The role of EOSC in applying FAIR principles to research data

Sara Garavelli, CSC - IT Center for Science
& EOSC Association Board of Directors
sara.garavelli@csc.fi

30 | 11 |2023 – **EU Pollinator Week**

“The value of data lies in its use and re-use”

European strategy for data (COM(2020)66 final)

- Research data need to **integrate with other data** and **interoperate with applications or workflows** for analysis, storage and processing
- The overall research domain is a very **dynamic and heterogeneous environment** with many different stakeholder communities
  - Disciplinary interoperability frameworks exist in several communities (=culture, data types and infrastructures used by that community, maturity level and dynamics, etc)

*There is the need to **generate added value** from the resources today available and to **break down silos and broker across borders and disciplines***

EOSC vision in a nutshell

EOSC is the European web of FAIR data and related services for research

Research data that is easy to find, access, interoperate and reuse (FAIR)
Trusted and sustainable research outputs are available within and across scientific disciplines

Three main objectives:

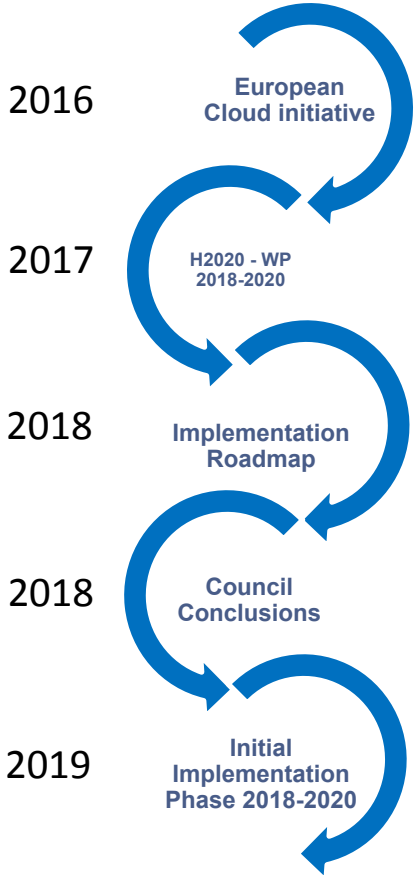
- **Ensure that Open Science practices and skills are rewarded and taught, becoming the 'new normal'**
- **Enable the definition of standards, and the development of tools and services, to allow researchers to find, access, reuse and combine results**
- **Establish a sustainable and federated infrastructure enabling open sharing of scientific results**

Strategic Research and Innovation agenda (SRIA)
eosc.eu/sria-mar

Why EOSC?

- To **enhance scientific research** by providing access to a weath of data and research outputs from various fields and institutions
- To **foster innovation** by making it easier for researchers to share, collaborate and build on each other´s work
- To **improve transparency and reproducibility** of research by making data and methods more open and accessible
- To **reduce data silos and duplication of efforts** by promoting data sharing and reuse
- To **facilitate interdisciplinarity research and cross-sectoral collaboration** by making data and resources from different domains more easily available

A journey started in 2016 …



- Vision
- EOSC roadmap
- Political support (Council Conclusion 2018)
- Targeted resourcing (HE INFRAEOSC 2018-2020)
- Communities at work (EOSC ‘Coalition of Doers’)
- Prototype Portal (eosc-portal.eu)
- Governance 2019-2020
- Targeted resourcing (HE INFRAEOSC 2021-2024)
- EOSC becomes a Co-Programmed Partnership & the EOSC Association is established
- EOSC is recognised by the Council of the European Union action #1 of the policy agenda 2022-2024 of the European Research Area
- The first procurement to operate the EOSC services is awarded

EOSC Association – the Voice of the Community

General Assembly #6, 22-23 May 2023

Over 260 organisations in Europe (RPOs, RFOs, service providers, universities, etc) have joined the EOSC Association. Over 400 experts are mobilised by the EOSC A to contribute to the development of EOSC (EOSC Association Task Forces LINK)

EOSC Co-Programmed Partnership

EOSC STEERING BOARD

European Commission

DG-RTD
DG-CNECT

Anna Panagopoulou

Member States and Associated Countries

Volker Beckmann

MoU

Partnership Board

MS/AC collective representation

Exchange of information

EOSC ASSOCIATION

President
Karel Luyben

Secretary General
Ute Gunsenheimer

BOARD

General Assembly

(including organisations performing research, funding research, providing services, organisations mandated by MS/AC and other organisations)

The FAIR principles are at the foundation of EOSC

- **March 2018: Study "Cost of not having FAIR research data"** by PwC EU Services. Overall cost estimated at €10.2bn/year for the EU alone

