

# EOSC CZ Initiative

Matej Antol, Ph.D.



14.11.2024

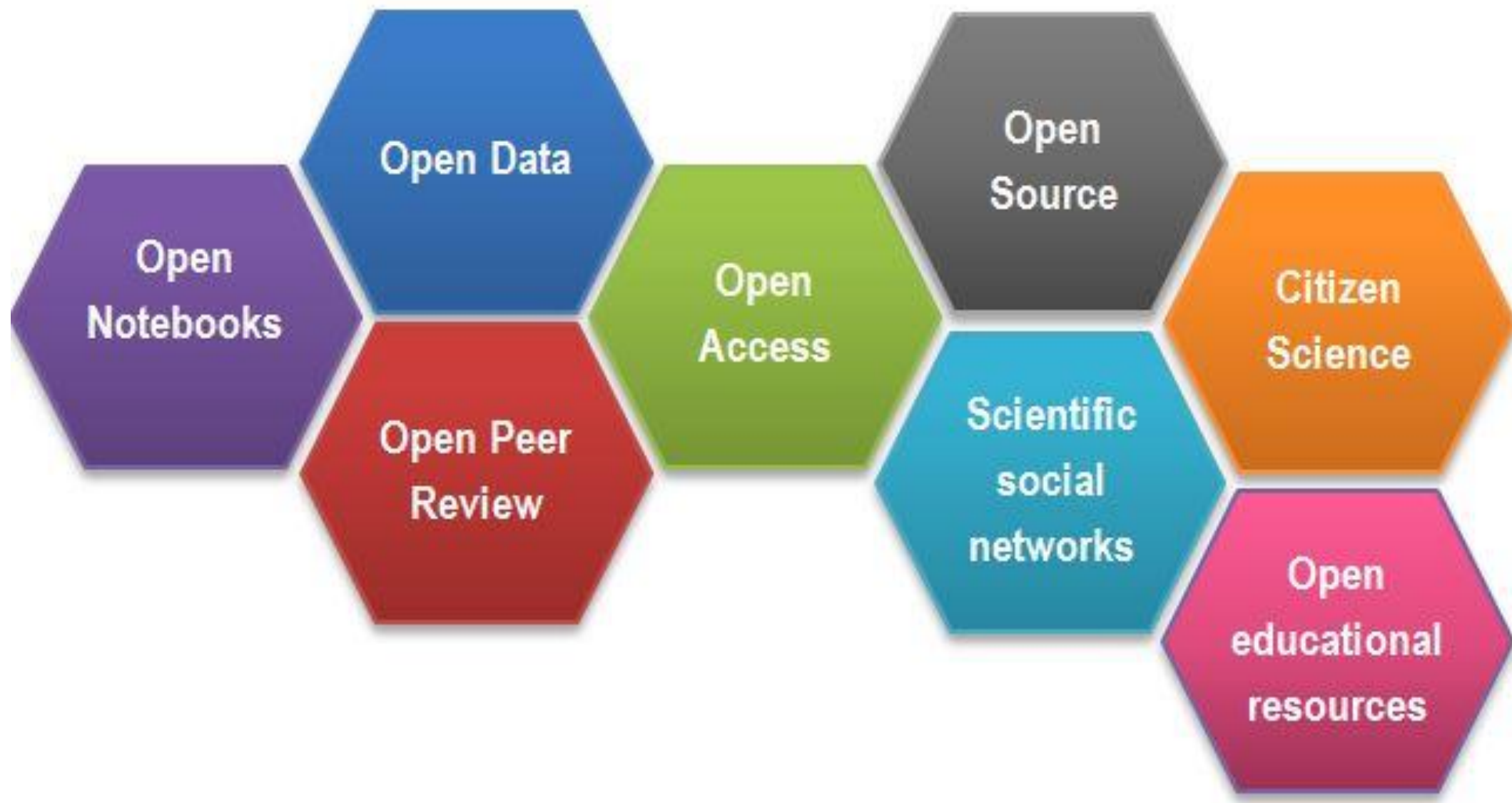


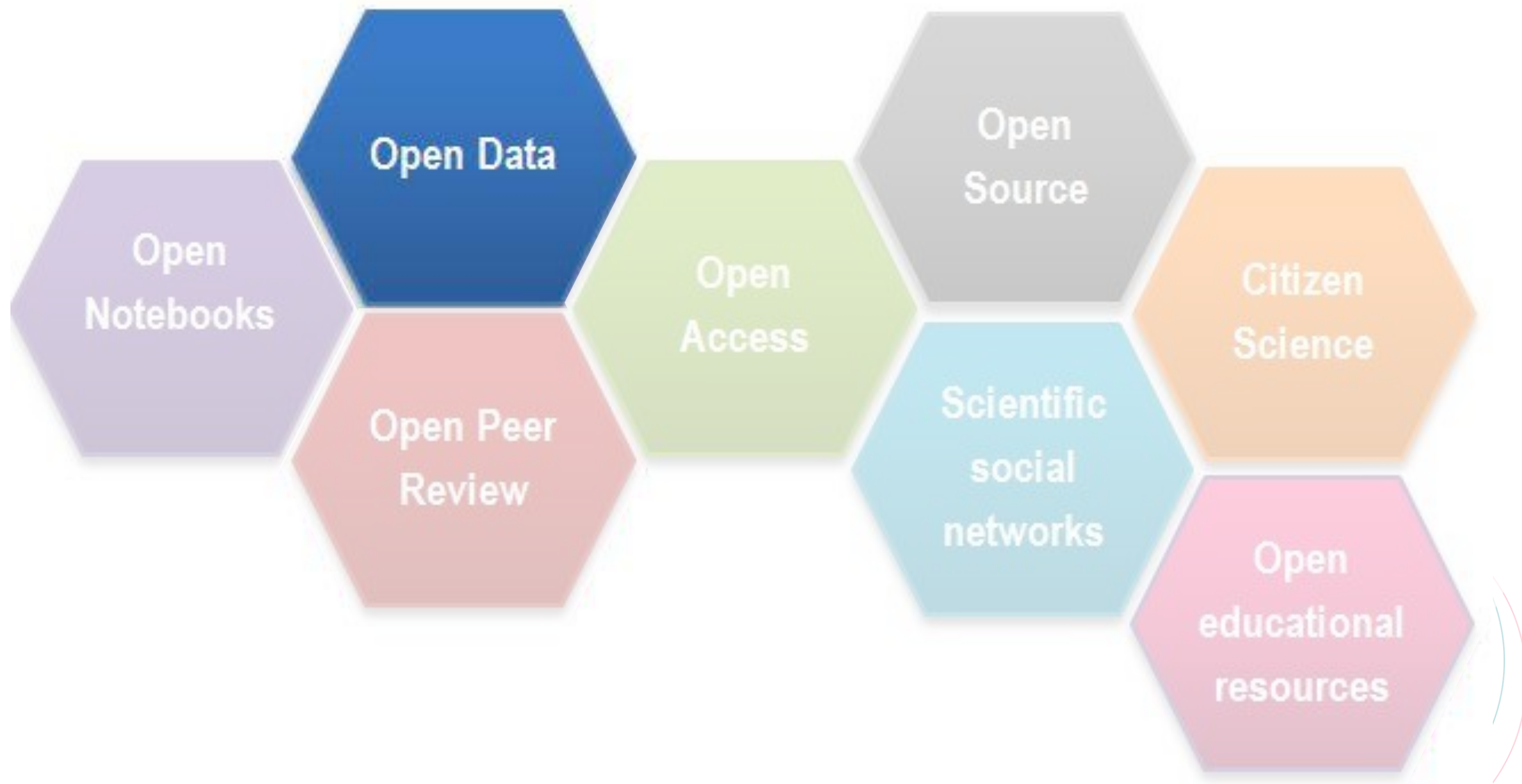
Spolufinancováno  
Evropskou unií

# About me

- Research group co-lead
  - Intelligent Systems for Complex Data
  - Faculty of Informatics, Masaryk university, CZ
  - <https://disa.fi.muni.cz/complex-data-analysis>
- CEO @ CERIT-SC, ICS MU
  - one of three partners of the national e-infrastructure
  - [www.cerit-sc.cz/](http://www.cerit-sc.cz/)
  - <https://www.e-infra.cz/>
- Principal Project Manager of the IPs EOSC CZ
  - systemic, strategic project within the Open Science
  - initiative creating national environment for FAIR
  - <https://www.eosc.cz/>







# ~~OPEN~~ -> FAIR data

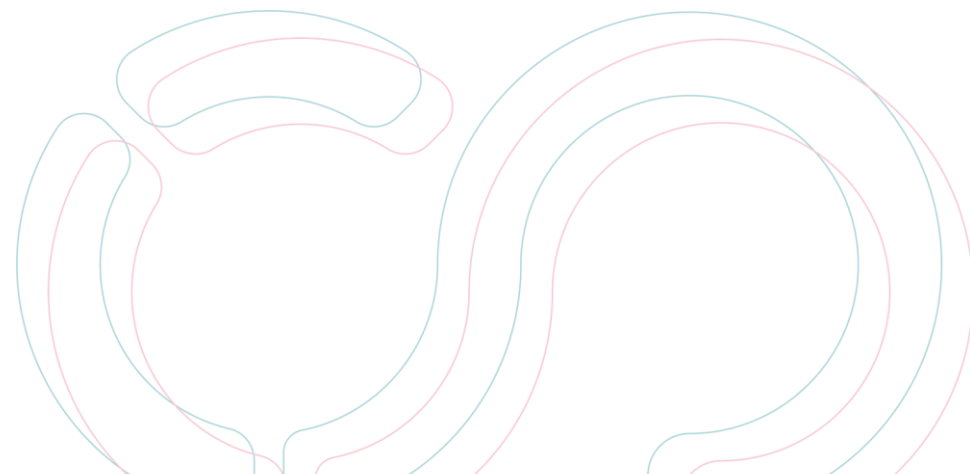
- Concept of open data simply **can't serve as a guiding principle for all research data**
  - Some are sensitive
  - Some have licencing restrictions
  - Some provide competitive advantage
  - Missing incentives
  - ...
- Pragmatically, so-called **FAIRness of data may and should be pursued**
  - Findable
  - Accessible
  - Interoperable
  - Reusable
- FAIR data == **well managed data**



# In this presentation

- Brief history of research data
- The EOSC CZ initiative
- Czech National Data Infrastructure and National Repository  
Platform for FAIR  
research data
- Research infrastructures in Czechia
- e-INFRA CZ

# Brief history of research data



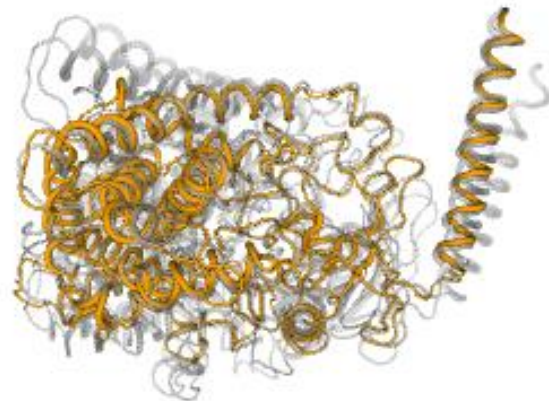






# Chemistry, structural biology and life

- PDB (protein data bank) was established in 1971
  - 60.000 depositors
  - 500.000 data entries: experimentally obtained structures
- A couple of weeks ago, the Nobel Prize for Chemistry was awarded for computational protein design and protein structure prediction -- AlphaFold
- Both PDB and AFDB now used by millions for countless applications such as decomposing pathogens and antibiotic resistance



