# EOSC CZ and the National Repository Platform

EOSC CZ Secretariat & Michal Růžička <ruzicka@ics.muni.cz> Masaryk University, Institute of Computer Science, CERIT-SC

Conference on Digital Transformation of Universities
University of Pardubice
2025-09-10











## **Contents**

- EOSC & EOSC CZ Projects
- Research & FAIR Data as Scientific Results
- Related Communities and Working Groups
- National Data Infrastructure
  - Overview of selected components



# European Open Science Cloud (EOSC)

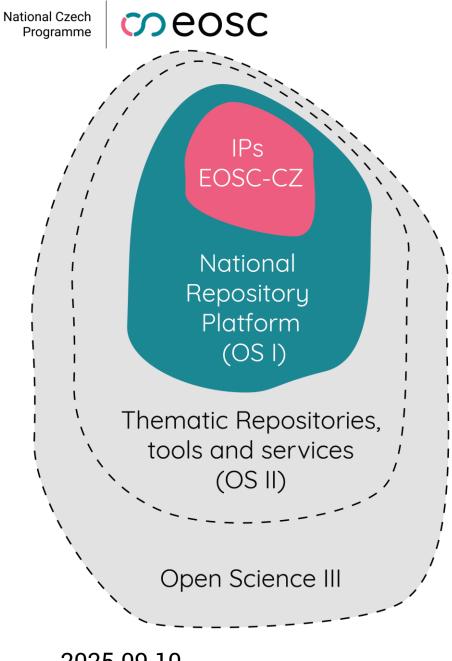
- Initiative to support data management the entire data life cycle
  - Since 2016, mainly at the EU level, the search for a suitable setting through a series of projects
  - Currently the concept of EOSC Federation, de facto a large meta-infrastructure
- The key concept is data management
  - Not just any data, but FAIR data i.e. well managed, described, reliably stored and reusable
  - FAIR data is not necessarily Open data support for data access management is included
- Ecosystem of data and related services
  - Built on existing foundations, i.e. mainly large research infrastructures and einfrastructures



## **EOSC CZ**

- Supported through OP JAK (MEYS)
- Series of projects:
  - EOSC-CZ, CARDS, NRP, partly Research Environments, OS II (submitted), OS III (planned)
- Related to Data Management Plans (DMPs)
- Supports objectives:
  - Create an environment for working with FAIR research data
  - Primarily Preserve, Share, Reuse
  - Plan and collect is a matter of research
  - Process and Analyse are related to (e-)infrastructures



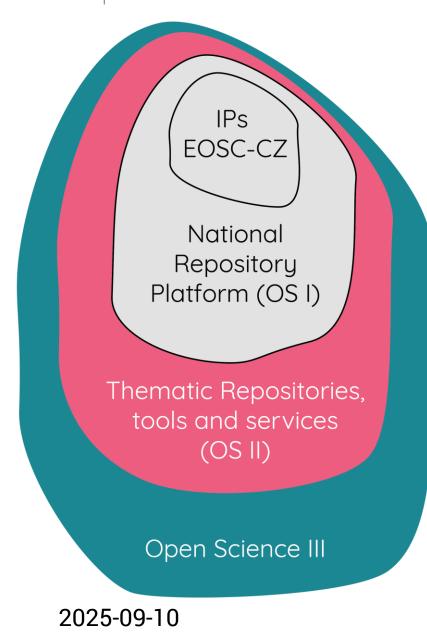


#### **IPs EOSC-CZ – fundamentals for EOSC implementation in CZ**

- Organizational (Secretariat) https://www.eosc.cz/en/about-eosc-cz/secretariat-eosc-cz
- Technical (National Metadata Directory) https://nma.eosc.cz/
- Educational (Training Centre) https://www.eosc.cz/en/training

#### National repository platform - "technical core"

- Repository systems (DSpace, CESNET Invenio, ASEP-ARL)
- **Pilot repositories**
- Core services (PIDs, DSW, licenses, ...)
- Compliance and UX (Cybersecurity, ServiceDesk, ...)
- **Training** technical aspects



