikki Lehvaslaiho, Adil Hasan, Josefine Nordling, & Pauli Assinen. (2020). D4.1 *An assessment of FAIR-uptake among regional digital repositories* (1.0). Zenodo. https://doi.org/10.5281/zenodo.4045402

www.eosc-nordic.eu | twitter.com/EOSC_Nordic | YouTube

Q&A

Towards a network of European FAIR-enabling TDRs

Photo by Eileen Pan on Unsplash

Scope **Today**

- initial network of existing and aspiring Trustworthy Digital **Repositories** (TDRs)
- increase the number of FAIR-enabling TDRs
- initial focus on **Europe** and **EOSC**

Scope **Later**

- wider network of existing and aspiring trustworthy **data services**
- increase the number of FAIR-enabling trustworthy data services
- focus reach **beyond Europe** on other federated research (meta)data infrastructures

FAIRsFAIR D4.4 Coordination Plan for a sustainable **network of FAIR-enabling Trustworthy Digital Repositories**
https://doi.org/10.5281/zenodo.5726691

Discussions

Menti.com

Please enter code **3208 0926**

Which of these options best describes your background?

FAIRsFAIR team — 7
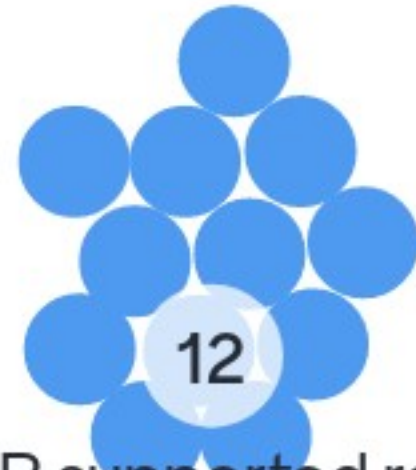
FAIRsFAIR supported repository — 12

SSHOC team — 6
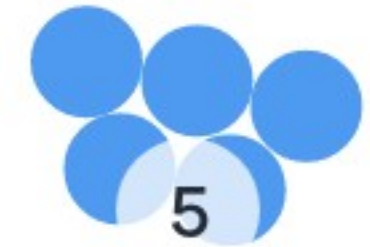
SSHOC supported repository — 5
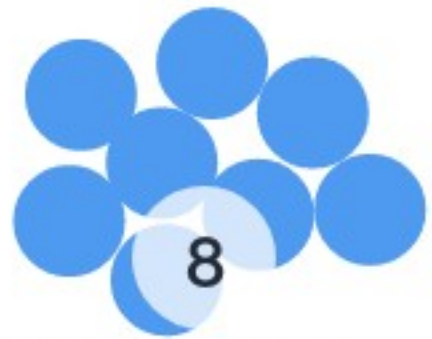
EOSC-Nordic team — 8

EOSC-Nordic supported repository — 3

European Commission — 2

European Certified Repositories — 12

CoreTrustSeal Board — 1

EOSC Task Force Long Term Data Preservation — 2

EOSC Executive Board — 0

Nestor Working Group — 2

In your perspective, which other stakeholders should be involved?

RDA

Research funders

OpenAire

Users

Clarin

Funders

RDA

Advisory involvement from funders

RDA

79

In your perspective, which other stakeholders should be involved?

COAR

RDA / RDA Europe

RDA

Research infrastructures

Funders

RDA

Maybe ERICs? (e.g. DARIAH)

funders

RDA

79

In your perspective, which other stakeholders should be involved?

RDA	Funders and Research organisations.	NRENs
RDA	GLAM repositories	Wider data services across the lifecycle
Funders	OpenAIRE	Research councils

79

In your perspective, which other stakeholders should be involved?

nestor

COAR

Disciplinary networks such as Elixir

RDA

users

OpenAire

policy makers

EOSC Board of Directors

Higher Education Institutes

79

In your perspective, which other stakeholders should be involved?

funders

national policy makers and funders

National networks such as the German National Researchdata infrastructure (NFDI)

RDA

RDA

LIBER

other cross-disciplinary Cluster Projects, if TDR network ist multidisciplinary with different certifications

Researchers

ERICs

In your perspective, which other stakeholders should be involved?

Funders

ESFRI

individual research infrastructures

digital preservation experts, e.g. OPF, dpc, nestor

All domain repository maintainers and instance managers, such as CLARIN.

ERICs

Funders

Reseach infrastructures

Senior researchers with data reuse experience, who can give feedback on how they perceive the repo quality and support

In your perspective, which other stakeholders should be involved?