III. Niklas Elmehed © Nobel Prize  
Outreach  
David Baker  
Prize share: 1/2



III. Niklas Elmehed © Nobel Prize  
Outreach  
Demis Hassabis  
Prize share: 1/4



III. Niklas Elmehed © Nobel Prize  
Outreach  
John M. Jumper  
Prize share: 1/4

The Nobel Prize in Chemistry 2024 was divided, one half awarded to David Baker "for computational protein design", the other half jointly to Demis Hassabis and John M. Jumper "for protein structure prediction"

# Astronomy, physics and the universe

- Since 2000, The Sloan Digital Sky Survey (SDSS) collects data of galaxies
  - More than 500 contributors from 13 countries
  - Millions of data entries: 2.5 million galaxies and 400.000 quasars
- Its data helped prove the accelerating expansion of the universe, leading to the Nobel Prize in Physics in 2011 for the discovery of dark energy.
- Currently, it is a resource for for over 10.000 scientists and contributed to the discovery of over 500.000 red dwarfs and 20.000 asteroids



© The Nobel Foundation. Photo: U. Montan  
Saul Perlmutter  
Prize share: 1/2



© The Nobel Foundation. Photo: U. Montan  
Brian P. Schmidt  
Prize share: 1/4



© The Nobel Foundation. Photo: U. Montan  
Adam G. Riess  
Prize share: 1/4

The Nobel Prize in Physics 2011 was divided, one half awarded to Saul Perlmutter, the other half jointly to Brian P. Schmidt and Adam G. Riess "for the discovery of the accelerating expansion of the Universe through observations of distant supernovae"





# Art, provenance, culture and history

- Getty Provenance Index (GPI) exists since 1980
  - Data from more than 50 institutions worldwide, including museums, galleries, and auction houses, as well as individual collectors and research scholars.
  - Over 1.5 million records: over 300,000 individual works of art, 270,000 owners and 200,000 auction sales.
- Significant impact on restitution efforts, particularly for art looted during World War II, just as the painting called “Girl” by Jean-Baptiste Greuze called
- Big impact on Art History and Provenance Research, with over 1.000 scholarly publications, spanning fields like art history, economics of the art market, cultural heritage studies and legal research on restitution.

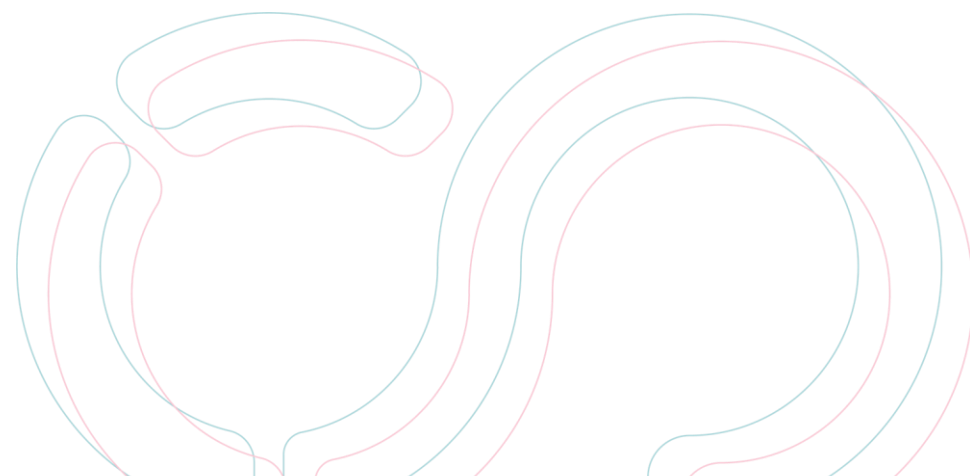




# DATA IS THE NEW OIL



# The EOSC CZ initiative



# European Open Science Cloud (EOSC)

- Technological, program and process foundation supporting better research data management
- Including
  - Capacity
  - Tools
  - Access management
  - Interfaces for analysis
  - ...
- **Federated** – not a single product, service or single monolithic environment
- **Interconnected** – across research domains and EU states



# EOSC CZ initiative

- EOSC CZ working groups
- Federated ecosystem of FAIR research data and services supporting Czech science
  - National Data Infrastructure (NDI)
  - In its heart, a National Repository Platform (NRP)
- Education and development of skills – there is no use of a hammer without a capable hand
- Funded via P JAC via 5 projects / Open Science project calls, together ca 180 mil. EUR

2023

2024

2025

2026

2027

2028

2029

Systemic Project EOSC CZ

Systemic Project CARDS

First Open Science Call: Single Project of National Repository Platform (NRP)

Second Open Science Call: Project(s) for Thematic Domain Clusters

Third Open Science call: Development of Human Competencies

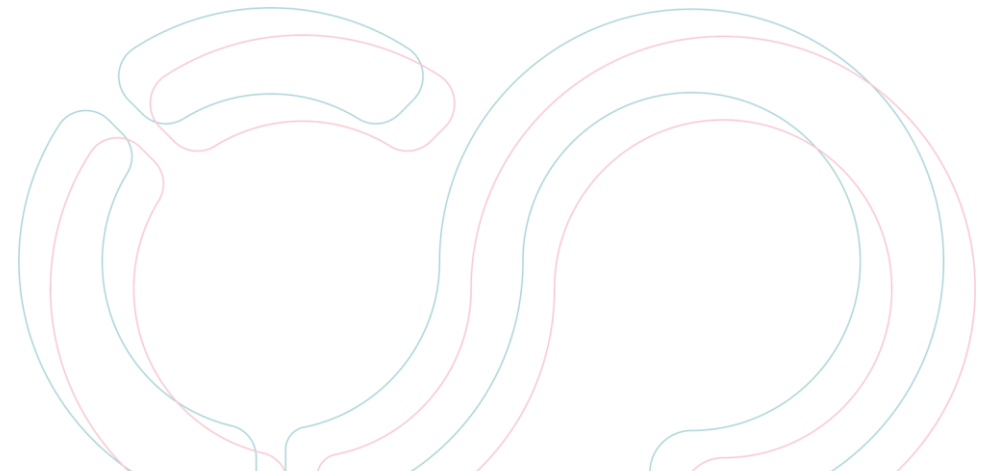
# 12 working groups

## 4 cross-cutting

- NDI Architecture
- Metadata
- Core Services
- Education and Human Resources

## 8 thematic working groups

- Bio/Health/Food
- Social Sciences
- Physics
- Humanities and the Arts
- Material Sciences and Technology
- Environmental Sciences
  
- Data Management for AI and ML
- Sensitive Data





**DAVID ANTOŠ**  
CESNET / e-INFRA CZ  
| National Data Infrastructure Architecture |



**PETRA ČERNOŠKOVÁ**  
National Library of Technology  
| Metadata |



**RADKA ŘÍMANOVÁ**  
Central Library of Charles University  
| Education and Human Resources |



**MICHAL RŮŽIČKA**  
ICS Masaryk University / e-INFRA CZ  
| Core Services |



**JIŘÍ VONDRÁŠEK**  
Institute of Organic Chemistry and Biochemistry of the CAS / ELIXIR - CZ  
| Bio/Health/Food |



**JINDŘICH KREJČÍ**  
Institute of Sociology of the CAS  
| Social Sciences |



**JIŘÍ CHUDOBA**  
FZU (Institute of Physics of the Czech Academy of Sciences)  
| Physical Sciences |



**JAN HAJIČ**  
Charles University / LINDAT / CLARIAH-CZ  
| Humanities and the Arts |



**MAREK CEBCAUER**  
J. Heyrovský Institute of Physical Chemistry of the CAS  
| Materials Sciences and Engineering |



**JAN MARTINOVIČ**  
IT4Innovations National Supercomputing Center VŠB - TUO  
| Data Management for Artificial Intelligence and Machine Learning |