#### OS II - "domain-specific outputs"

#### **Based on the inputs from 8 thematic working groups**

- Domain-specific and interdisciplinary activities
- Development of thematic repositories and tools for FAIR data management
- Bio/Health/Food, Matech, AI & ML, Social Sciences, Physics, Humanities & Arts, Environmental Sciences, Sensitive Data
- https://www.eosc.cz/en/working-groups

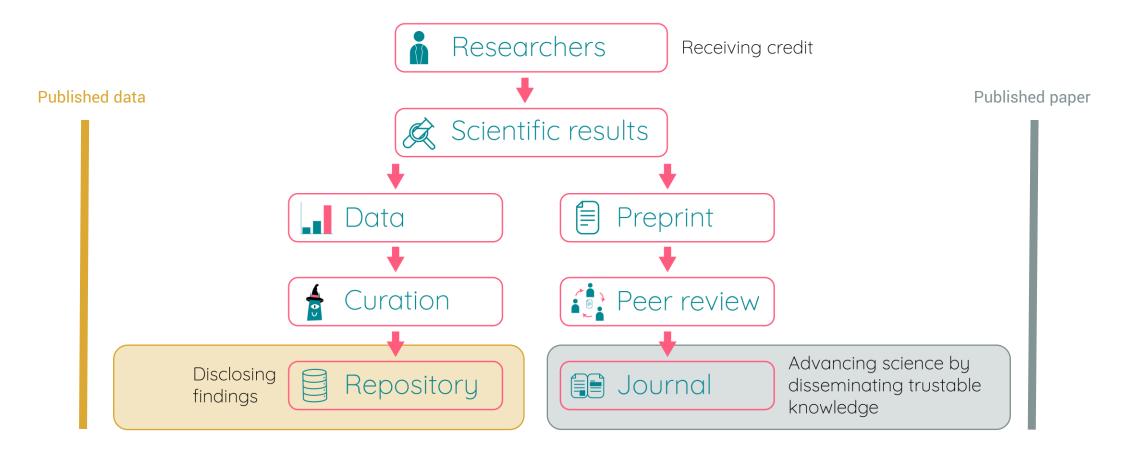
#### OS III – "development of knowledge and skills"

# Development of human resources and competencies in research data management

- Involvement of institutions where the topic of data management is in early stages
- Communication and educational activities
- Establishing institutional support for FAIR data management
- Funded through mini-projects



#### Peer reviewed research



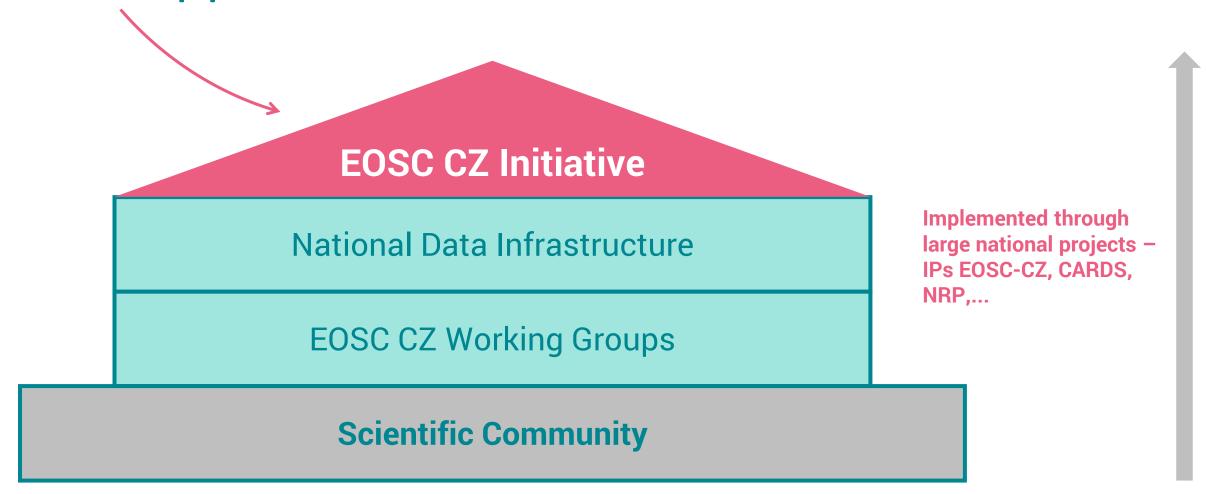
This **ALSO** is a figurative checkmark in your career