- **7 June 2021: Declaration 'Riding the Next Wave of Research Data: Leveraging the COVID-19 response towards advancing data interoperability** (workshop under the Portuguese Presidency of the EU)
    - Coronavirus crisis:
        - Pressure on the research community to speed up scientific discovery
        - Response: global collaboration and significant increase in early and open sharing of knowledge, data and tools.
    - Hurdles:
        - Ability to quickly find and integrate different data types, such as virus and host genomics, clinical, epidemiological and social science data, originating from laboratory, hospital or industrial settings across countries, organisations and communities.
    - Pre-requisite for impactful science:
        - Be able to ride the next wave of data and to increase the volumes of relevant FAIR and open data and their reuse in multidisciplinary science.
        - Call for a wide mobilisation across the European Union towards developing – in the context of the European Open Science Cloud - coherent interoperability frameworks that build on research community practices to enable interdisciplinary research and innovation.

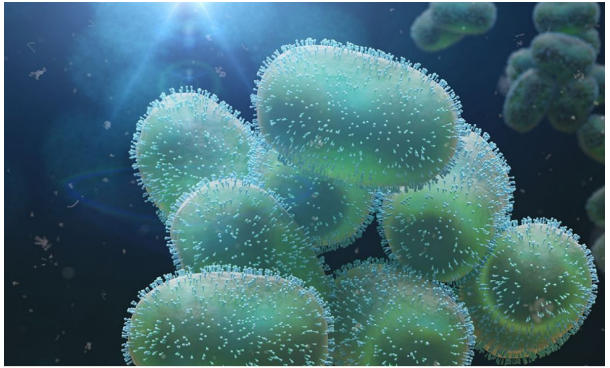
FAIR-related activities in EOSC

- Baseline knowledge on FAIR practices in Europe (**European Research Data Landscape study**, 2022)
- **Core assessment criteria for FAIRness** (FAIR data maturity model as an RDA Recommendation, 2020)
- **FAIR assessment tools** (further work on convergence)
- **FAIR implementation Profiles and FAIR use cases** feeding future Interoperability Frameworks
- **FAIR activities at national and institutional levels** (in-kind contributions through the EOSC Partnership)
- **European Policy papers on FAIR**
  - EOSC FAIR data literacy
  - FAIR data sovereignty in EOSC
  - FAIR data productivity (under preparation)
- **FAIR thematic demonstrators** like:
  - The science clusters (for instance: PANOSC: the photon and neutron open science cloud)
  - BlueCloud 2026: Blue Cloud data lakes for ocean sustainability

The new Pathogens Portal builds on EOSC success

Launch of the new Pathogens Portal in July 2023

- Builds on the same framework as the European COVID-19 Data Portal
- Enables sharing and analysis of pathogen data from across the world.
- Extends to the pathogen world beyond viruses
- Contains data on more than 200 000 pathogen species and strains. All data are open and FAIR
- Helps scientists, healthcare, and public health professionals to collaborate and enhance pathogen surveillance worldwide, especially during public health outbreaks and pandemics.
- Wide effort supported by EMBL and a series of EU-funded projects including EOSC-Life (no. 824087) and BY-COVID (no. 101046203)
- News release: https://www.embl.org/news/science/pathogens-portal-launch/