Registries	Other repository membership organisations	Research funders
Digital Preservation organisations (e.g. DPC)	RDA / RDA Europe	Research infrastructures
RDA	(relevant) ICT standardisation organisations	Scientific Unions such as IUPAC

In your perspective, which other stakeholders should be involved?

Research funders

Research projects, researchers, universities

DG CNECT

Traditional archives

depends on the aim of the network

Archives

Digital Preservation networks

All ERICs (e.g. CLARIN, DARIAH, ...), National Initiatives (e.g. NFDI in Germany)

Depends an the goals of such a network

79

In your perspective, which other stakeholders should be involved?

Technology or platform providers

research infrastructures

DiSSCo members (Distributed System of Scientific Collections) - European RI under development (https://www.dissco.eu/)

National EOSCs (like the Nordic one, there are being created in other countries, typically one per country.

NFDI (Germany)

Funders

NRENs

79

Would you want a European TDR network to come into existence?
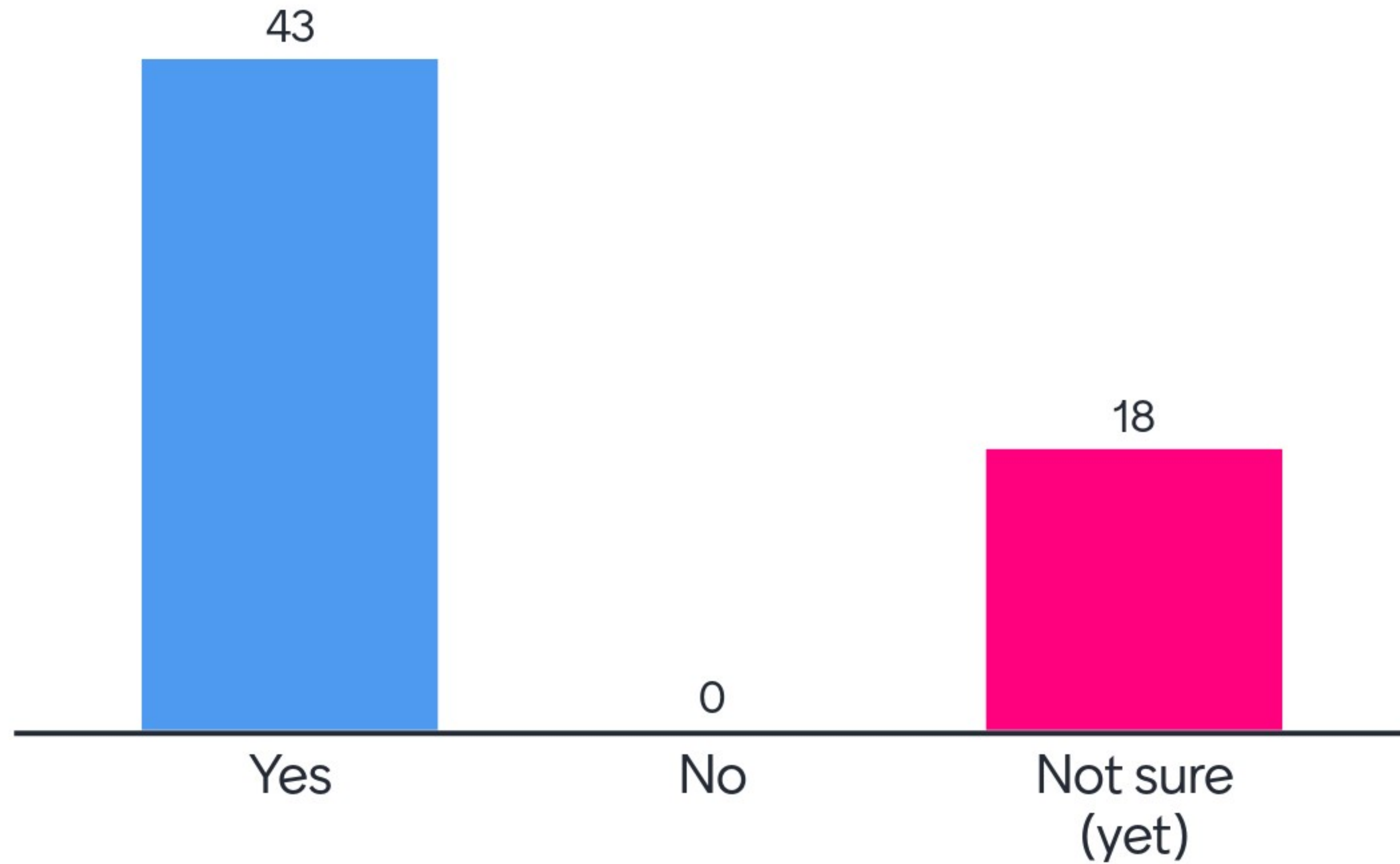
In your opinion, what should be the possible goals of such a network?

