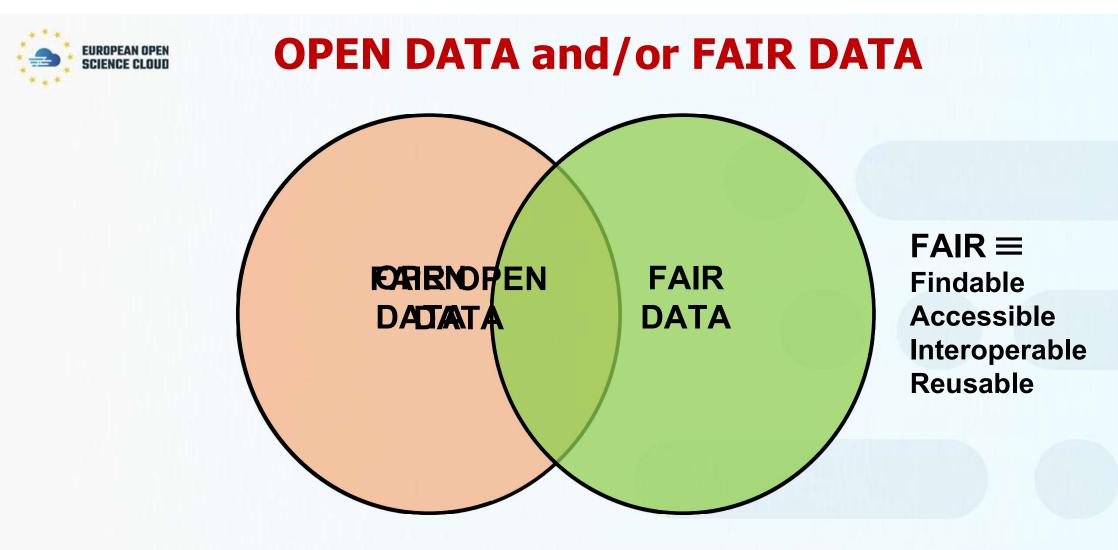


# EOSC-A suggestions for the EURAMET call on Metrology support for digital transformation

### Karel Luyben President EOSC Association

Open Consultation on Metrology for Digital Transformation, 9th November 2021

EOSC Association | www.eosc.eu



Towards "as FAIR as possible" and "as open as possible"



#### DIGITAL OBJECT Data, code and other research outputs

At its most basic level, data or code is a bitstream or binary sequence. For this to have meaning and to be FAIR, it needs to be represented in standard formats and be accompanied by Persistent Identifiers (PIDs), metadata and documentation. These layers of meaning enrich the object and enable reuse.

### IDENTIFIERS

### Persistent and unique (PIDs)

Digital Objects should be assigned a unique and persistent identifier such as a DOI or URN. This enables stable links to the object and supports citation and reuse to be tracked. Identifiers should also be applied to other related concepts such as the data authors (ORCIDs), projects (RAIDs), funders and associated research resources (RRIDs).

### STANDARDS & CODE

### Open, documented formats

Digital Objects should be represented in common and ideally open file formats. This enables others to reuse them as the format is in widespread use and software is available to read the files. Open and well-documented formats are easier to preserve. Data also need to be accompanied by the code use to process and analyse the data.

### METADATA

### **Contextual documentation**

In order for Digital Objects to be assessable and reusable, they should be accompanied by sufficient metadata and documentation. Basic metadata will enable data discovery, but much richer information and provenance is required to understand how, why, when and by whom the objects were created. To enable the broadest reuse, they should be accompanied by a plurality of relevant attributes and a clear and accessible usage license.

# SCIENCE CLOUD

Building the EOSC ecosystem collaboratively with all stakeholders through the EOSC Partnership Enable interdisciplinary research to address societal challenges

Reduce fragmentation by federating existing research infrastructures

> Develop a Web of FAIR Data and Services (including publications and software)

Support Open Science and contribute to the Digital Single Market

Offer EU researchers the digital resources they need to practise Open Science

> Stimulate the emergence of a competitive EU cloud sector

Give Europe a global lead in research data management



# "A web of scientific insight"

- Web of FAIR Data and related Services
- Federation of relevant existing and future data sources
- Virtual space where science producers and consumers come together
- An open-ended range of content and services
- Meeting all European data requirements
- In interaction with other regions of the world





## Twinning the data- to the e-infrastructures

EOSC is a data-infrastructure and could be seen as a twin sister (or brother) of the European e-infra-structure organisations (yin/yang). The last offering the store, compute and connect services used by EOSC to offer the servicing of data and creating interoperability. The combination forms the EOSC-ecosystem





# **Guiding principles for EOSC**

The overarching principle for developing EOSC is that research has to be at the centre of the EOSC initiative.

Multi-stakeholderism

EOSC will succeed if and only if it follows a multi-stakeholder approach;

Openness

EOSC will ensure research artefacts be 'as open as possible, as closed as necessary';

### FAIR principles

EOSC research artefacts need to be findable, accessible, interoperable and reusable;

Federation of infrastructures

EOSC will federate existing and upcoming data- and e-infrastructures;

Machine-actionable

EOSC will strike the right balance between machines and people in delivering the services that will serve the needs of European scientists.