Credit: dottedyeti/stock.adobe.com

EOSC-Life

BY-COVID

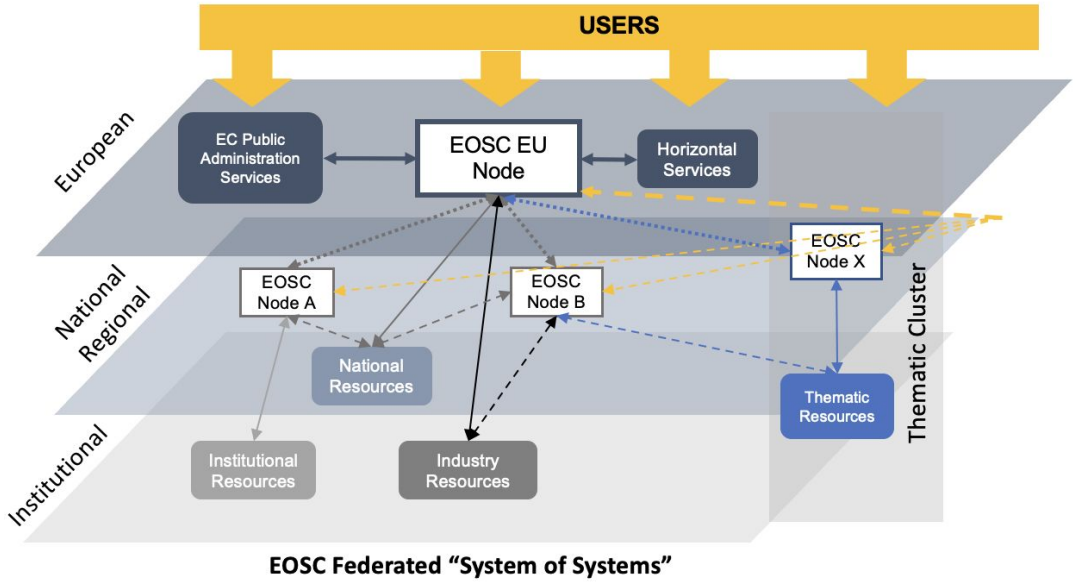
Slide courtesy of the EC

The approach of the FAIR demonstrators

- **Making FAIR data a reality within the community**
  - Adoption of common FAIR data policies
  - Standard metadata and data stewardship within the community
  - Adoption of common templates for Data Management Plans
- **Developing and providing services for scientific data**
  - FAIR data discovery and access service
  - Virtual research environment with distributed computing and storage
  - Virtual labs with analytical services
- **Connecting these resources to the EOSC Federation**

The EOSC Federation: from prototype to operation

- A **federation of distributed systems**, combined into a system of systems
- **consisting of multiple "nodes"** that are interconnected and can **collaborate to share and manage information and resources (FAIR data & services)**
- **across thematic and geographical boundaries**



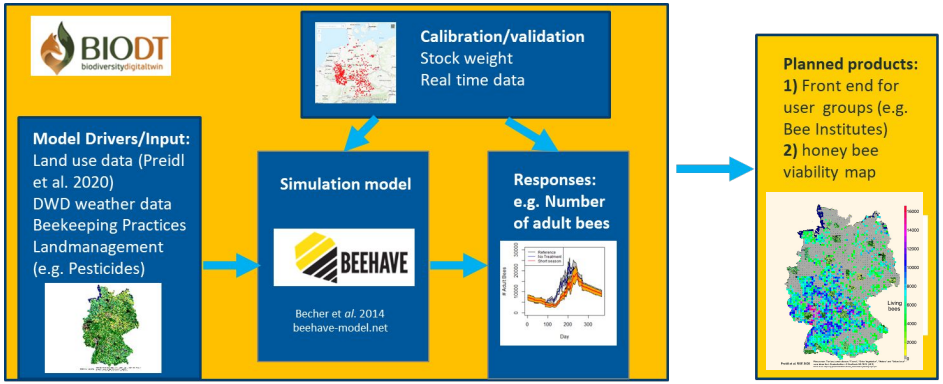
The first EOSC node will be operational in 2024 and will act as blueprint for the other nodes

EOSC and pollinators: two different worlds?

An example from the Biodiversity Digital Twin project (biodt.eu)

- Honey bees and other pollinators are exposed to multiple stressors such as pesticides, forage gaps, diseases and management failures. To better understand the impacts of these multiple stressors on bee vitality in agricultural landscapes, **a honey bee pollination digital twin will be developed** in the BioDT project

- The objective of the use case is **to produce a map for Germany that can be used to assess the quality of the landscape for honey bee vitality and productivity given environmental and beekeeping management scenarios**.



UFZ HELMHOLTZ Centre for Environmental Research

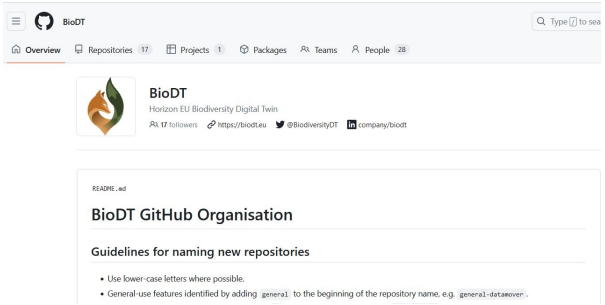
CSC

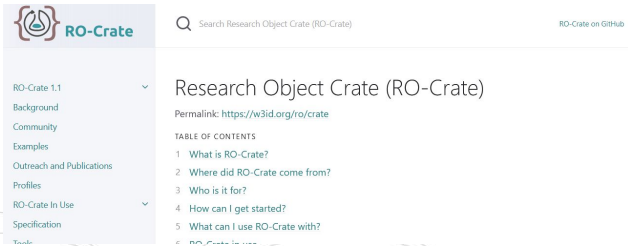
VŠB TECHNICKÁ UNIVERZITA OSTRAVA | IT4INNOVATIONS NÁRODNÍ SUPERPOČÍTAČOVÉ CENTRUM

EOSC and pollinators: two different worlds?