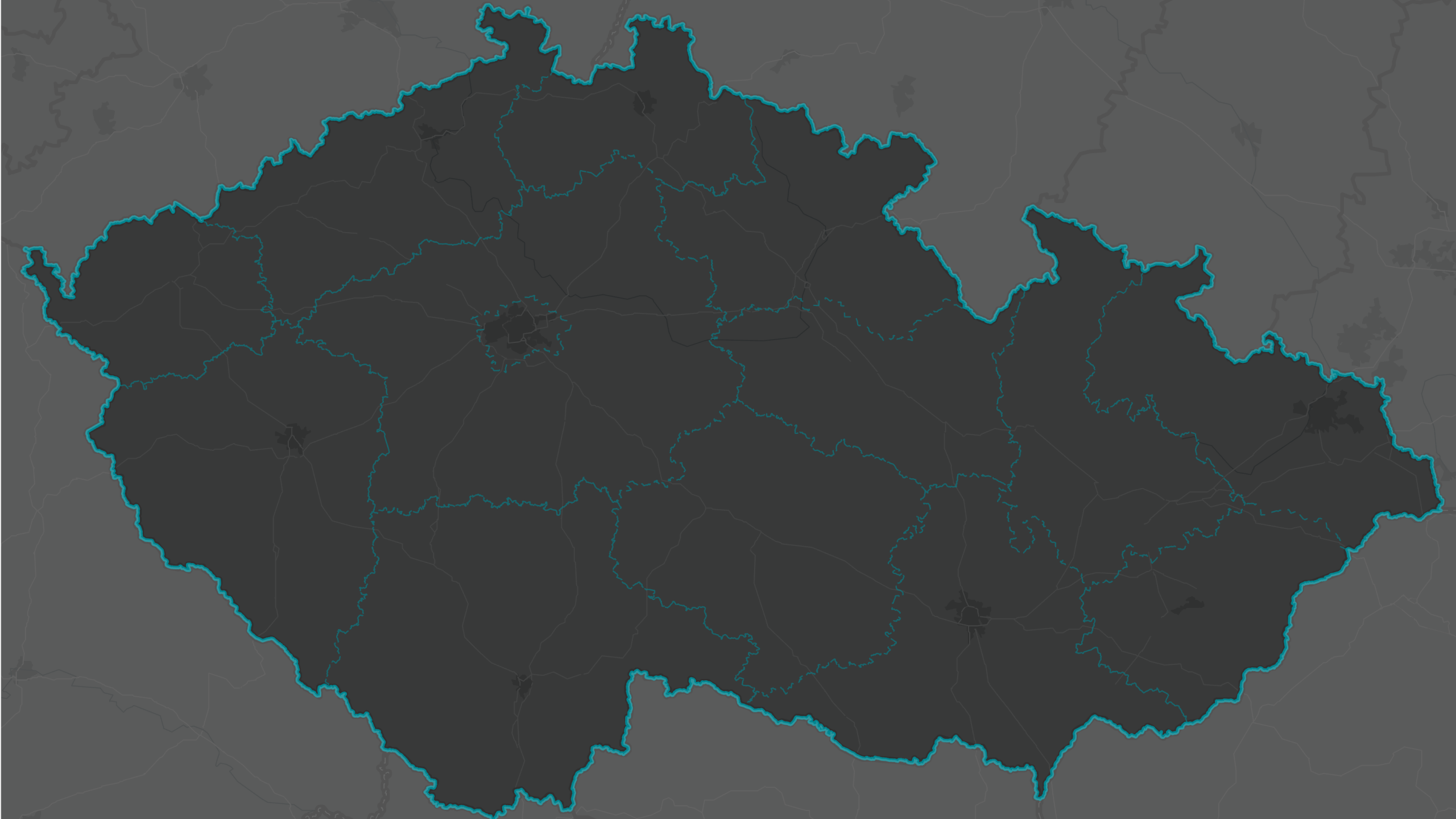


**JANA KLÁNOVÁ**  
Masaryk University / RECETOX  
| Environmental Sciences |

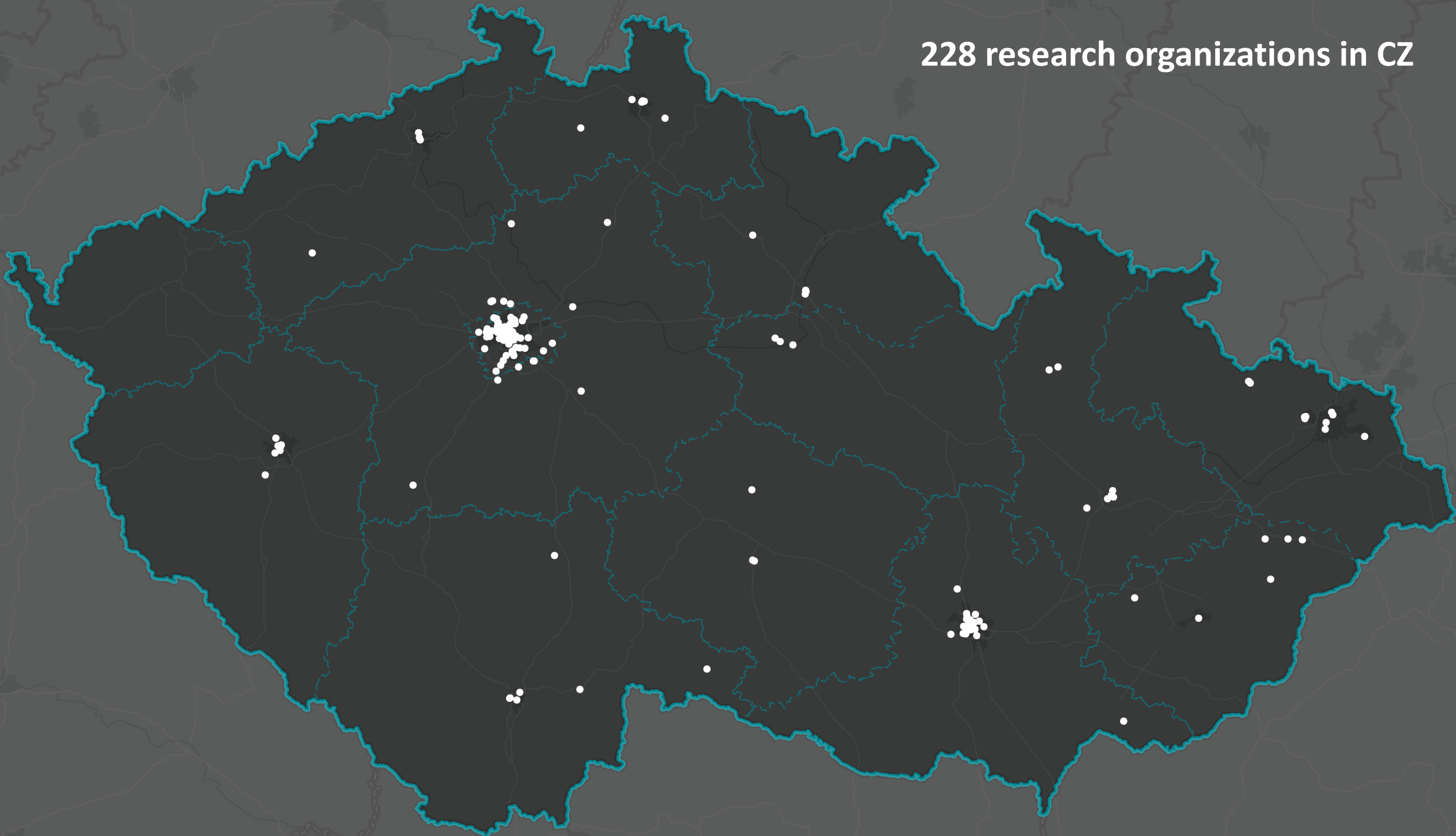


**ZDENKA DUDOVÁ**  
BBMRI.cz, Masaryk Memorial Cancer Institute  
| Sensitive Data |



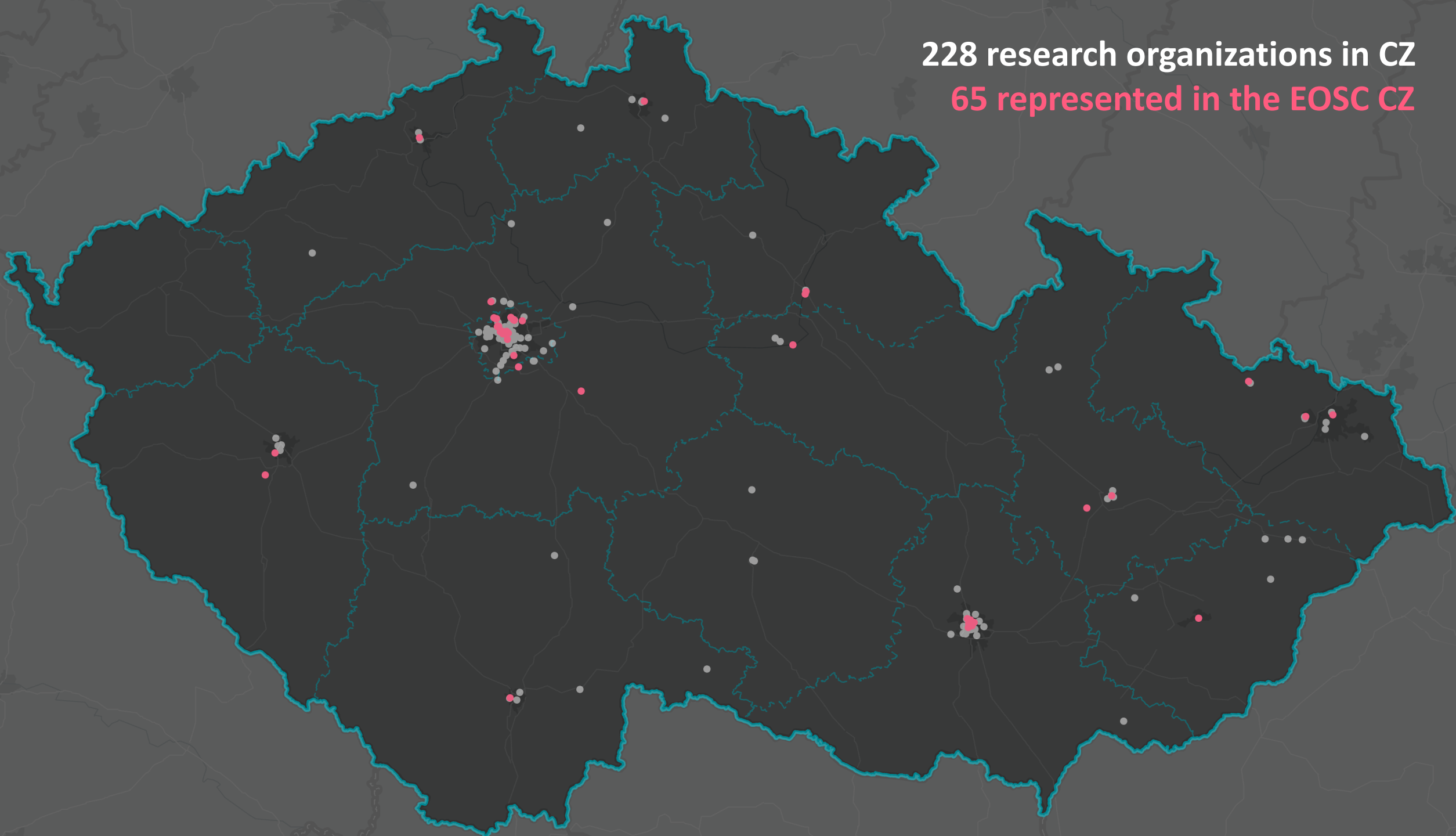


228 research organizations in CZ



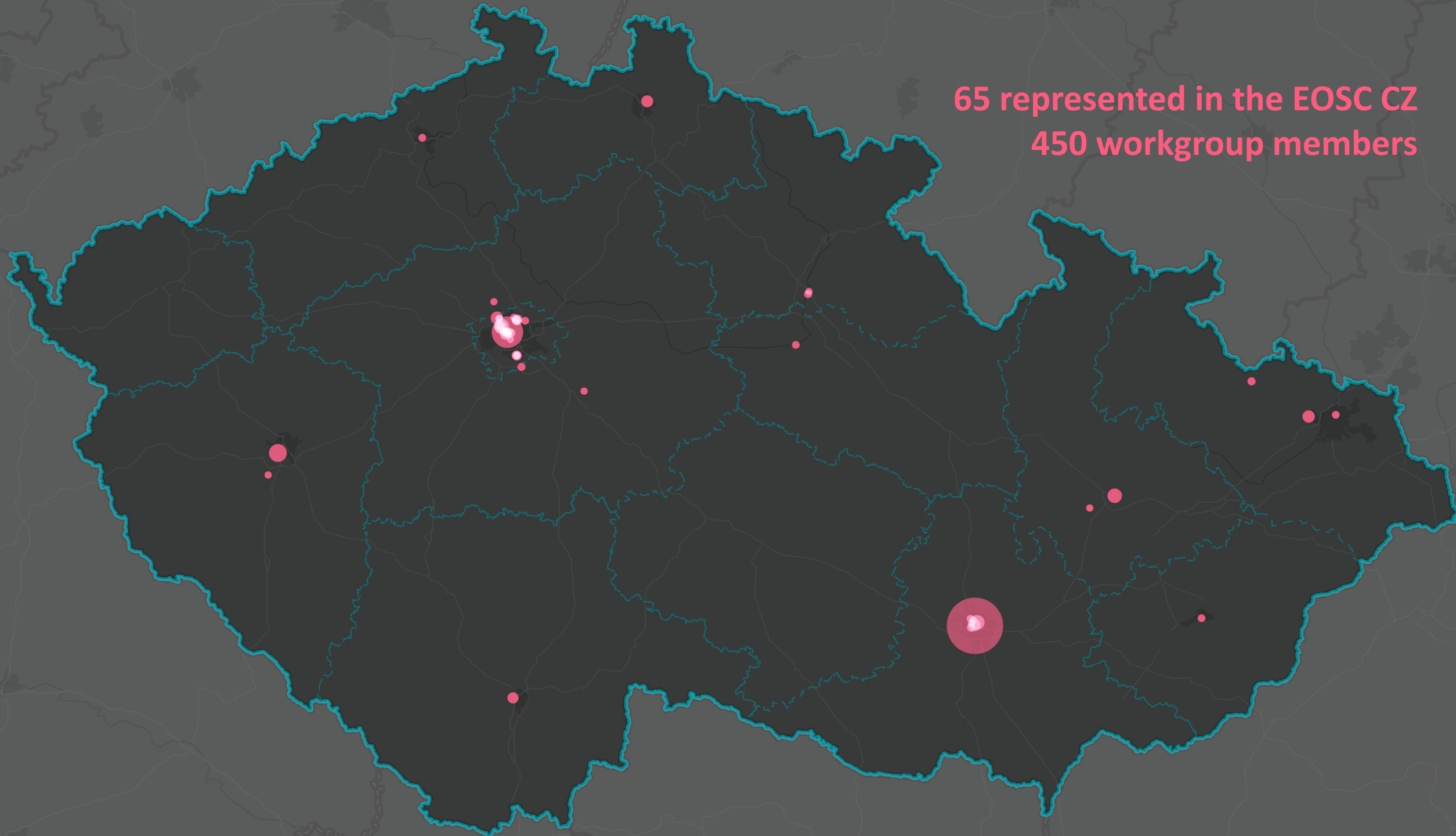
228 research organizations in CZ

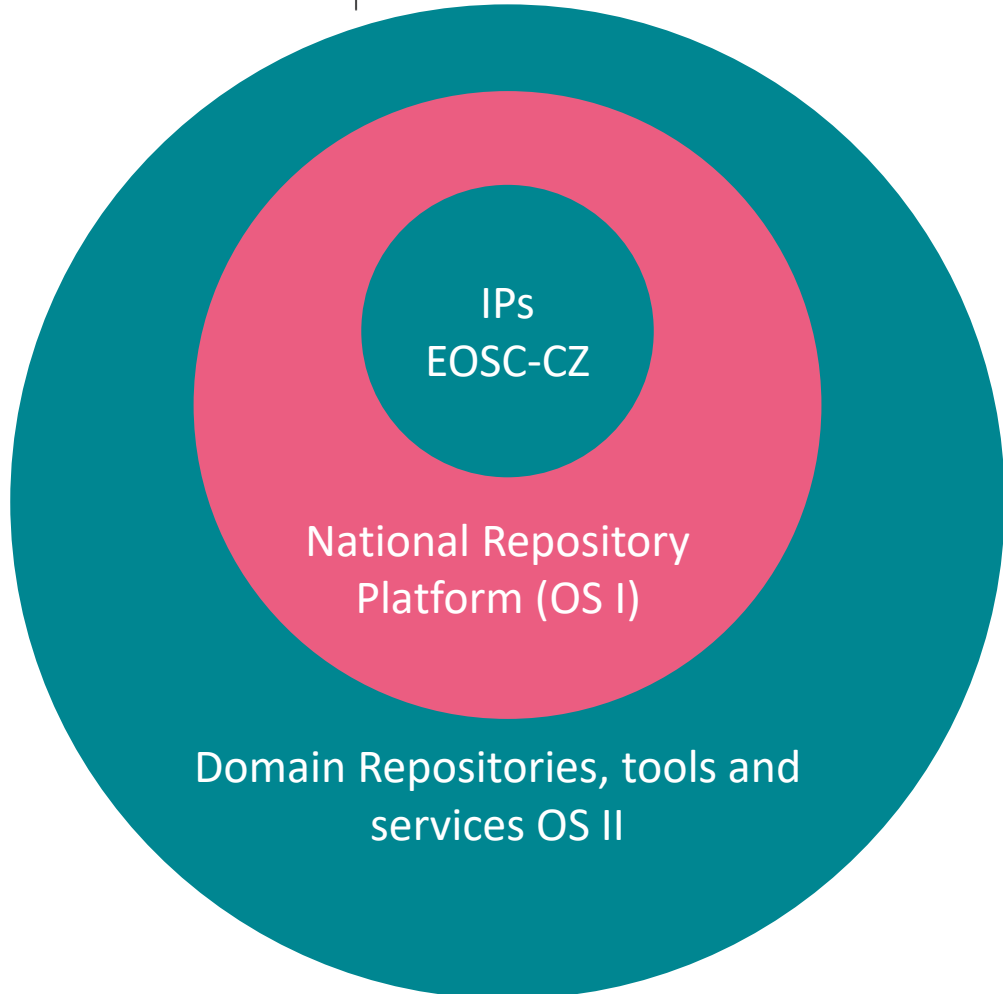
65 represented in the EOSC CZ





**65 represented in the EOOSC CZ**  
**450 workgroup members**





- IPs EOSC-CZ (since 2023) – Fundamentals for EOSC implementation in CZ
- Organizational (EOSC Secretariat) – <https://www.eosc.cz/en/secretariat>
  - Technical (National Metadata Directory) -- <https://nma.eosc.cz/>
  - Knowledge and skills (EOSC CZ Training Centre) -- <https://www.eosc.cz/en/training-centre>
  - Coordinated by Masaryk University, Brno

National repository Platform (OS I, since 2024) – “technical core”

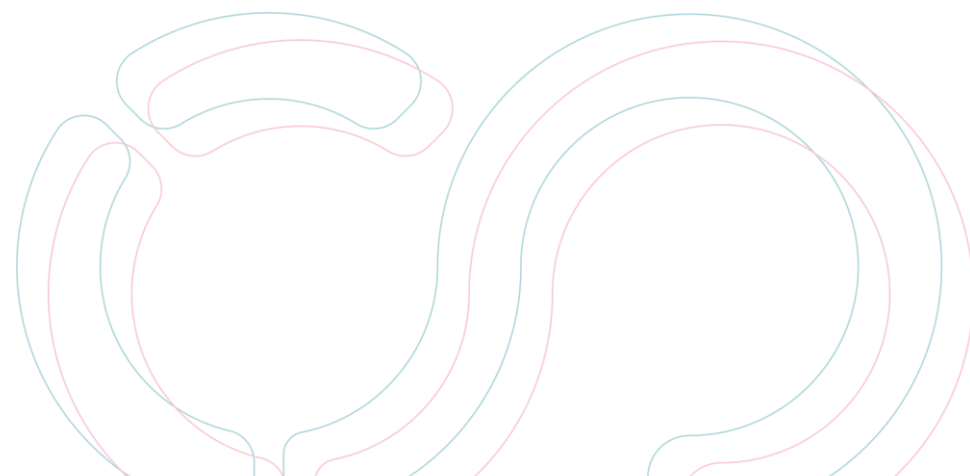
- Architecture of repository platforms (dspace, cesnet invenio, asepl arl) (50+ PB user capacity)
- First exemplary repositories
- Core services (PIDs, DSW, licenses, ...)
- Compliance and UX (cybersecurity, ServiceDesk, ...)
- Training – technical side of things
- Coordinated by CESNET (Czech NREN)

OS II (since 2025) – “domain specifics”

- Under preparation, content not clear yet
- Based on expertise of the 8 thematic / discipline workgroups