## **EOSC** Association

# An Exciting Journey



## **EOSC Association: Mission**

Advancing the European Open Science Cloud to accelerate the creation of new knowledge, inspire education, spur innovation and promote accessibility and transparency

- To provide a single voice for advocacy and representation for the broader EOSC stakeholder community in Europe
- To promote the alignment of European Union research policy and priorities with activities coordinated by the Association (SRIA)
- To ultimately enable seamless access to data through interoperable services that address the entire research data life cycle, from discovery to storage, management, analysis and re-use across borders and scientific disciplines



# **EOSC Association: Milestones**

- Four founding members (CESAER, GEANT, GARR, CSIC)
- Was incorporated as AISBL on Wednesday 29<sup>th</sup> July 2020
- Obtained Royal Decree on Friday 11<sup>th</sup> September 2020
- First General Assembly on 17-12-2020 elected President and Board
- Research Performing; Research Funding and Service Proving organisations
- Now ~ 150 members and ~ 60 observers (62% 8% 30%) (May 2021)
- A European Co-programmed Partnership, between the EC and the EOSC Association, MoU to be signed 23 June 2021
- Joining the EOSC Association = Joining the EOSC Partnership!



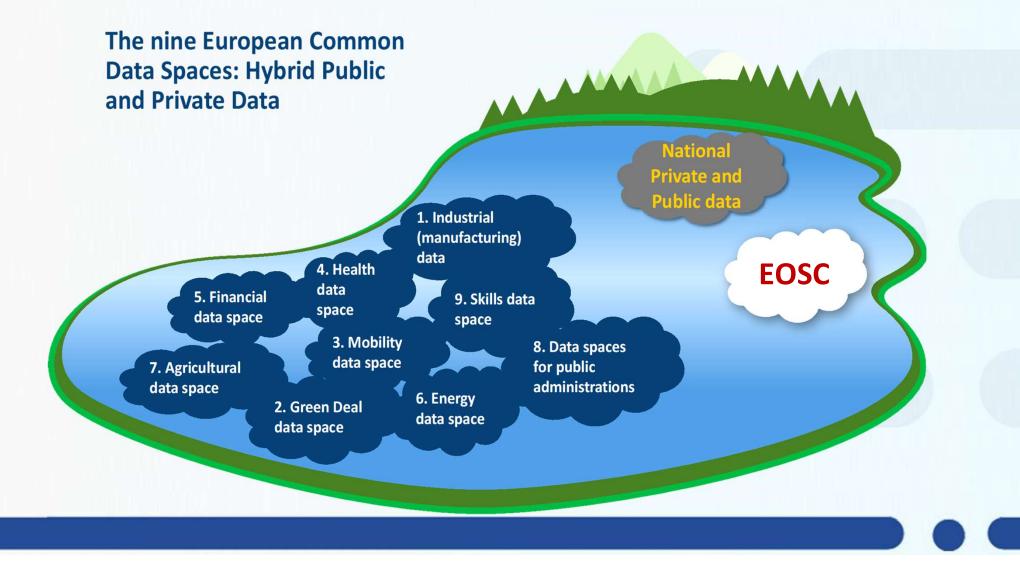
Tasks for EOSC Association to see to:

- Develop and govern federating core
- Manage the AAI
- Manage PID policies
- Manage compliance framework
- Manage trusted certification

- Outreach to stakeholders
- Monitor services and transactions
- Manage the 'EOSC' trademark(s)
- Contribute to Horizon EU policy



# **EOSC in wider context**





## **Position of European Commission**

From EC slides

"EOSC is the basis for a science, research and innovation data space that will bring together data resulting from research and deployment programmes and will be connected and fully articulated with the sectoral data spaces." (European Data Strategy, COM(2020) 66 final)





# FAIR and data interoperability in EOSC

- EOSC interoperability framework is premised on FAIR principles
- Currently defining a minimum metadata model (and crosswalks) to ease discovery over existing federated research data and metadata
- Need clear, publicly-available definitions for all concepts, metadata and data schemas → challenge is that not all communities apply standards
- Need repositories of semantic artefacts and mapping across these
- Aim for all metadata and measurements to be machine-readable

The application of standards and mappings between communities is where EURAMET expertise can help



# **Quality of research data**

## Objective

Explore what are the most relevant quality dimensions considered in different communities and how they can be incorporated in EOSC.

## Approach

- Measures are context dependent and defined by research communities.
- Quality attributes and dimensions (e.g. accuracy, completeness, conformity) vary within and across disciplines and are closely related to the research subject they are intended to measure.
- Identify common measures and align them with the FAIR principles



# Access to unstructured & non-standard data

## Objective

Lack of structure and heterogeneity of non-standard data, i.e. data for which there is either no previously defined data model or which is not organised in a predefined way (e.g. social media streams), innovative, reusable access methods are needed.

## Approach

- Structure such data and make it accessible using AI
- Al as an integral part of the life cycle of research data; need to be managed and made accessible in the same way as the data and the scientific publication itself.
- Include concepts for and implementations of knowledge graphs in EOSC which link data, software, publications and other entities (e.g. authors) to one another.





EOSC Association | <u>www.eosc.eu</u>



# Timeline