Sharing experiences

Advocacy

Knowledge exchange

Standards

Developing best practices

A voice for Trusted Data Repositories at higher Policy Levels

Communication

Share best practices and tips

Peer-to-Peer support

89

In your opinion, what should be the possible goals of such a network?

Promoting the standardisation of approach	Data backups across repos	Sharing knowledge and best practices
Best Practice	Best practice share	Common service components when possible
Support repositories in enabling FAIR data sharing/publishing/archiving	Mutual Support	Advocacy

In your opinion, what should be the possible goals of such a network?

To reduce duplication of effort	Alignment	Standardisation of tools and services
Making the case for long-term sustainable funding for Trusted Data Repositories	Coordination, standardization	Common voice
Improve Open Science	Best practice	Share knowledge/experience

89

In your opinion, what should be the possible goals of such a network?

Standards	Peer to peer support	Exchange of experiences. knowledge and common challenges
Knowledge sharing	Succession / Fall-back plans	Common practices and standards
Best pracices	Enhanced findability of resources	mutual support and information exchange

89

In your opinion, what should be the possible goals of such a network?

Acting as a voice for all relevant repositories - institutional, disciplinary, etc

Support, Sharing of expertise, evolution of best practice

Peer-to-peer support and enforcing best practices.

Advocacy for Long Term Preservation

Goal is to link the distributed / federated repositories into one environment.

Facilitate data interchangePromote standards and common approachesDevelop methods

Definition of interfaces to allow the use of data across institutions.

Harvesting posibilites on metadata level

Contribute to standard development

89

In your opinion, what should be the possible goals of such a network?

Development of common tools to support the curation of data & metadata

Developing best practice guidelines

Best practices

Share knowledge, best practices, support

Sharing experiences

share our practices and learn from others

Collaborating for funding

To advocate for TDRs at the EU levelTo jointly apply for funding as a networkTo promote TDRsTo provide peer-to-peer support

Best practice

89

In your opinion, what should be the possible goals of such a network?

Connection point for Global Collaboration and Knowledge Sharing

Avoid overlap and harness synergies

Promote preservation specialist training.

Stronger push towards platform developers to note requirements of certification and FAIR

Essential to have a federated core of data on which to build eosc

Awareness of the differences between storage, retention, curation and preservation across the lifecycle and data services

Should not be about providing "back up" across repositories -- too many IP/legal issues

Harmonisation of FAIR practices

Not a technical infrastructure

89

In your opinion, what should be the possible goals of such a network?

Exchange of information, best practices. Very careful about metadata aggregation & central search - but a possible goal, too.

Are we talking a technical network?

Distributed repositories that provide their metadata/data in a standardised way using user licenses like CC-BY

Raising digital preservation awareness

Shared best practices, consistency/standardised technologies, metadata standards, etc.

Starting point for new repositories

Establishing recommendations, best practices, standards. Maybe important to promote technical solutions to be interoperable, like Linked Data.

Best practises + How to guarantee sustainability of repository.

How to assure data quality

89

In your opinion, what should be the possible goals of such a network?

Community of practice

Sharing and support for metadata standards

How to create repository with useful data!

Advocacy

Connect regional and disciplinary initiatives that are already providing support for their communities - cultivate new networks where there is a need

Stronger voice at European tables

cross disciplinary exchange

Support

Technical interoperability

89

In your opinion, what should be the possible goals of such a network?

Funding

Preservation

For Researchers - Standardised, multilingual and interconnected discovery layers

Promotion of Open Science to researchers

Advocacy of economic impact (and ROI) for data sharing

Promoting the value of Trusted Data Repositories to the wider research community

Allignment of community standards

All of the previously mentioned

89

What are the biggest challenges to progress towards a network of Trustworthy Digital Repositories?

Funding	Lack of consistency in aims	Agree on a common vision
Time	Money	Communications
Complexity	Co-ordination, differences of requirements from different communities	create added value to the individual repositories without diverting their own ressources

55

What are the biggest challenges to progress towards a network of Trustworthy Digital Repositories?

Time and resources.

Lack of time exceeds the amount of interest

Many actors

money and time

Interoperability

Common vision

Agreement as to what the core purpose of the network is

need to find a coordinator

Too diverse goals

55

What are the biggest challenges to progress towards a network of Trustworthy Digital Repositories?