  - Bio/Health/Food, Matech, AI & ML, Social Sciences, Physics, Humanities & Arts, Enviro, Sensitive Data
- see <https://www.eosc.cz/en/working-groups>
- Coordinated by Charles University

OS III (???, since 2026???)

# Czech National Data Infrastructure





# On repositories

- Definition – information system for digital archivation, i.e. to provide storage, security, integrity, authenticity and access to digital documents in long-term.
- Practical take – data storage with metadata, access control and responsibility for the data
- Number of types
  - Institutional – e.g. for a specific university
  - Thematic / Domain – e.g. Protein Data Bank
  - Generic (catch-all) – e.g. Zenodo
  - National, ...
- Number of attractive properties
  - Typically free, secure, licence-respecting, available, indexed, possibly with additional services supporting FAIR principles
  - Trustworthiness
    - i.e. CoreTrustSeal, Nestor Seal, ISO
    - guarantees regarding licences, policies, etical standards, integrity, ...

# Infrastructure components

## • Repository platform

- Number of platforms -- CESNET Invenio, Clarin-DSpace, ASEP ARL, Islandora
- Total of 50+ PB of user data storage capacity
- Offered to research communities to create and operate specific repositories

## • Services

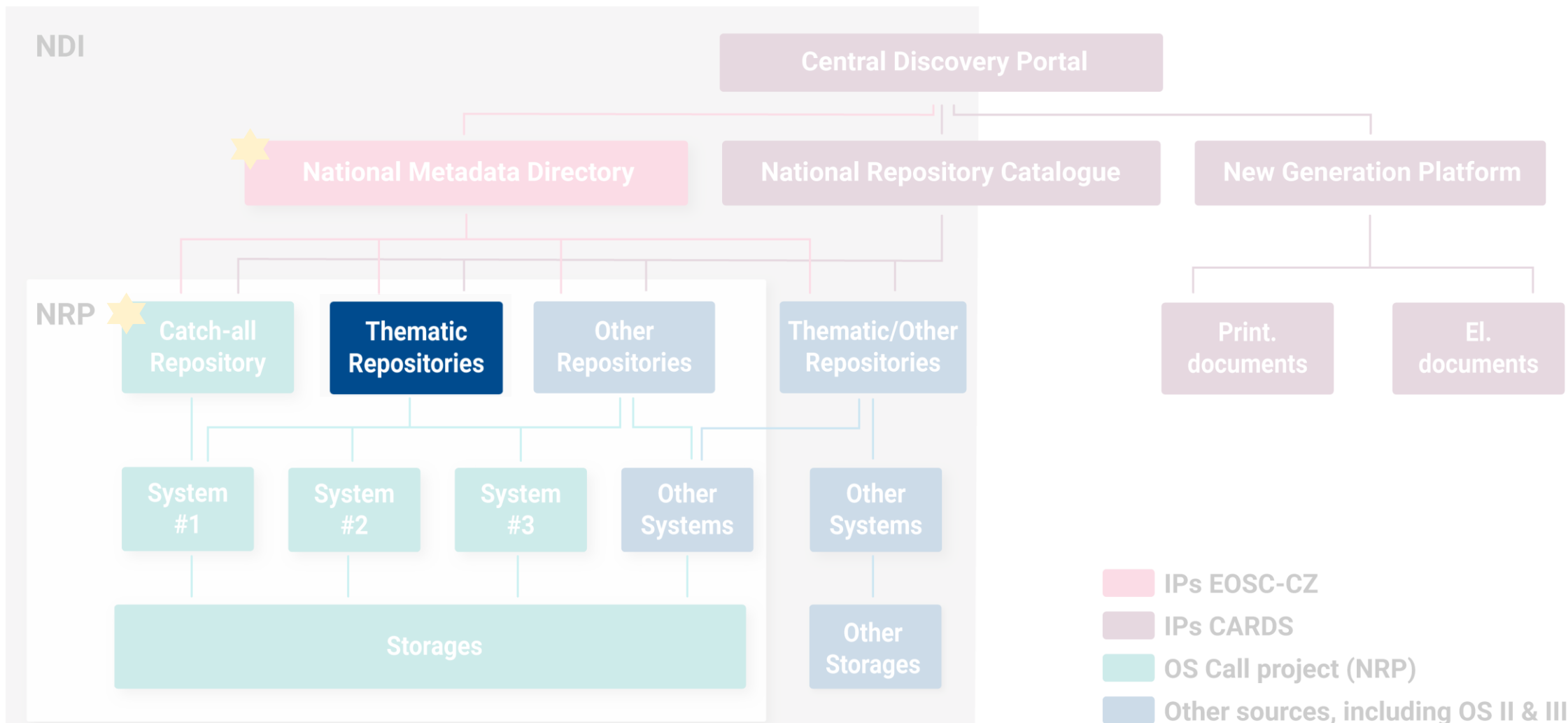
- Support for Data Management Planning
- Support for persistent identifiers
- AAI
- FAIRificator
- Interfaces to computing environments for analysis
- ...