This is a figurative checkmark in your career

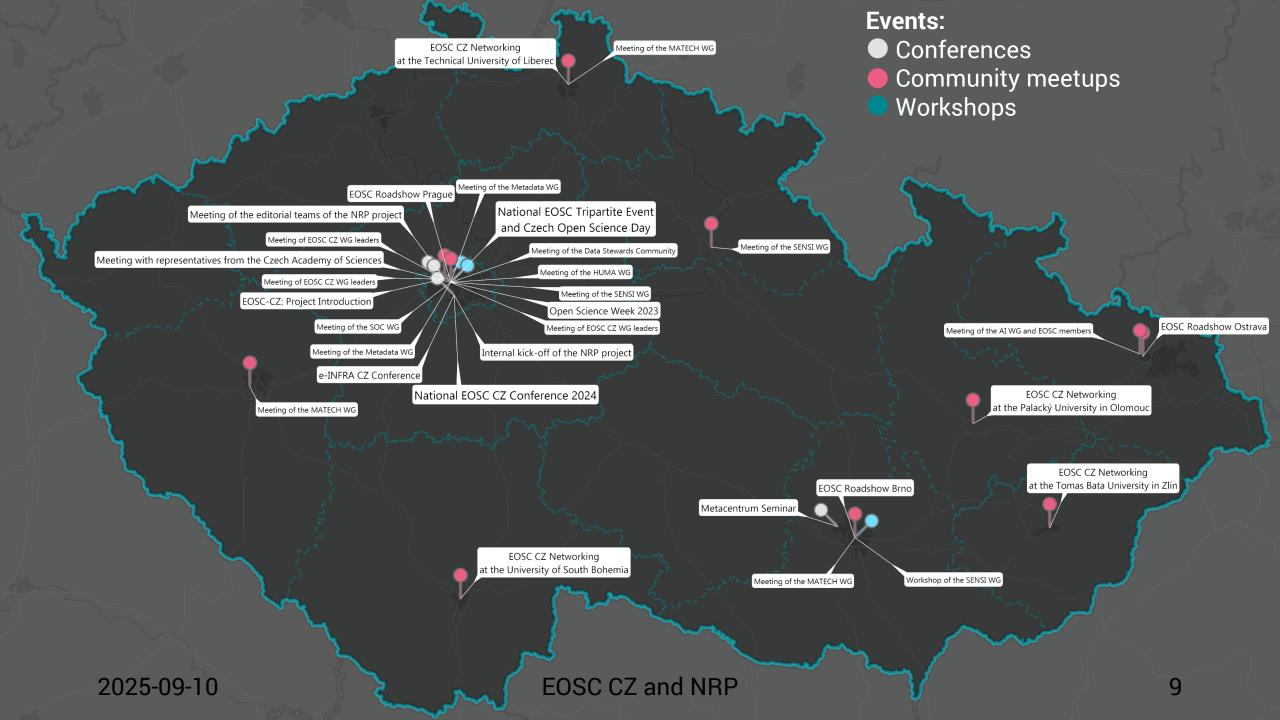
2025-09-10 EOSC CZ and NRP 7

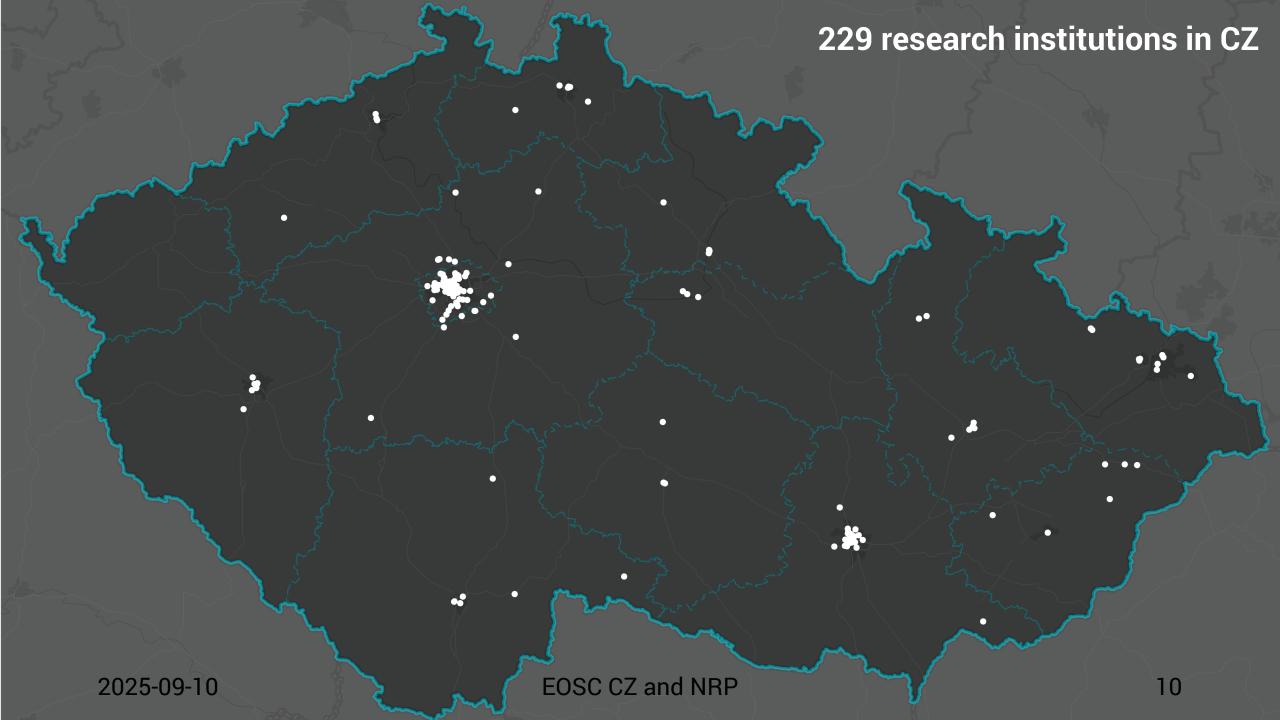


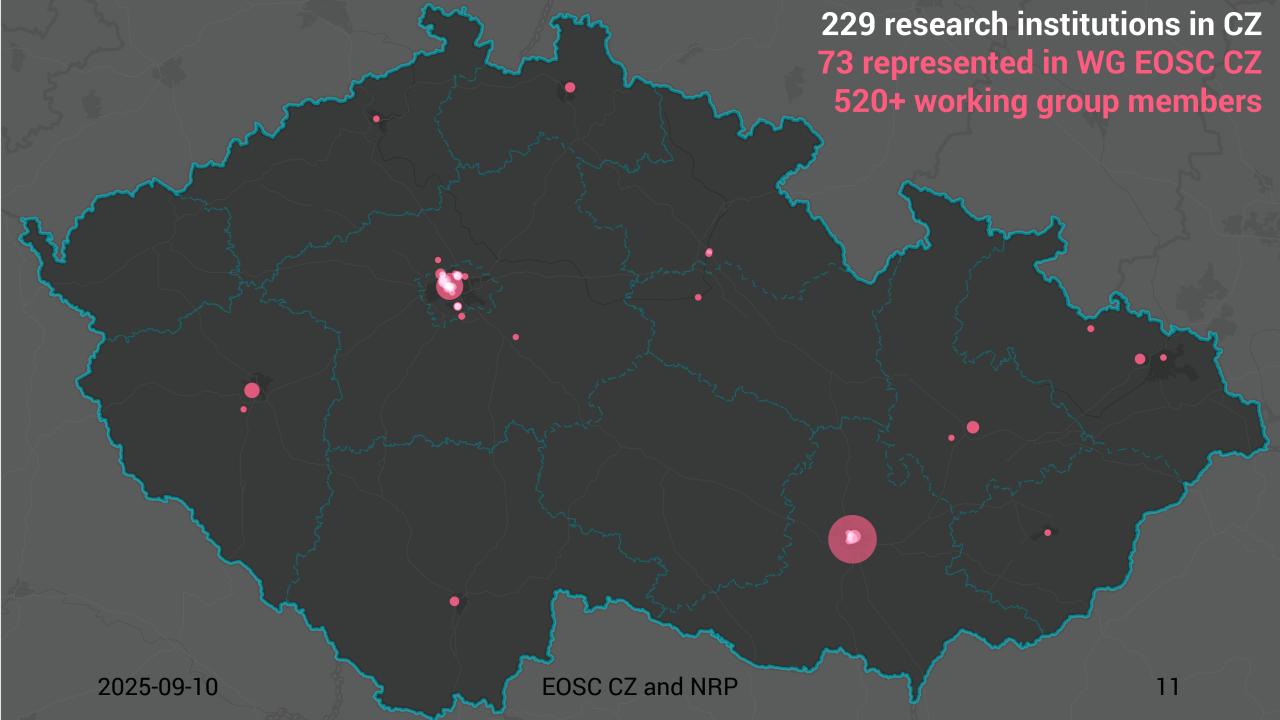
## FAIR Support in Czechia



2025-09-10 EOSC CZ and NRP 8







# **EOSC CZ Working Groups**









Core Services

Education and Human Resources