An example from the Biodiversity Digital Twin project

- **Land cover data is stored locally or can be downloaded** from Pangea repository; weather data is streamed from German weather forecast (DWD) --> In the future EOSC can facilitate **data access and harmonization**
- **Different simulation models are available** (BEEHAVE has been selected in this case) □ EOSC will improve **discoverability of resources** (incl. models)
- The **resulting workflows will be published on the BioDT Github** in 2024
- **Metadata descriptions** for BEEHAVE will be developed as RO-CRATES
- **The generated data could become part of the EOSC federation and thus reusable also in other domains**





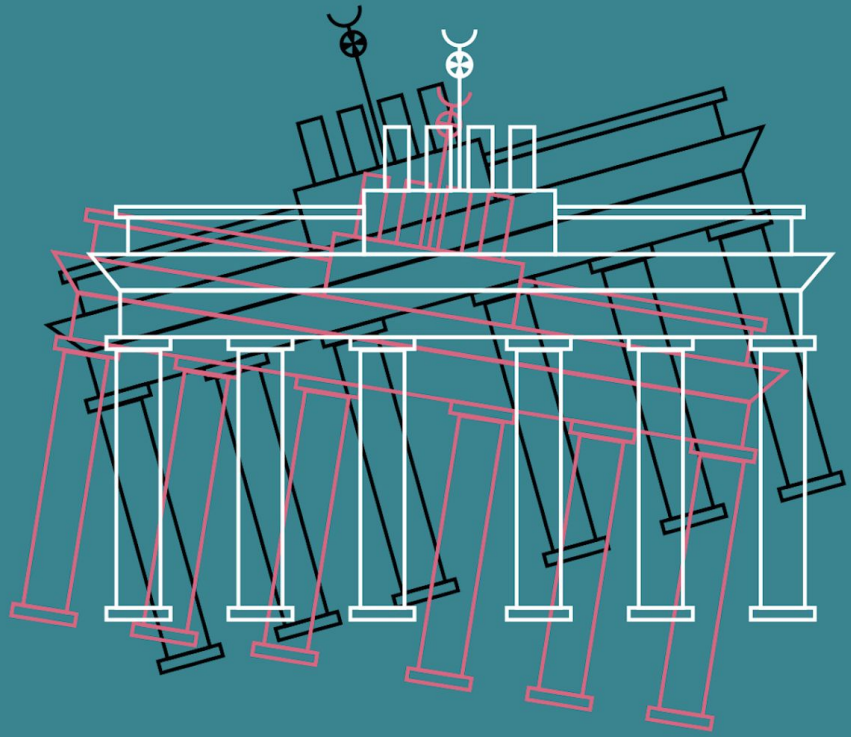
The BioDT project is already contributing to EOSC, other contributions are welcome!

Conclusions

- EOSC is a federated initiative that aims to **promote the production and reuse of research data**
- EOSC helps researchers manage and make their data FAIR. This ensures that **new generated data can be easily discovered and reused** by other researchers
- EOSC provides a platform for researchers to **share and integrate data** from various sources and disciplines (the operational phase of EOSC will start in 2024): **different communities and researchers are invited to contribute to the EOSC federation with data and services!!!**
- EOSC fosters **interdisciplinary collaboration** by providing a common framework for researchers from various fields to work together and share knowledge and resources.
- Scientific research often require access to advanced research infrastructure, such as high-performance computing facilities, data repositories, and laboratories. **EOSC will help coordinate access to these resources**, ensuring that researchers have the tools they need to carry out their work.
- EOSC will continue the work **on aligning practices and standards for high-quality FAIR data** as well as the **development of tools & services to find, access, reuse and combine FAIR research results**
- EOSC will continue to advocate **for Open science becoming the 'new normal'**

eosc

EOSC Symposium 2024



21 - 23 October / Berlin, Germany

Connected to vpn.csc.fi.

eosc.eu | #eoscsymposium2024

UNDER THE PATRONAGE OF BETTINA STARK-WATZINGER MDB

Federal Ministry of Education and Research

eosc | Focus

Funded by the European Union

In cooperation with

nfdi Nationale Forschungsdaten Infrastruktur

ZBW Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

EOSC Association AISBL

Rue du Luxembourg 3
BE-1000 Brussels, Belgium
+32 2 537 73 18
info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931
VAT number: BE0755 723 931

Many Thanks... and stay tuned