## • Training and education

- EOSC Training centre
- CZ Data Steward community
- Methodological support for platforms, DMP and more



# National Data Infrastructure (NDI)



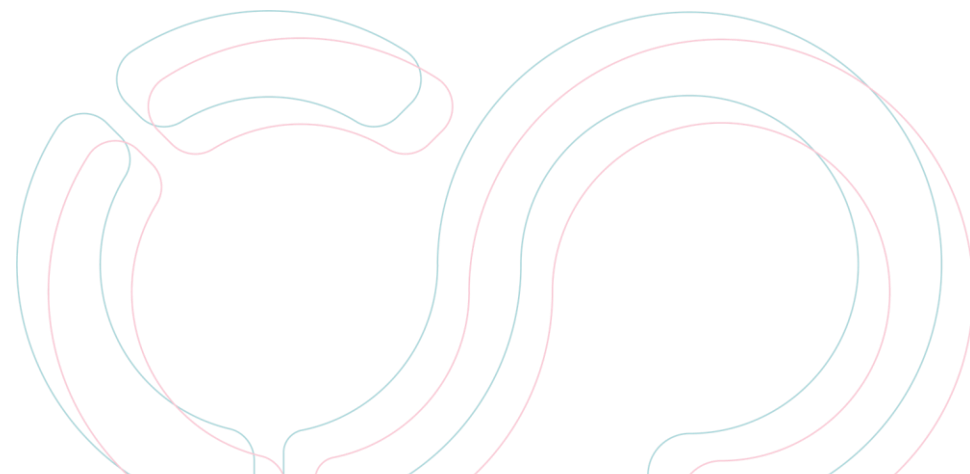


- ○ Physical sciences and engineering
- ○ Energy
- ○ Environmental sciences
- ○ Health and food
- ○ Social sciences and humanities
- ○ e-Infrastructures

● ELI Beamlines



# e-INFRA CZ



# e-INFRA CZ

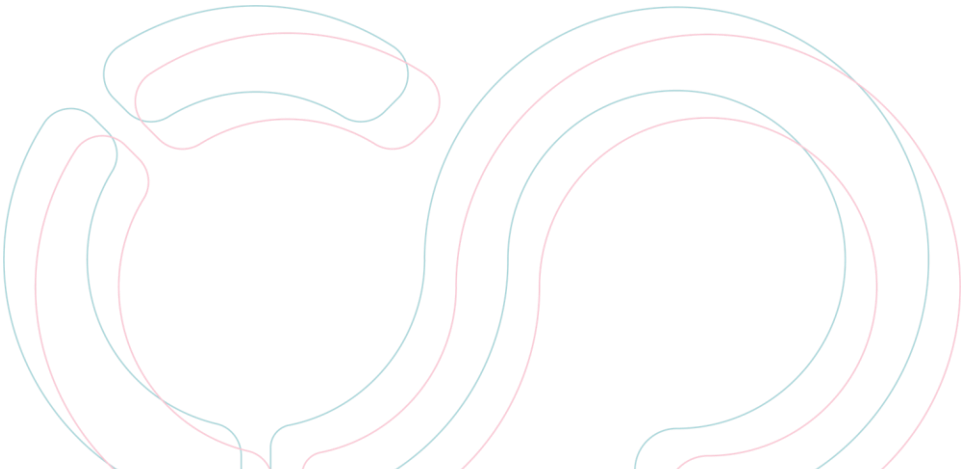
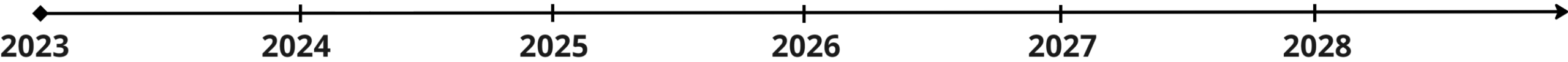
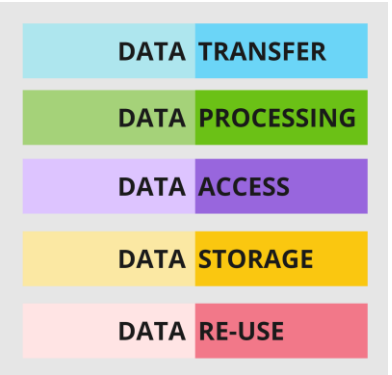
Single national e-INFRASTRUCTURE, consortium of three:

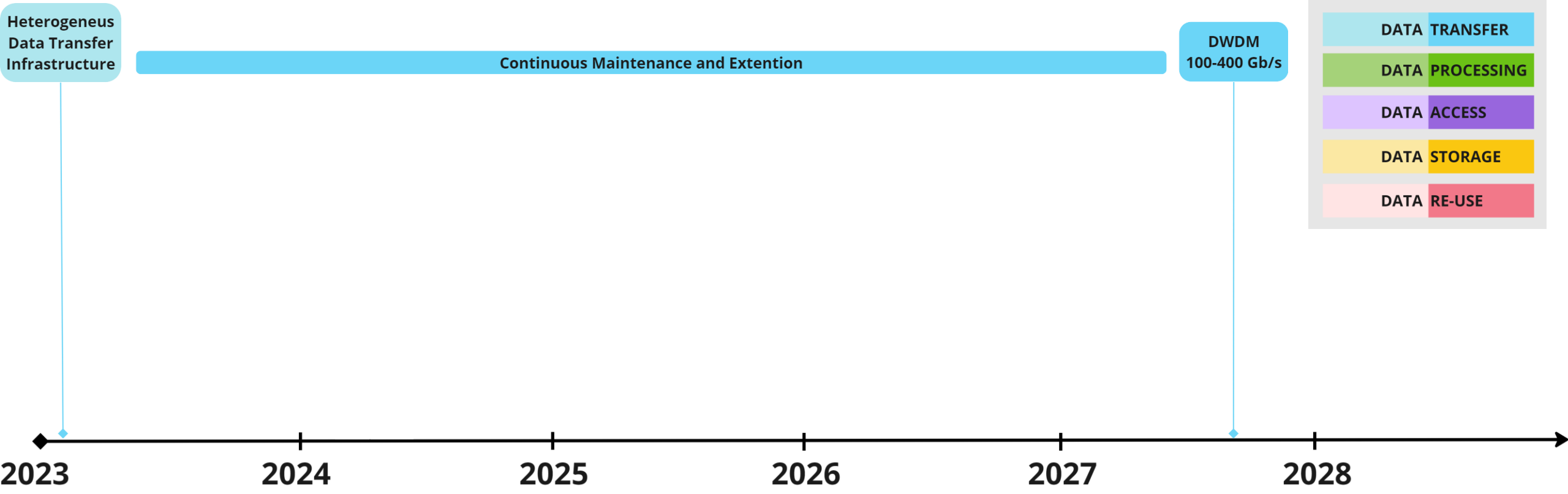
**CESNET** is an association of universities and the Academy of Sciences of the Czech Republic, which operates and develops the national e-infrastructure for science, research, and education, including a computer network, computational grids, data storage, collaboration environments, and offers a wide range of services.

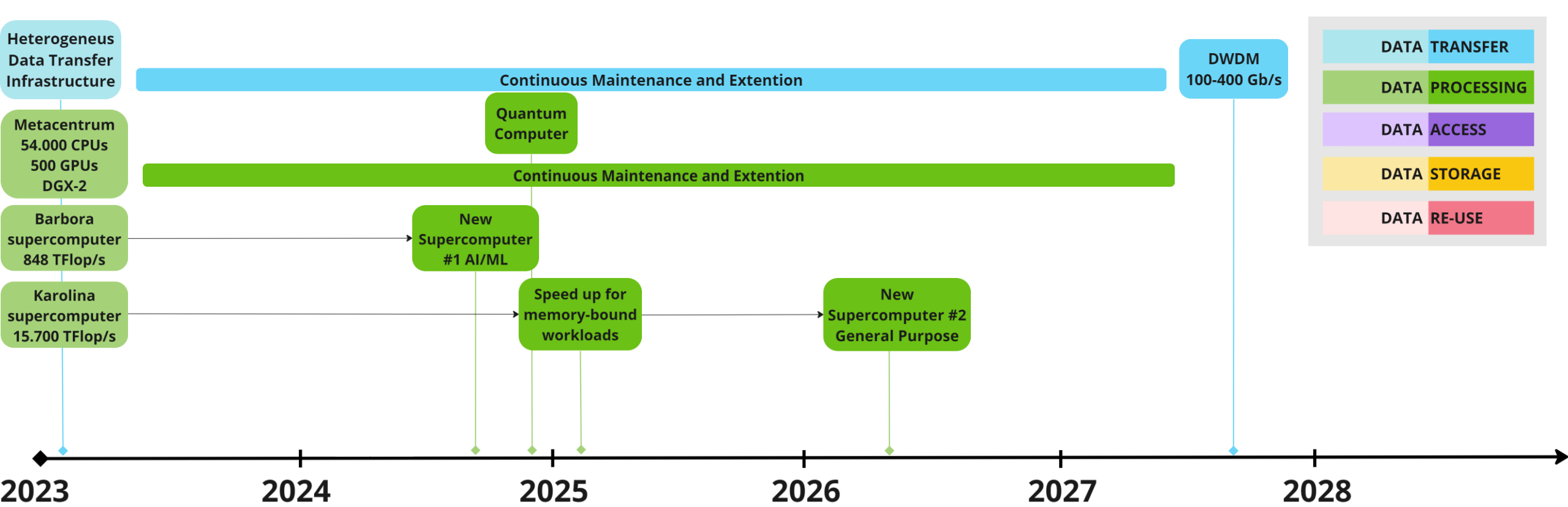
**IT4Innovations National Supercomputing Center** at **VSB – Technical University of Ostrava** is a leading research, development, and innovation center in the field of high-performance computing (HPC), data analysis (HPDA), artificial intelligence (AI), quantum computing (QC), and their applications in other scientific, industrial, and societal fields, operating the most powerful supercomputing systems in the Czech Republic.

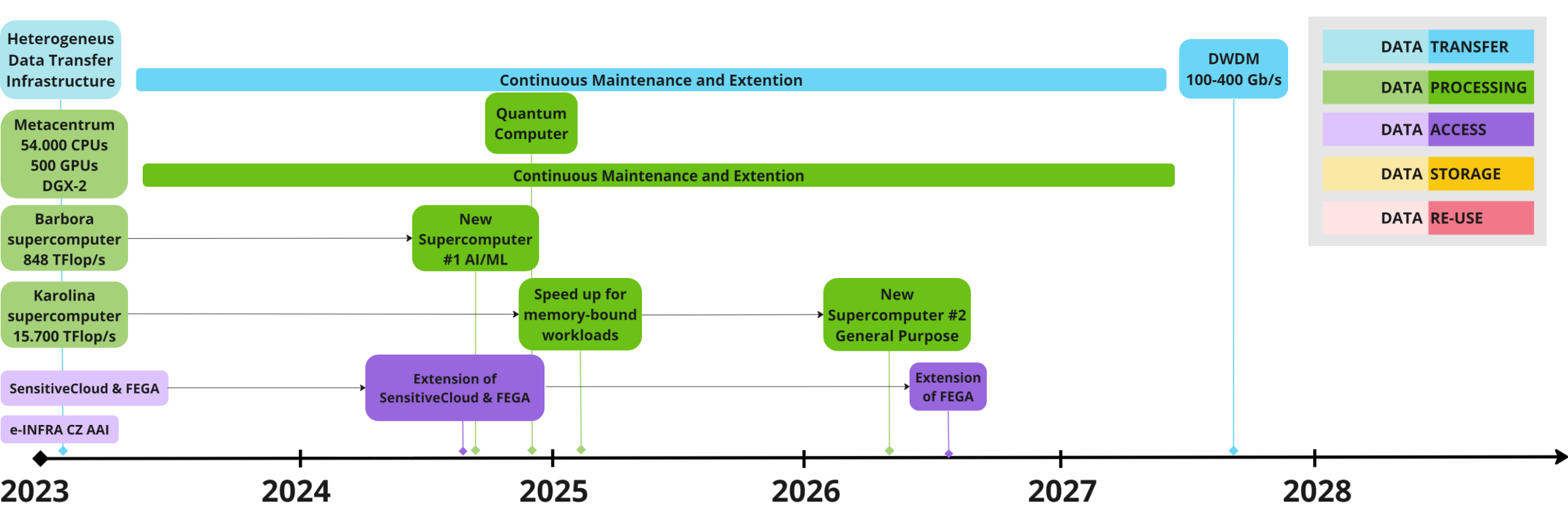
**CERIT-SC** at **Masaryk University** is a national center operating computational and data infrastructure for research and development.



