

EOSC a rozvoj národní infrastruktury pro výzkumná data v České republice

Jiří Marek, vedoucí Sekretariátu EOSC CZ

02. 06. 2025



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

MUNI
ICS

Obsah prezentace

- Hodnota výzkumných dat
- Národní datová infrastruktura
- První dostupné služby
- Výpočetní kapacity
- Podpora práce s daty
- Projekty iniciativy EOSC CZ
- Vzdělávání
- **Data veřejného sektoru vs. Výzkumná data**
- Co nás čeká dál

Hodnota výzkumných dat

Recenzní řízení výzkumu



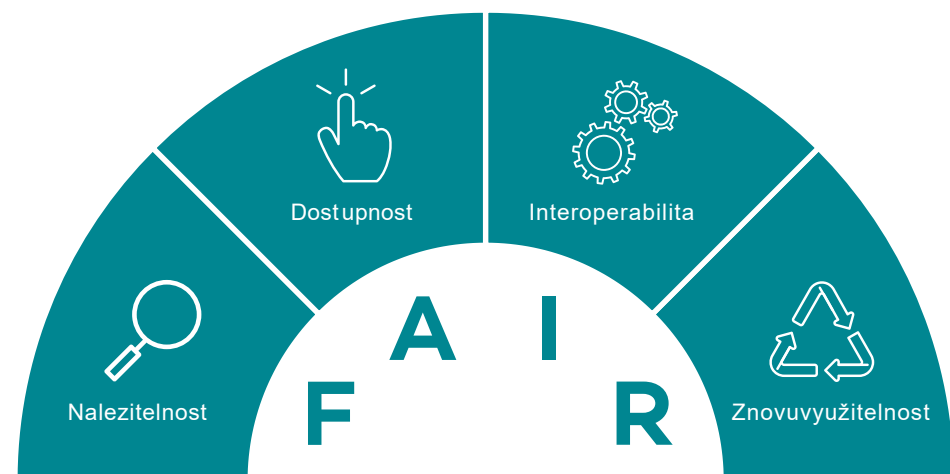
FAIR data

Findable – metadata, PIDs, repozitáře

Accessible – dostupnost, řízení přístupu

Interoperable – preferované formáty

Reusable – read me file, licence



Podpora FAIR v České republice



Konference, workshopy, setkání s komunitami

- **28 školení** a workshopů
 - cca **5000 registrací** ze 125 výzkumných institucí
- **4 konference** se 100+ účastníky
 - Osobní i fyzická účast
- **35 fyzických setkání** komunit a pracovních skupin
 - 3 roadshow v Praze, Brně, Ostravě
 - 8 EOSC CZ Networking – České Budějovice, Liberec, Olomouc, Zlín, AV ČR, Pardubice, Brno, Hradec Králové
- **+ více než 100 online setkání pracovních skupin**
- **Tisíce lidí jsou neustále v kontaktu**
 - 300+ sledujících na sociálních sítích
 - 300+ odběratelů EOSC CZ newsletteru
 - 4000+ zhlédnutí obsahu na YouTube