Funding

Generalist vs Specialist visions of repository roles

money, time, coordination

maintenance of trustworthiness (which costs time, money and diligence)

Different standards

lack of resources

Funding

Geographical scope - since some of the purposes are global needs not just European ones

different needs in different communities/disciplines

55

What are the biggest challenges to progress towards a network of Trustworthy Digital Repositories?

Research funders not aligned

Lack of funding of individuals from the repositories who would be working actively on setting it up

Coordination

Varying resources, maturity levels and interest. Involvement of all parties that would benefit from it.

Time for active participation

Clarity of purpose

Understanding of different domains and disciplines

Common vision and goals

There are too many repositories types

55

What are the biggest challenges to progress towards a network of Trustworthy Digital Repositories?

Define clear value, goals, infrastructure, funding and a plan to implement it.

clear goals

Tendency to only network within a discipline

Clear articulation of the benefits for repository stakeholders.

Finding the vision/mission

Agreement on overall aims - is this intended to be a high-level advocacy network or a more technical support network?

Networking or competing? We cannot exclude this totally.

Information security

Geographical scope (inclusion of non-EU countries)

55

What are the biggest challenges to progress towards a network of Trustworthy Digital Repositories?

What about repositories without long-term preservation mission? They would need to be part of the network

Lack of a common understanding on what trustworthy means

Engaging community of repositories, showing the added value

Time available to do the work, and if not funded be allowed to use inhouse funds.

need specifics rather than generics?

Funding at concrete, participating repositories.

Organising takes serious time

A clear way to harvest institutional gains from a sheared vision

Different technical, organizational and best practices approaches in different disciplines, which arise naturally from the character and structure of the domain-specific data.

What are the biggest challenges to progress towards a network of Trustworthy Digital Repositories?

Assuring the federated structure .. Limit the central functionality to serve the existing repositories.

What other repository community efforts are you familiar with?

COAR	CESSDA	CoreTrustSeal
nestor	CESSDA	CLARIN, NFDIs
CESSDA	CLARIN	Dutch Digital Heritage Network

48

What other repository community efforts are you familiar with?

CLARIN

WDS

COAR

Many ERICs

Portage

A number in GLAM world (can't remember names)

National initiatives

CoreTrustSeal

Re3data.org

48

What other repository community efforts are you familiar with?

DINI	COAR	Clarin
ERICs like CLARIN, CESSDA, DARIAH	CESSDA	CMEMS
NFDI	EIFL	WDS

48

What other repository community efforts are you familiar with?

Existing user groups - particularly those supporting open source tools

CoreTrustSeal

CMEMS

OPF

COAR

CLARIN, DARIAH

Core PID providers

WDS

CoreTrustSeal Community

48

What other repository community efforts are you familiar with?

IOC/IODE Quality Managment Framework	Open Preservation Foundation	WDS
DPC	ENVRI-FAIR	ELIXIR
Digital Preservation Coalition	DCC	CLARIN

48

What other repository community efforts are you familiar with?

DPC

RDA

CLARIN ERIC overall approach (certification, two step, plus best practices for all of language resources community). Also, national EOSC initiative trying to schedule the best roadmap for common repository network in the Czech Republic.

48

What might be the next logical step to take?

Lunch...	Who would coordinate after FAIRsFAIR?	Propose aims and goals
Shout from the rooftops	Create a project team ..	Set out options, present, discuss and seek consensus
Work on a vision	Investigate funding opportinities	Approach developing the shared vision

34

What might be the next logical step to take?

Work on the purpose of the network	mailing list or discussion forum	Strategy
Establish working group	Present team at RDA	Clarify if this is intended to provide high-level advocacy or low level support
Establish a coordination team	Project management approach	set the goal/common vision

What might be the next logical step to take?

Perhaps start small with some initial hubs in the network, from the coalition of the willing (and able).

Get someone to provide seed corn funding

Connect to similar efforts in other parts of the world

Help repositories to improve, even without certification. A sort of light certification, maybe using some automated tools (here FAIR can help)

Establish what not to do because others are already doing it

Circulate the list to potential participants to express their interest to take active participation in the network

Identify funding source

Look at existing (technical) interfaces and APIs for existing solutions

Establish a discussion forum to identify Terms of Reference for such a network.

What might be the next logical step to take?

A shared vision of a smaller group that can then be opened for wider discussion

Papers/Workshops at iDCC and iPRES

Propose a European project and find funding, in order to organize the efforts.

Form an RDA group

Keep participants of this session in the loop

There seems to be interest in the network. Charting the willingness and ability to commit could be next as well as considering the main core and purpose of the network

Perhaps try to influence the HE WP 2023-24 infra calls to include a call for a repo networking activity…

Wrap-up

Thank you!

For further questions: ingrid.dillo@dans.knaw.nl