Materials Sciences and Engineering

Metadata



Data Management for Artificial Intelligence and Machine Learning



Social Sciences

#### **Open Platform for EOSC CZ Implementation**

- Main EOSC CZ Building Blocks (scientists for scientists)
- 4 cross-sectional and 8 field-specific groups
- Open to new members at any time
- Operating throughout the duration of the initiative









2025-09-10 EOSC CZ and NRP 12

Physical Sciences



# **EOSC CZ Working Groups**

#### How do WGs work?

- Regular meetings
- Notes and recordings
- Open membership on a voluntary basis

### Why get involved?

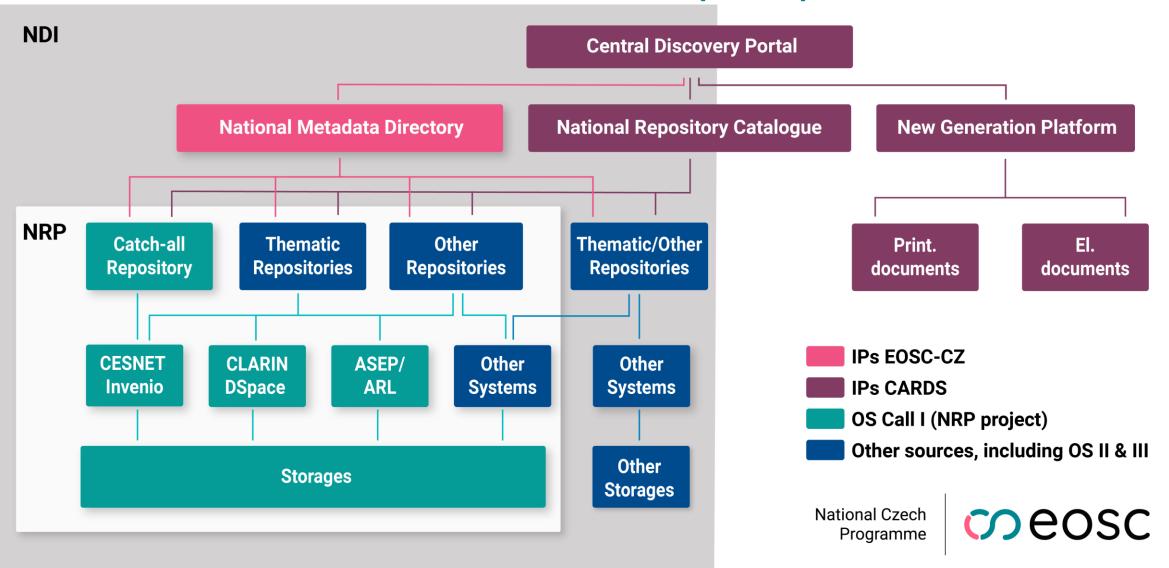
- Expertise and awareness
- Building connections
- Opportunity to make an impact