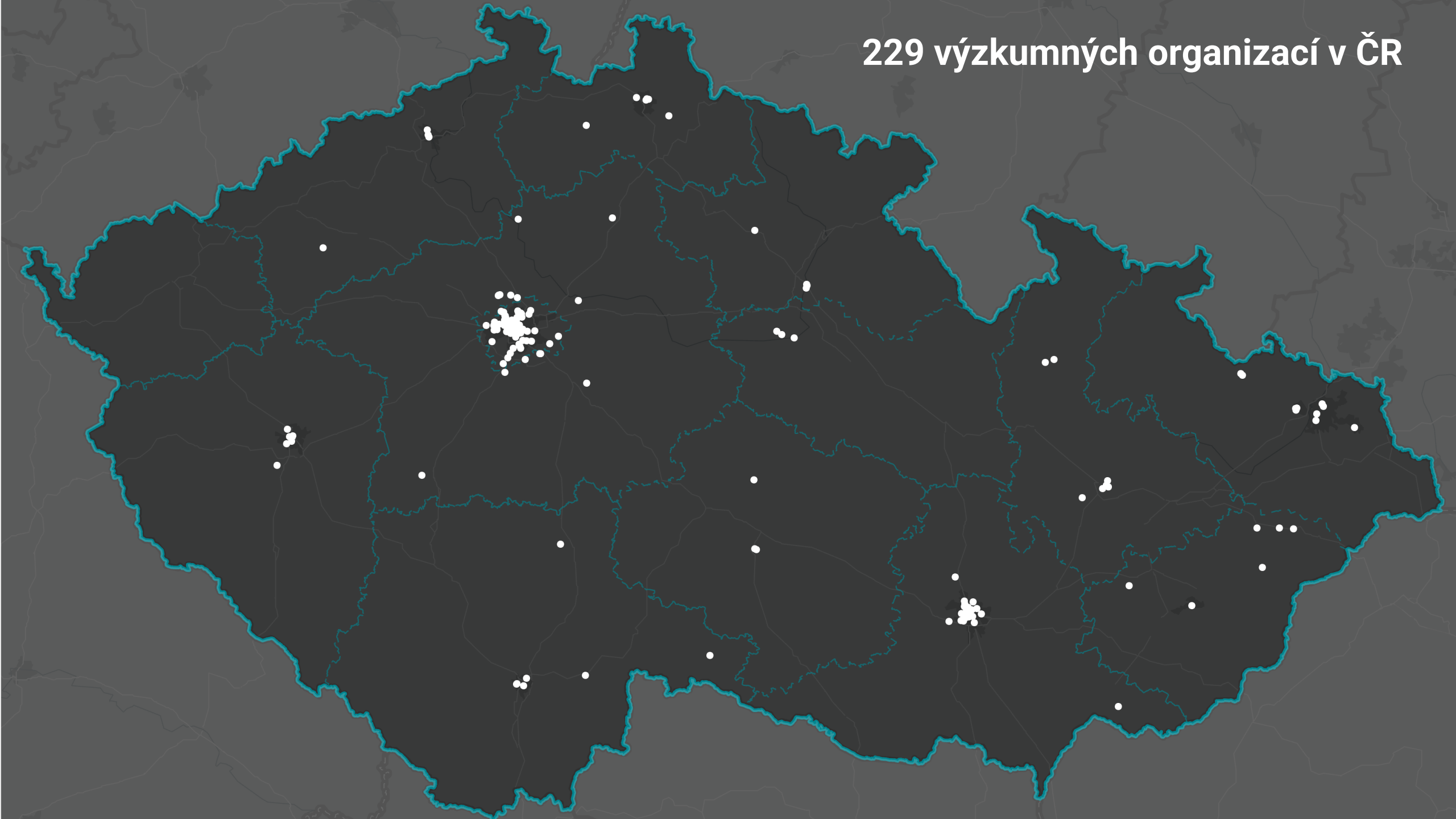


Události:

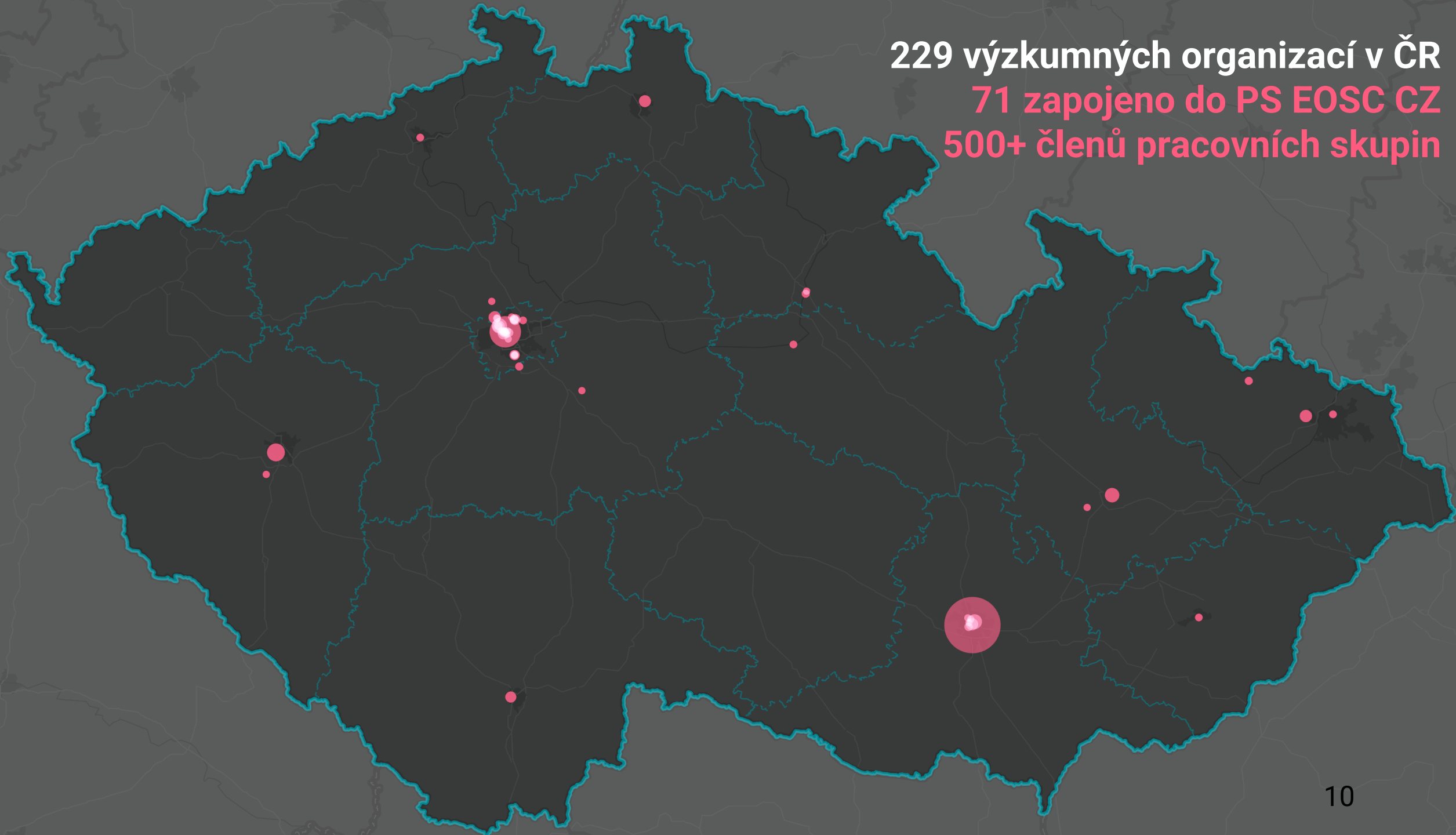
- Konference
- Setkání komunit
- Workshopy



229 výzkumných organizací v ČR



229 výzkumných organizací v ČR
71 zapojeno do PS EOSC CZ
500+ členů pracovních skupin



Vědci budují pro vědce



Architektura národní datové
infrastruktury (NDI)



Metadata



Základní služby



Vzdělávání a lidské zdroje



Bio/zdraví/potraviny



Materiálové vědy a technologie



Data management pro umělou
inteligenci a strojové učení



Sociální vědy



Fyzikální vědy



Humanitní vědy a umění



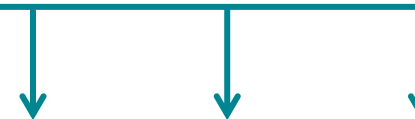
Environmentální vědy



Citlivá data

Pracovní skupiny EOSC CZ

- hlavní komunita EOSC
- 4 průřezové: vstupy do NRP
- 8 oborových: vstupy do OS II
- otevřené zájemcům (kdykoliv přijímá nové členy)
- Vědci pro vědce - tajemníci PS
- fungování po celou dobu trvání iniciativy



Otevřená platforma pro implementaci EOSC

Národní datová infrastruktura

Praktické výstupy

Úložné kapacity

Hardware

- Fyzická, distribuovaná úložná infrastruktura
- Uživatelská kapacita = **50 PB**, vysoká replikace dat
- Umístění až na 5 různých lokalitách

Repozitářové systémy

- CESNET Invenio, CLARIN-DSpace, ASEP-ARL

Repozitáře

- [Národní metadatový adresář](#) (2024)
- [Catch-all repozitář](#) (2025)
- **Oborové (tematické) repozitáře**
 - Piloty (4, OS I, od konce 2025)
 - Ostatní od 2025/2026