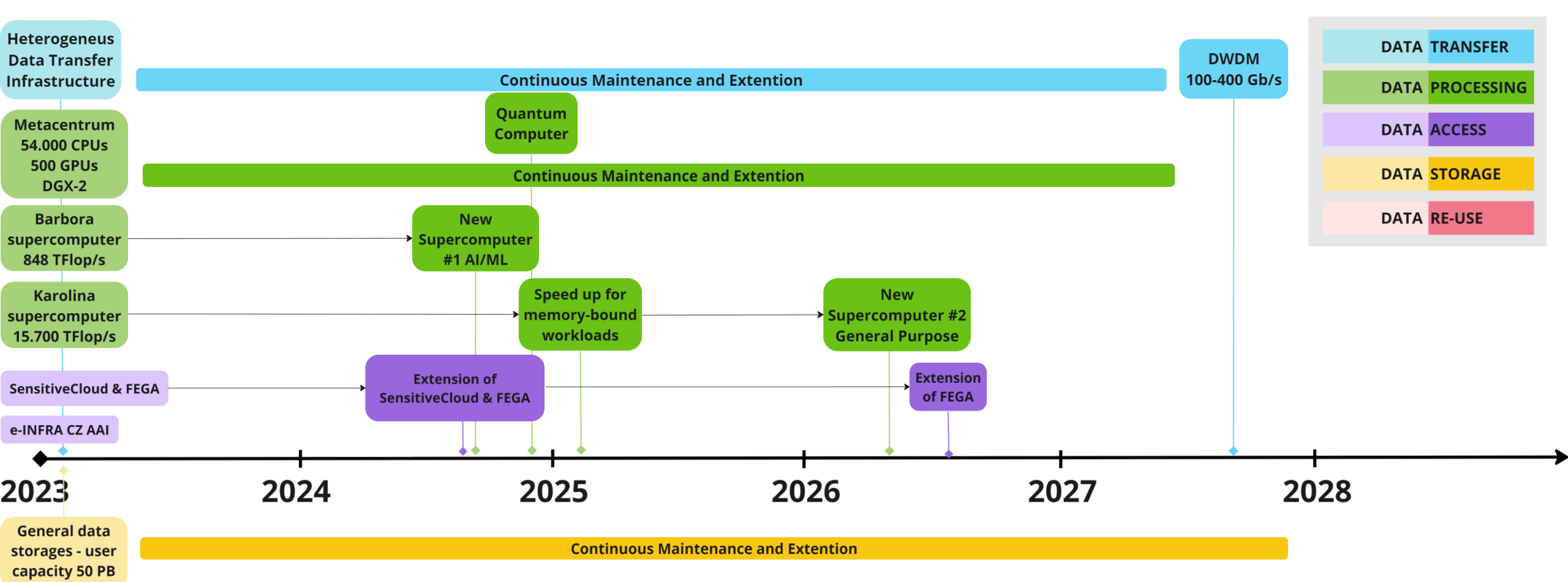


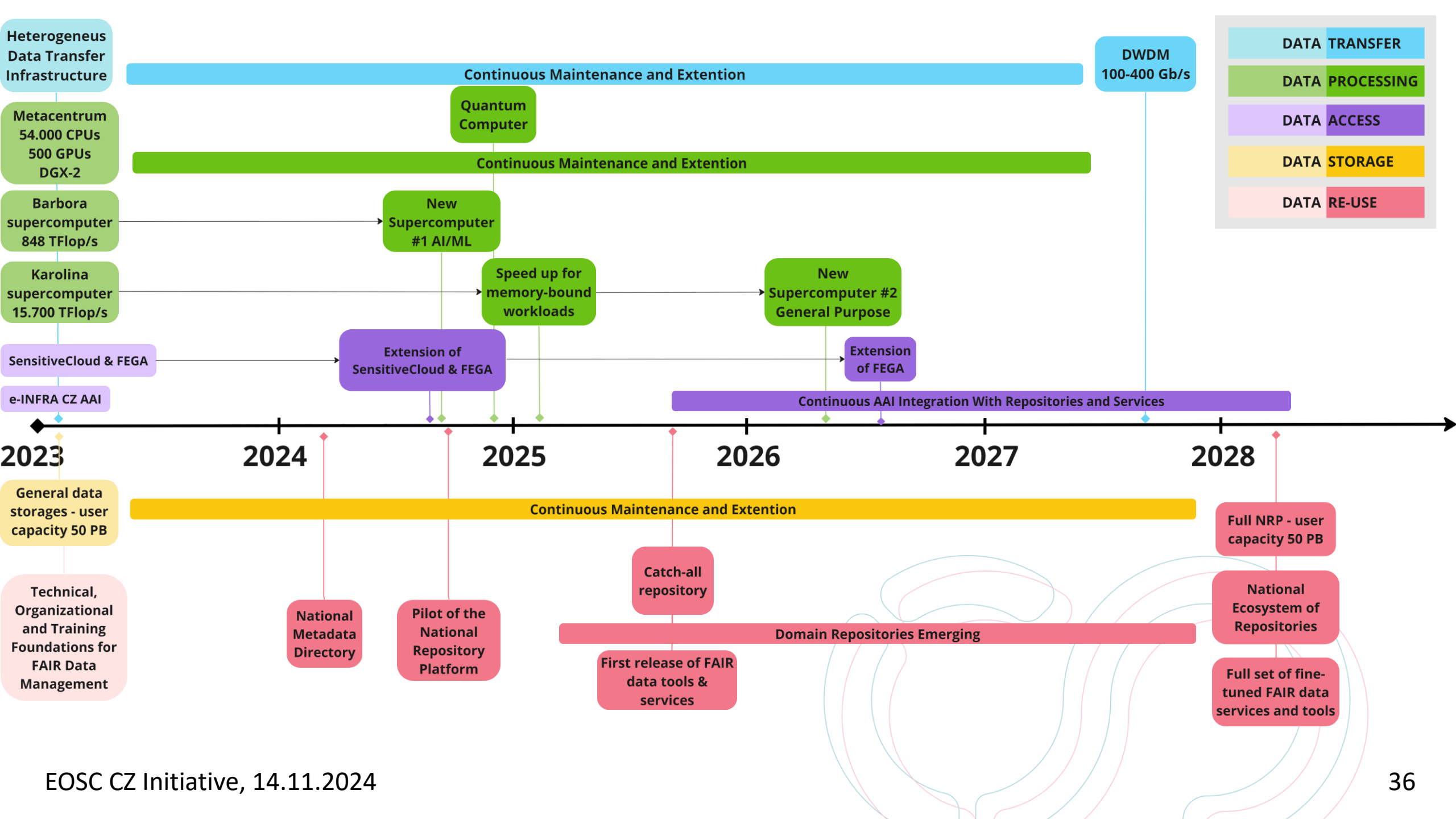




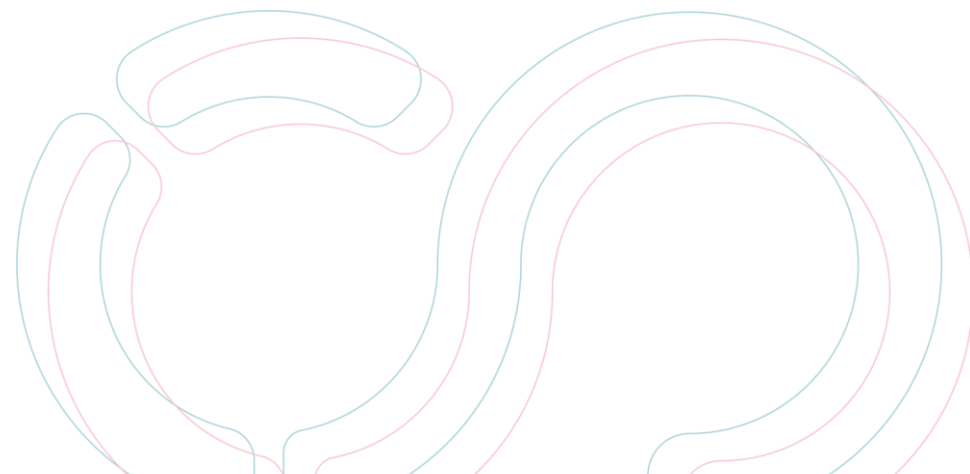








# First glimpses at the EOSC CZ services



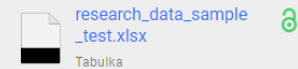
# National "catch-all pilot" Repository

## 🔗 Měření průměrné teploty v Praze: vzorový detailní záznam

Licence:



Soubory:



Identifikátor objektu:

DOI: 10.48700/datst.7n287-zc761

Stav záznamu:

Veřejný

V komunitě:

General community

Překlad názvu: [angličtina](#) Average temperature measurement in Prague: example of detailed recordTvůrci: [Vyčítalová, Hana](#) [Černošlávková, Petra](#) (manažer dat)

Datum zveřejnění: 2022-01-03

Datum vytvoření datové sady: 2018-07-01/2018-07-25

Datum sběru dat: 2018-04-01/2018-06-30

Jazyk: čeština, angličtina

Vydavatel: Národní technická knihovna

Klíčová slova: [cs](#) teplota [cs](#) klima [cs](#) Praha [cs](#) Česká republikaOborové kategorie: [Přírodní vědy](#) || [Vědy o Zemi a související environmentální vědy](#) || [Meteorologie, vědy o atmosféře](#) || [Výzkum klimatu](#)Abstrakt: [čeština](#) | [angličtina](#)

Měření průměrné teploty v Praze (hlavní město České republiky) během dubna, května a června 2018.

Metodologie: [čeština](#) | Teplota byla měřena každou hodinu a z hodnot byla vypočítána průměrná denní teplota.Technické informace: [čeština](#)

K zaznamenání byl použit měřicí přístroj, hodnoty byly zaznamenány ve stupních Celsia. Hodnoty byly zpracovány pomocí excelové tabulky. Zobrazení dat nevyžaduje speciální software.

Poznámky:

Vazby na/z dalších zdrojů:

**Název:** Collecting Grey Literature – Institutional Repository versus National Aggregator | **Autoři:** Černošlávková, Petra; Vyčítalová, Hana | **Rok:** 2018 | **DOI:** 10.26069/greynet-2018-000.009-gg

Projekt: ID34F57 | Evropská agentura pro životní prostředí

Práva: Creative Commons Uveďte původ 4.0 Mezinárodní licence





- Files up to 500 GB
- Supports DOI for easy referencing



# National Metadata Directory



- Single point of contact for research data – uniform format and metadata


National Czech Programme  **Národní Metadatový Adresář** Čeština


[← ZPĚT NA VÝSLEDKY HLEDÁNÍ](#)  


Vydáno: 15. 12. 2011


## Air Traffic Control Communication


Lidé	Šmídl, Luboš
Vloženo	None
Jazyk	eng
Vydavatel	University of West Bohemia, Department of Cybernetics
Typy zdroje	Other , corpus
Témata	<input type="text" value="speech corpus"/> <input type="text" value="acoustic model"/>
Alternativní identifikátory	<b>ID</b> ZCU_CZ_ATC  <b>HANDLE</b> <a href="http://hdl.handle.net/11858/00-097C-0000-0001-CCA1-0">http://hdl.handle.net/11858/00-097C-0000-0001-CCA1-0</a> 
Abstrakt	Corpus contains recordings of communication between air traffic controllers and pilots. The speech is manually transcribed and labeled with the information about the speaker (pilot/controller, not the full identity of the person). The corpus is currently small (20 hours) but we plan to search for additional data next year. The audio data format is: 8kHz, 16bit PCM, mono.



Identifikátory objektu 


 [Originální záznam](#)

Exportovat 


 [JSON](#)

API Odkazy 

 [API souborů](#)  
 [Tato položka](#)

Citace 

Šmídl, Luboš. (2011). Air Traffic Control Communication [Data set]. University of West Bohemia, Department of Cybernetics.

Style  





# Identifikátory CZ – Portal for Persistent

Identifikátory CZ

 **identifikatory.cz**

Stránky o perzistentních identifikátorech

 English

[Identifikátory](#) ▾ [Služby](#) ▾ [O nás](#) [Novinky](#)



[Domů](#) / [Identifikátory](#)

## Identifikátory

Zjistěte víc o jednotlivých perzistentních identifikátorech (PID). Perzistentní identifikátory jsou nástroje, které slouží k jednoznačné identifikaci osob, organizací a dalších objektů (např. knih, článků, datových sad) v systému vědecké komunikace.