#### Are you interested in becoming a member?



https://www.eosc.cz/en/working-groups

## National Data Infrastructure (NDI)



## **NDI Outputs: Storage Capacities**

#### **Repositories**

- Catch-all repository (2025)
- Thematic (domain-specific) repositories

4 pilots: Molecular Biophysics Database\*, National Repository for Biodiversity Data, Repository for Biological Imaging Data, ArchaeoVault (end of 2025)

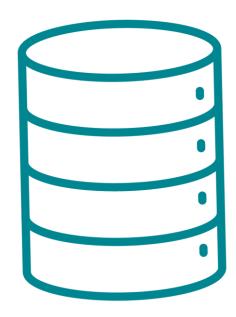
- Others from 2025/2026
- National Metadata Directory

#### **Repository systems**

CESNET Invenio, CLARIN-DSpace, ASEP-ARL

#### **Hardware**

- Physical, distributed storage infrastructure
- Total of 50+ PB of user data storage capacity



\* in production mode already



# Al-ready (Valuable) Data



Your (author)

Citable (DOI) + Findable

Accessible + Interoperable

Reusable (licence)

Machine actionable (metadata)

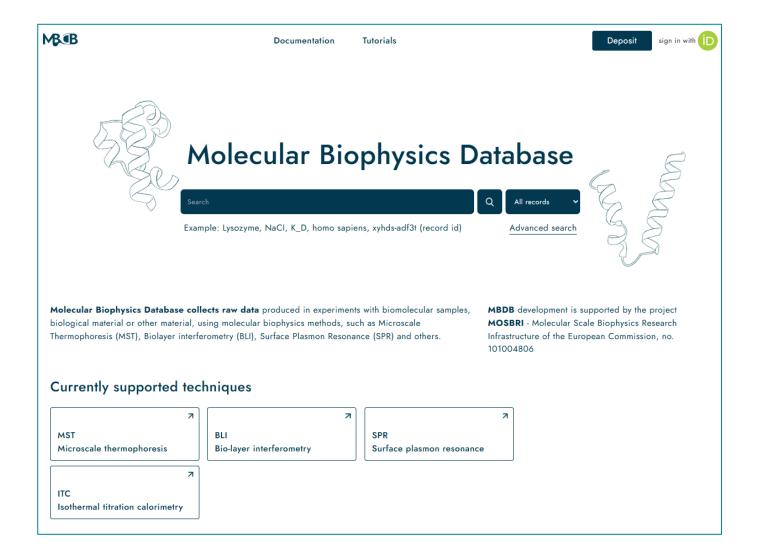
= Al-ready record (dataset)