Nástroje a služby

- Podpora pro [data management plánování](#) (DMP)
- Správa **metadatových profilů**
- Podpora **práce s licencemi**
- Podpora práce s [persistentními identifikátory](#)
- Podpora pro **FAIRifikaci výzkumných dat**
- Automatizace sběru dat
- Elektronické laboratorní deníky
- Celková kybernetická bezpečnost systému a compliance

Výpočetní kapacity

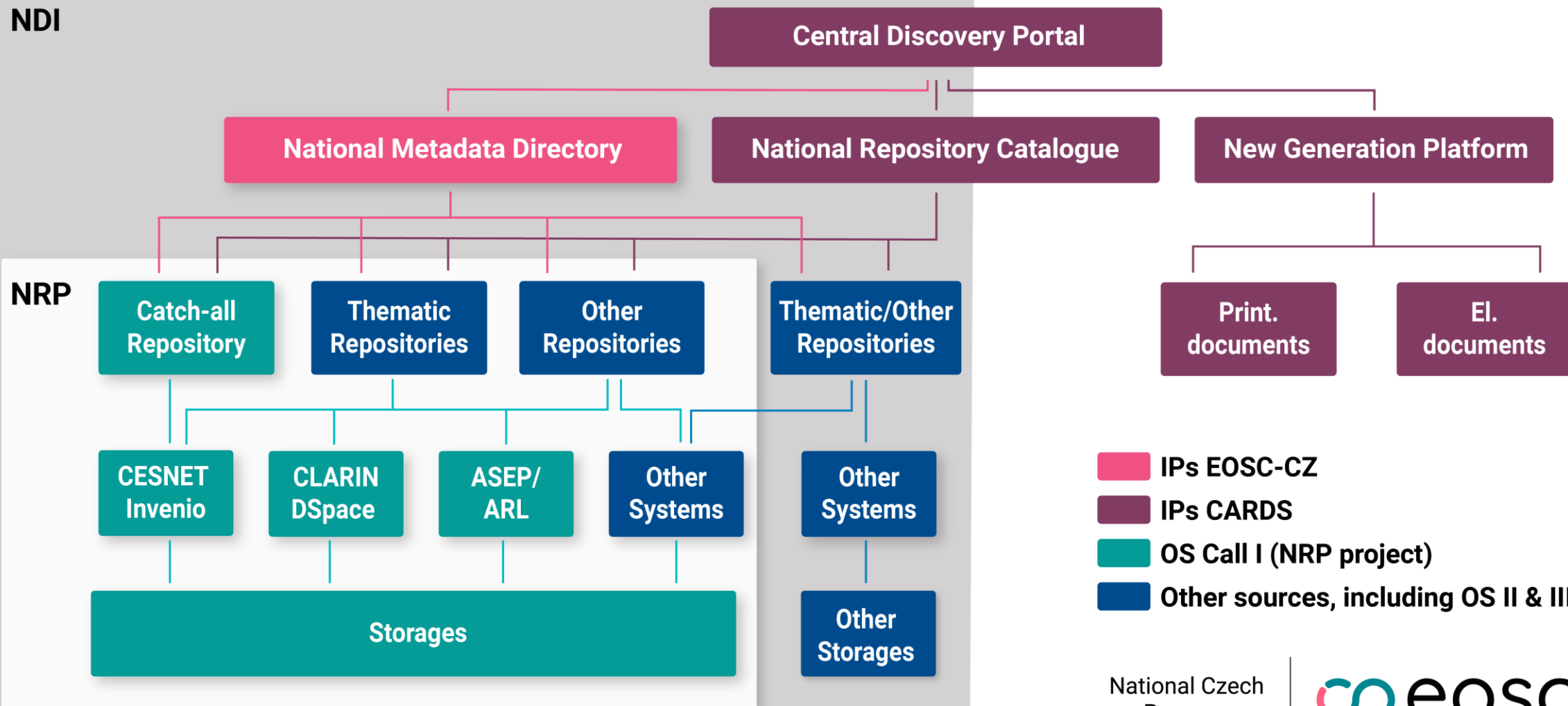
- Konsorcium e-INFRA:
 - CESNET: Metacentrum
 - CERIT-SC: Cloudová a gridová řešení
 - IT4I: Výkonné superpočítače