[ORCID pro osoby](#)

[DOI pro objekty](#)

[ISBN pro knihy](#)

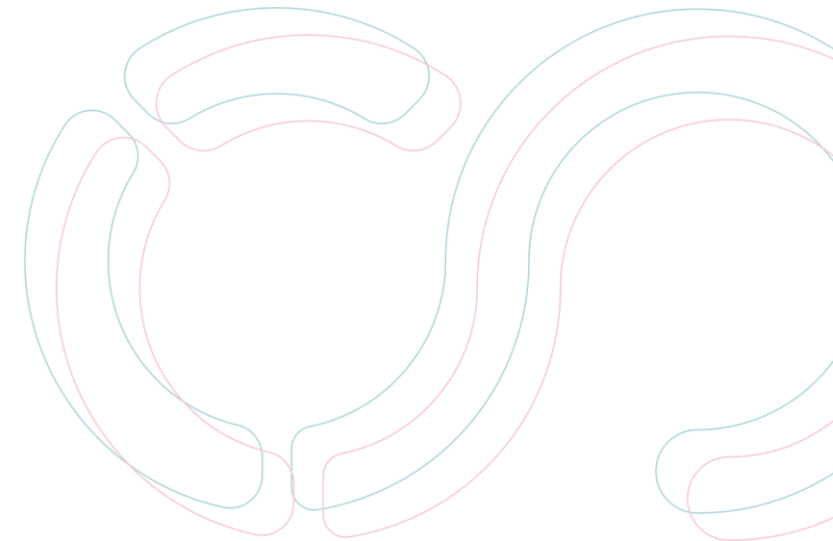
[ISSN pro časopisy](#)

[ISMN pro hudebniny](#)

[ROR pro organizace](#)

[IGSN pro vzorky](#)

[Další identifikátory](#)



# Sensitive Cloud

## Environment for processing sensitive data

- Virtual desktop
- Computing resources
- Secure applications
- Storing, sharing and cooperation on sensitive data
- VPN, Kubernetes



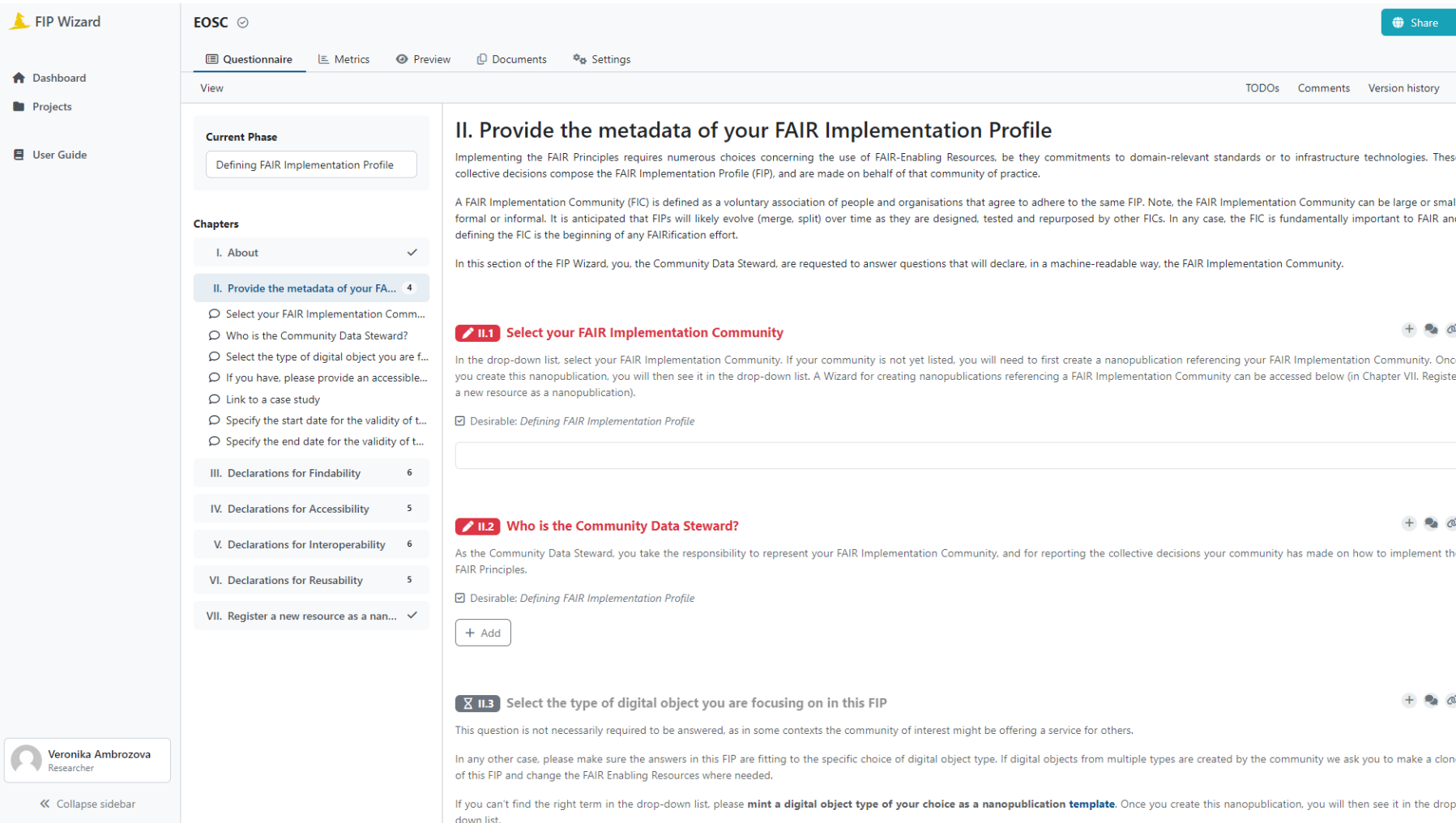
# Data Stewardship Wizard

- Comprehensive Tool for Data Management Planning

The screenshot displays the 'My Experiment' interface. On the left is a sidebar with navigation options: DS Wizard, Users, Knowledge Models, Projects (List, Importers), Documents, and Settings. The main content area is titled 'My Experiment' and includes tabs for Questionnaire, Metrics, Preview, Documents, and Settings. Below these are 'View' and 'Import answers' options. The 'Current Phase' is set to 'Before Submitting the Proposal'. A 'Chapters' list on the left shows 'IV. Processing data' selected. The main content area displays the 'IV. Processing data' section, which includes a question: '1 Will you be using a shared working space to work with your data?'. Below the question are buttons for 'Horizon 2020 DMP', 'Horizon Europe DMP', and 'Science Europe DMP'. The question has two radio button options: 'a. No, participants in the project each have different collections of data and tools' and 'b. Yes'. A comment from 'Albert Einstein' is visible, asking 'Will this work space be run by dedicated specialists?'. The interface also shows a 'Comments' sidebar on the right with a 'Share' button and a 'Version history' section.



# FIP Wizard (fair implementation profile)



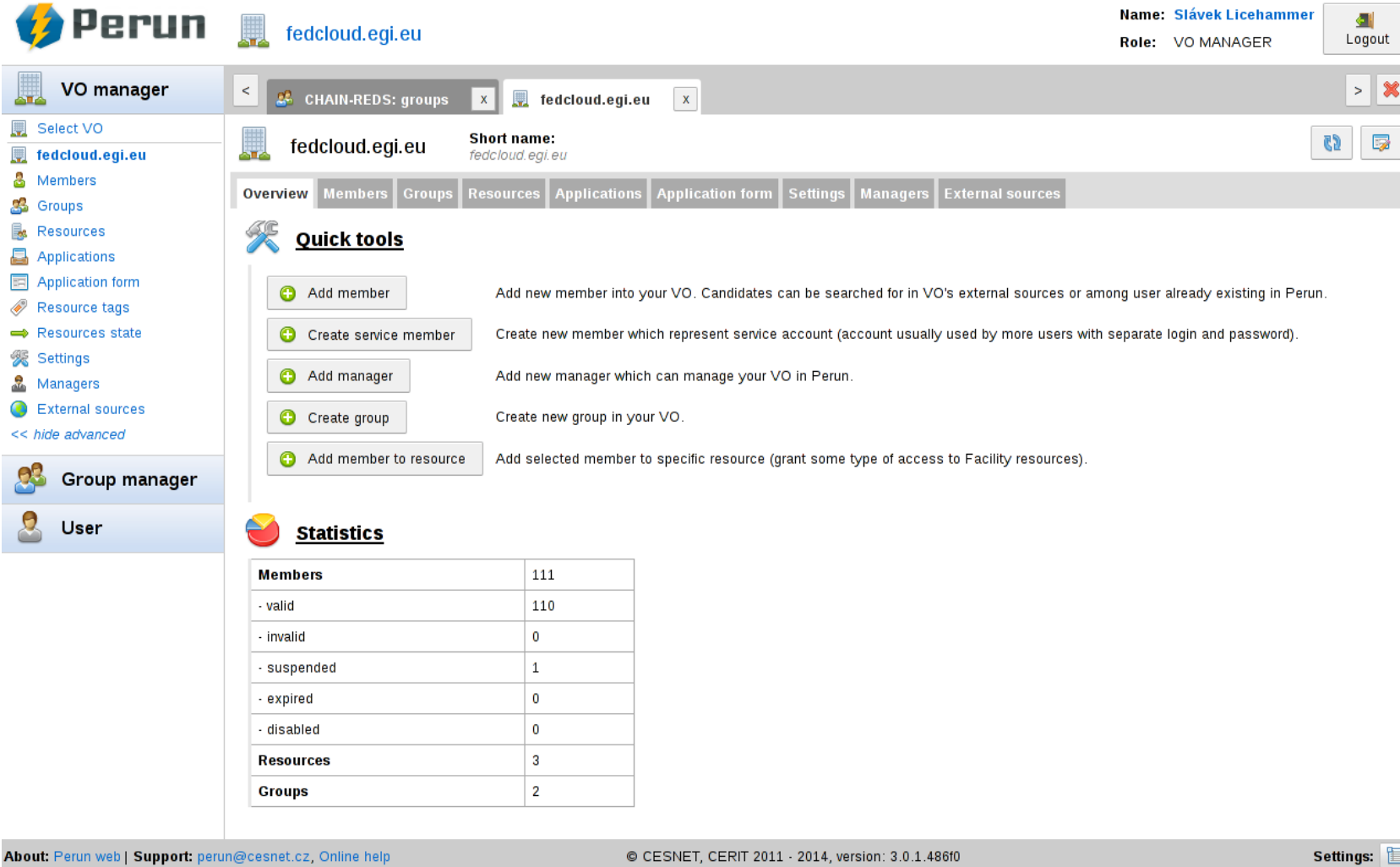
The screenshot shows the FIP Wizard interface. The left sidebar contains navigation links: FIP Wizard, Dashboard, Projects, and User Guide. The main content area is titled 'EOSC' and includes a 'Share' button. The current phase is 'Defining FAIR Implementation Profile'. The 'Chapters' list on the left shows 'II. Provide the metadata of your FAIR Implementation Profile' as the active section. The main content area displays the title 'II. Provide the metadata of your FAIR Implementation Profile' and a description: 'Implementing the FAIR Principles requires numerous choices concerning the use of FAIR-Enabling Resources, be they commitments to domain-relevant standards or to infrastructure technologies. These collective decisions compose the FAIR Implementation Profile (FIP), and are made on behalf of that community of practice.' Below this, there is a paragraph defining a FAIR Implementation Community (FIC) and a note that the user, as the Community Data Steward, is requested to answer questions that will declare, in a machine-readable way, the FAIR Implementation Community. The interface is divided into three sub-sections: 'II.1 Select your FAIR Implementation Community', 'II.2 Who is the Community Data Steward?', and 'II.3 Select the type of digital object you are focusing on in this FIP'. Each sub-section has a 'Desirable: Defining FAIR Implementation Profile' checkbox and a '+ Add' button. The 'II.1' section has a text input field. The 'II.2' section has a text input field. The 'II.3' section has a text input field. The bottom left corner shows the user profile for Veronika Ambrozova, Researcher, and a 'Collapse sidebar' button.