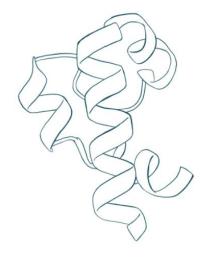
**VALUABLE SCIENTIFIC RESULT** 

2025-09-10 EOSC CZ and NRP 16



# Molecular Biophysics Database





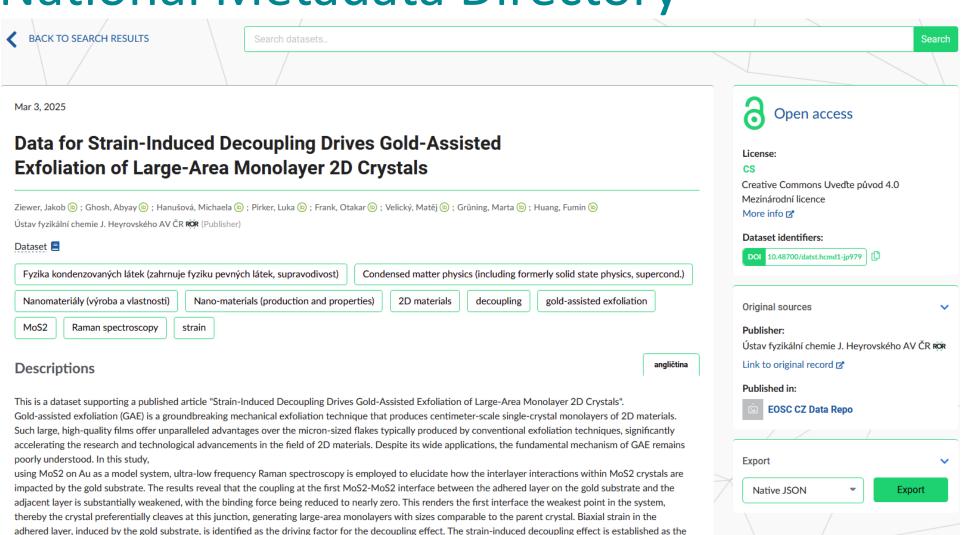


https://mbdb-data.org/



primary mechanism of GAE, which can also play a significant role in general mechanical exfoliations.

## **National Metadata Directory**





https://nma.eosc.cz/



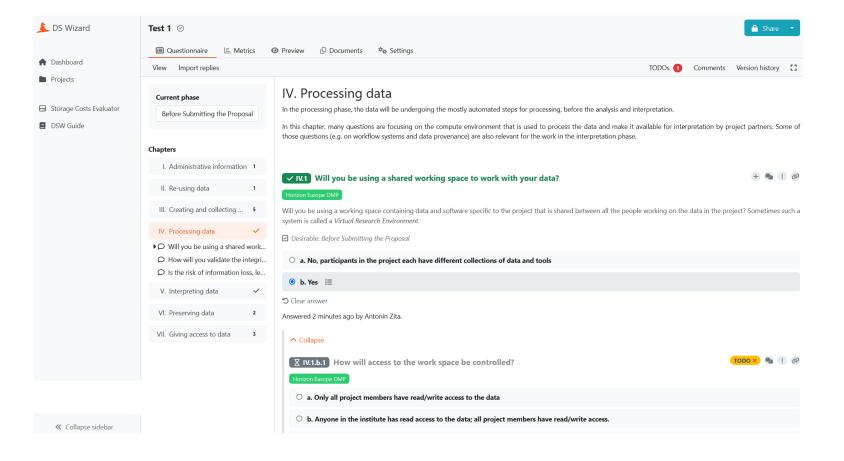
# NDI Outputs: Tools and Services

- Support for <u>data management planning</u> (DMP)
- Metadata profile management
- Support for license handling
- Support for working with <u>persistent identifiers</u>
- Support for FAIRification of research data
- Automation of data collection
- Electronic laboratory notebooks
- Overall cybersecurity and system compliance



# Data Stewardship Wizard

#### Research data management planning tool









## Persistent Identifiers



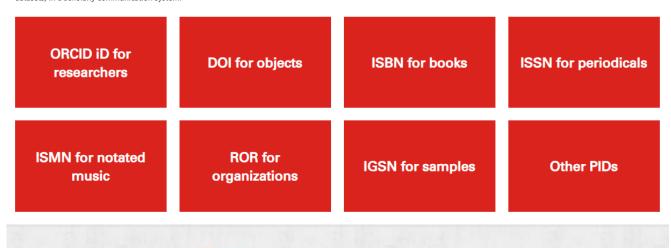
Persistent Identifiers



Home / Persistent Identifiers

#### Persistent Identifiers

Learn more about each persistent identifier (PID). Persistent identifiers are tools that are used to uniquely identify people, organisations, and other objects (e.g., books, articles, datasets) in a scholarly communication system.



About us

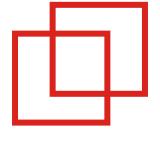
National ORCID Centre

FAQs - ORCID

Persistent Identifiers

DOI for objects

ORCID iD for researchers



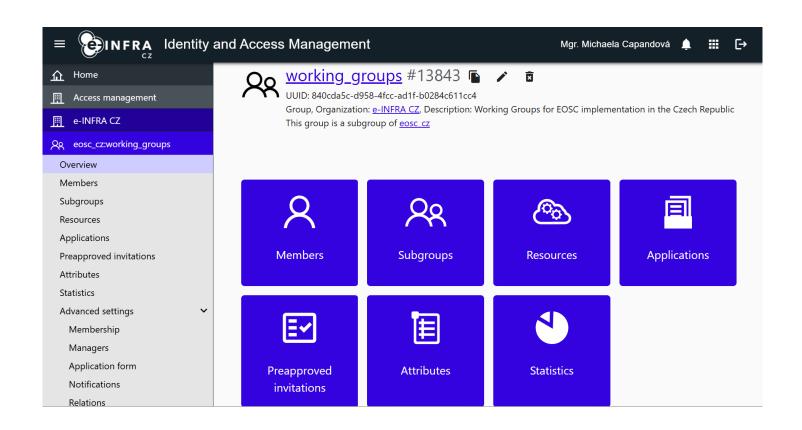


https://identifikatory.cz/en/



### Authentication and Authorization Infrastructure

Enabling users from different institutions to easily access data and services.



- Access and identity management
- Group and role management
- Permission assignment



https://perun-aai.org/



## SensitiveCloud







## Secure environment for storing, sharing

#### and processing sensitive data

- Primarily designed for work with your own sensitive data
- Gradually extended to support controlled data sharing
  - · Main technical component for handling sensitive data within NDI
- Includes storage, computing resources and support for ready-to-use web applications









## **NDI Outputs: Summary**



**Storage capacities** 



**Tools and services** 



**Computing capacities** 



## **Useful Links and Contacts**

#### **EOSC CZ Website**



**EOSC CZ Newsletter** 



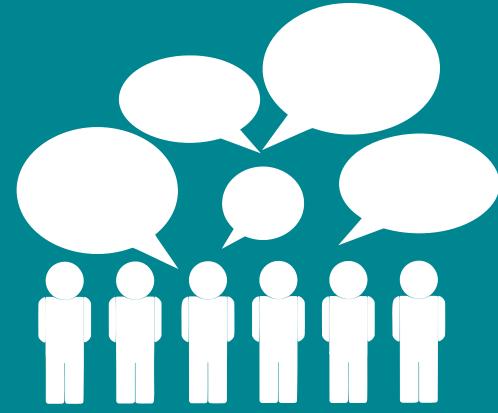
- Any questions?
  - o info@eosc.cz
- Ideas for a lecture or a training?
  - o <a href="mailto:events@eosc.cz">events@eosc.cz</a>
- · Get in touch with our PR
  - o pr@eosc.cz





@eosccz.bsky.social

# Thank you for your attention



Source: Communicate communication conference 2028004 by OpenClipart-Vectors from Pixaba