- Maps FAIR-ness of scientific conduct **interactive questionnaire**

- Generates **nanopublications**, interoperable between communities



# Authentication and authorization infrastructure (AAI)



**Perun** fedcloud.egi.eu

Name: **Slávek Licehammer** Logout  
Role: VO MANAGER

VO manager

CHAIN-REDS: groups x fedcloud.egi.eu x

fedcloud.egi.eu Short name: fedcloud.egi.eu


Overview Members Groups Resources Applications Application form Settings Managers External sources

**Quick tools**

- + Add member Add new member into your VO. Candidates can be searched for in VO's external sources or among user already existing in Perun.
- + Create service member Create new member which represent service account (account usually used by more users with separate login and password).
- + Add manager Add new manager which can manage your VO in Perun.
- + Create group Create new group in your VO.
- + Add member to resource Add selected member to specific resource (grant some type of access to Facility resources).

**Statistics**

<b>Members</b>	111
- valid	110
- invalid	0
- suspended	1
- expired	0
- disabled	0
<b>Resources</b>	3
<b>Groups</b>	2

About: [Perun web](#) | Support: [perun@cesnet.cz](mailto:perun@cesnet.cz), [Online help](#) © CESNET, CERIT 2011 - 2014, version: 3.0.1.486f0 Settings: 

To allow users from different universities easy access to data and services.

- Access and identity management
- Management of groups and roles
- Rights' delegation
- System integration





## EOSC CZ: Towards the development of Czech national ecosystem for FAIR research data

Matej Antol, Jiří Marek, Michaela Capandová, Jaroslav Juraček, and Luděk Matyska

**Abstract**—This short paper presents a compact overview of the Czech approach to implementing the European Open Science Cloud and plans for developing a Czech national infrastructure for FAIR research data. Its purpose is to provide an all-encompassing summary of the near future of research data management in Czechia. As such, we deliberately attempt to explain complicated concepts in minimum words, sacrificing the precision of expression for compactness.

**Index Terms**—EOSC, EOSC CZ, FAIR data, National Data Infrastructure, National Repository Platform, Open Science

### I. INTRODUCTION

THE importance of data in research is continuously rising, while approaches to store, manage and share these data seem to fall behind. The value of the data is reduced by their considerable heterogeneity and lack of structure, which leads to low reproducibility and hinders scientific progress. Open Science (OS) [1] seeks to address some of these current issues, focusing on data availability and sharing, urging for more collaboration and emphasising research integrity. European Open Science Cloud (EOSC) [3], [4] is an international initiative that builds on the Open Science principles. EOSC seeks to create a common European research environment [5] to store, share and re-use research data and other digital objects without barriers. We call such data and objects FAIR [2] (Findable, Accessible, Interoperable, Reusable).

### II. EOSC CZ - INFRASTRUCTURE AND SERVICES FOR FAIR RESEARCH DATA

The establishment of fundamental principles for the Czech national EOSC implementation took place in 2021, resulting in the document called *Architecture of EOSC Implementation in the Czech Republic* [6]. The document represents the official start of the EOSC CZ initiative [7]. The primary tangible outcome of this initiative will be a National Repository Platform (NRP) - a core component of the National Data Infrastructure (NDI). NRP will be a federated ecosystem of distinct technological layers (see Fig. 1) and associated services (see below).

The data infrastructure will complement the existing Czech national e-infrastructure e-INFRA CZ [11] with all its services. NDI will be fully integrated at the European level [12]. NRP will interconnect with the already running parts of NDI: data repositories and services held at universities, Czech Academy

Authors are with the CERIT-SC centre, Institute of Computer Science, Masaryk University, 60200 Brno, Czechia. Contact at info@eosc.cz  
Manuscript published February 20, 2024

of Sciences and Research Infrastructures. Examples are environments such as LINDAT/CLARIN-CZ [13] for natural language processing, Czech-BioImaging [14] for biological and medical imaging or EIRENE RI [15] for human exposure. Next to the repositories themselves, the initiative plans to deploy and integrate several FAIR data-related services designed for NDI users. Notably:

- Central Discovery Portal (CDP) integrated into the New Generation Platform (PNG) will ensure the searchability and availability of all types of resources (electronic, digitized and printed) and research results.
- National Metadata Directory to search in NDI metadata.
- Single Authentication and Authorization Infrastructure (AAI) solution Perun [16] to guarantee data accessibility.
- Support for data management planning via Data Stewardship Wizard [17].
- Support for Persistent Identifiers (PIDs) [18].
- Support for data FAIRification.
- Data mgmt. tools such as OneData [19] or iRODS [20].
- Training [21] and university courses on data management.

### III. ACTIVE COMMUNITIES AND HOW TO PARTICIPATE

Researchers' engagement is vital for the EOSC CZ's success. Since 2021, as a reaction to the EOSC CZ Architecture document, 12 EOSC CZ working groups [22] have been established through a self-organizing community effort. These groups will be operational during the entire EOSC CZ initiative, and registration is continuously open to new potential members. A list of their members is publicly available. Currently, the initiative is in its initial implementation phase, and the active participation of scientists in the working groups

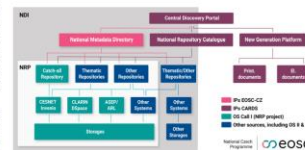


Fig. 1. NDI and NRP blueprint with five abstraction layers. Bottom-up: hardware infrastructure distributed across Czechia; three initial repository systems - CESNET Invenio [8], CLARIN-DSpace [9] and ASEP/ARL [10]; specific domain and other repositories, metadata directory and on top, Central Discovery Portal.

is the main guarantor for the NDI ecosystem to encompass and support all relevant research data management needs of research communities.

- The initiative is also closely connected with the National Open Science coordination team within the National Library of Technology. On top of that, collaboration is being established with the already existing national Open Science communities:
- Open Science working groups of the Association of Libraries of Czech Universities,
- national Data Steward Community and
- members of the institutional Open Science centres within Czech academic institutions.

### IV. HOW TO BENEFIT FROM THE EOSC CZ OUTCOMES

The NDI's ecosystem of services will be offered to the whole research community regardless of their active participation in the EOSC CZ initiative. The EOSC CZ Secretariat [23] and Training Centre [21] are already operational, providing consultancy, seminars and workshops for the Czech research ecosystem. The National Metadata Directory will be deployed in 2024, followed by the NRP with a portion of core services in 2025. By this time, the first domain and other repositories should also be emerging. This first phase will be completed in 2026, with an entire NRP and its services available. The initiative will concurrently foster the development of data management and other related skills for all Czech academia members. It will also encourage the systemic formation of data steward and curator roles across the academic ecosystem.

With this infrastructure, any reasonably interested Czech scientist should have sufficient information, know-how, skills, institutional support, and services to store, share, and reuse research data efficiently. These ambitions summarize the main objective of the EOSC CZ initiative.

### ACKNOWLEDGMENTS

The EOSC CZ initiative has active collaborators who significantly exceed the authors of this paper. Out of these, we would namely like to acknowledge the contributions of Radka Římanová, Klára Slanařová, Petra Čermohávková, Martin Svoboda, Miroslav Bartošek, David Antoš and Michal Růžička.

### APPENDIX: FINANCIAL SUPPORT FOR EOSC IN CZECHIA

Czech Ministry of Education, Youth and Sports (MŠMT) supports the EOSC CZ initiative [24] via two systemic projects and three open science calls:

- Individual Systemic Project (IPs) EOSC-CZ, coordinated by Masaryk University with two additional partners, supported with 18 mil. EUR to provide a fundamental organizational, technical, and training environment.
- IPs CARDS, coordinated by National Library of Technology, supported with 56 mil. EUR, to provide support for PIDs, research data description, and deliver the PNG.
- OS Call I, with an allocation of 50 mils. EUR, to create the NRP, its core services and related training.
- OS Call II, with an allocation of 36 mil. EUR to support domain-specific data management, repositories and related services over the NRP.
- OS Call III, scope of which is currently under discussion.

### REFERENCES

- [1] Manafó, M., Nosek, R., Bishop, D. et al. A manifesto for reproducible science. *Nat Hum Behav* 1 (2017). doi.org/10.1038/s41562-016-0021
- [2] Wilkinson, M.D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1), pp.1-9. (2016)
- [3] <https://eosc-portal.eu/>
- [4] <https://eosc.eu>
- [5] <https://eosc.eu/stra-mar/>
- [6] [https://www.msmt.cz/uploads/311/Architektura\\_Implementace\\_EOSC\\_CZ.pdf](https://www.msmt.cz/uploads/311/Architektura_Implementace_EOSC_CZ.pdf)
- [7] <https://www.eosc.cz/en>
- [8] <https://github.com/CESNET>
- [9] <https://github.com/clarin-dspace>
- [10] <https://asep-portal.lib.cas.cz/basic-information/dataset-repository/>
- [11] <https://www.e-infra.cz/en>
- [12] <https://eosc.eu/participate-collaboration/czech-republic/>
- [13] <https://lindat.cz/>
- [14] <https://www.eosc-bioimaging.cz/>
- [15] <https://www.eirene-ri.eu/>
- [16] <https://perun-as.org/>
- [17] <https://ds-wizard.org/>
- [18] <https://identifiers.cz/en/>
- [19] <https://www.cerit-sc.cz/management-of-data-workflows>
- [20] <https://irods.org/>
- [21] <https://www.eosc.cz/en/training-centre>
- [22] <https://www.eosc.cz/en/working-groups>
- [23] <https://www.eosc.cz/en/secretariat>
- [24] <https://www.dokc.eu/en/ostatnosti-za-analyzy/seznam-operaci-jerjencu>



**Matej Antol** is the principal project manager of the IPs EOSC-CZ. He is also the integration manager of the Czech e-infrastructure e-INFRA CZ and an executive director of one of its three partners, the CERIT-SC infrastructure. He has a long background in IT and research projects. His research activities focus on managing and analysing complex, high-dimensional data.



**Jiří Marek** is the General Secretary of the EOSC CZ initiative and head of the EOSC CZ Secretariat. He holds the role of the Open Science manager at Masaryk University and serves as a head of the CZARMA Open Science Task Force. He is also involved with activities regarding digitization of the public sector via open technologies (Open Cities, etc.)



**Michaela Capandová** is the secretary to the EOSC CZ Working Groups Metadata and Materials Sciences and Engineering. Her research in the biomedical field is focused on the development of cellular elements and biomaterials for lung tissue engineering. She loves electrospraying and scanning electron microscopy.



**Jaroslav Juraček** is the secretary to the EOSC CZ Working Group BioHealth/Food. Beyond that, he takes part in building the European Genomic Data Infrastructure and related activities at the national level. His focus is set on advancing open science initiatives and access to and utilization of genomic data for research and innovation.



**Luděk Matyska** is a full professor at the Faculty of Informatics, Masaryk University, with a long track in developing national and European research infrastructures. He is the director of the CERIT-SC, one of three members of the e-INFRA CZ steering board, the principal project manager of the NRP project, and chairman of the IPs EOSC-CZ steering committee.

Implementace European Open Science Cloud (EOSC) v České republice si klade za cíl vytvořit národní uzel této evropské iniciativy a podpořit dobrou praxi v oblasti správy výzkumných dat například vědeckými komunitami. V rámci implementace **Národní datové infrastruktury (NDI)** vznikne společná platforma pro sdílení, správu a přístup k datům a výpočetním zdrojům pro výzkumné účely. NDI bude podporovat jak vědecké, tak multidisciplinární výzkumné aktivity a bude zahrnovat širokou škálu vědních oborů a disciplín.

visit <https://www.eosc.cz/en>

read <https://arxiv.org/pdf/2402.13343>

# Thank you for your attention