Národní datová infrastruktura

NDI

NRP



IPs EOSC-CZ

IPs CARDS

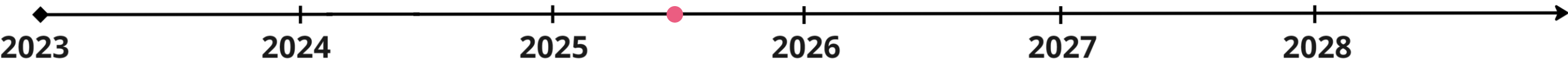
OS Call I (NRP project)

Other sources, including OS II & III

Konec první fáze

Fáze 1

- **Zapojení českých výzkumníků** a odborníků
- **Organizace pracovních skupin** a komunit
- **Kick-off 3 hlavních projektů**
- **Navržení** základní infrastruktury a služeb
- **Spuštění prvních** služeb
- **Zvýšení povědomí**



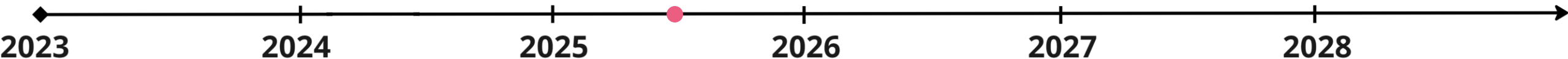
Začátek druhé fáze

Fáze 1

- Zapojení českých výzkumníků a odborníků
- Organizace pracovních skupin a komunit
- Kick-off 3 hlavních projektů
- Navržení základní infrastruktury a služeb
- Spuštění prvních služeb
- Zvýšení povědomí

Fáze 2

- Spuštění infrastruktury
- Spuštění všech základních služeb
- Integrace prvních repozitářů
- Podpora data management dovedností a know-how
- Integrace národního a mezinárodního ekosystému
- Návrh modelu udržitelnosti



Ukázky prvních funkčních služeb pro FAIR správu výzkumných dat

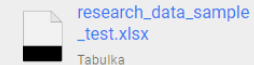
Catch-all repozitář (pilot)

🌱 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 record

Tvůrci: [Vyčítalová, Hana](#) [Černohlávková, 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 republika

Oborové 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.

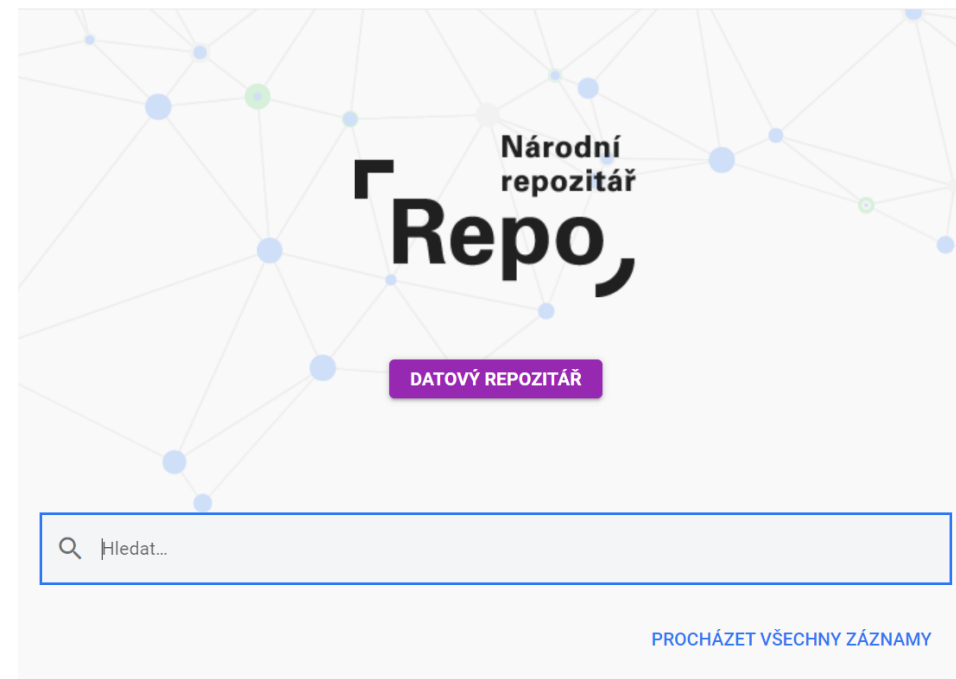
Poznámky:

Vazby na/z dalších zdrojů:

Název: Collecting Grey Literature – Institutional Repository versus National Aggregator | **Autoři:** Černohlávková, 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



Catch-all repozitář (pilot)

Nová datová sada

Datová sada

Název datové sady *

CS 

Toto pole je vyžadováno

 PŘIDAT PODNÁZEV

Jazyk *

 VYBRAT POJMY

Abstrakt *

CS           

Prostor pro detailní popis metodologie a technických informací naleznete v části „Popis datové sady“

Licence

 VYBRAT POJEM

Pokud se k datové sadě váže licence Creative Commons, vyberte příslušnou variantu licence z nabídky. Doporučujeme výběr nejnovějších verzí a to 3.0 Česko a 4.0 International.

Upozornění: Ujistěte se o tom, že datová sada může být pod licencí zveřejněna!

Vydavatelé *

 VYBRAT POJMY

POKRAČOVAT

 Tvůrci

 Popis datové sady

 Datace

 Informace o financování

 Vytvoření draftového záznamu

 Upload dat

Kontakt

support@narodni-repozitar.cz

Nápověda

www.du.cesnet.cz



Národní úložiště pro biodiverzitní data

Czech herbaria specimens repository

Search for a specimen in several formats, e.g. PRC 37 or PRC-000037...

Search



EOSC CZ

This repository is a part of the Czech
National Repository Platform [NRP](#)
[EOSC CZ](#)



Interoperability

Designed and maintained with
interoperability support for [JACO](#)
and [GBIF](#) as global aggregators.



Documentation

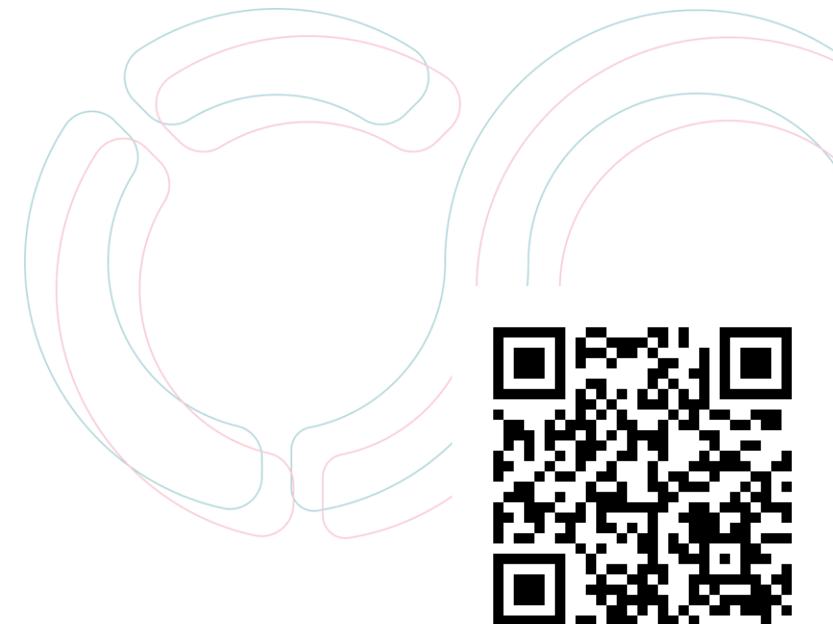
Please read the [documentation](#) to
understand how the repository
imports and publishes digital images.



Co-funded by
the European Union



National Czech
Programme



Databáze molekulární biofyziky



Documentation

Tutorials

Deposit

sign in with



Molecular Biophysics Database

Search

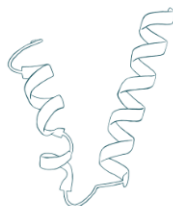


All records



Example: Lysozyme, NaCl, K_D, homo sapiens, xyhds-adf3t (record id)

[Advanced search](#)



Molecular Biophysics Database collects raw data produced in experiments with biomolecular samples, biological material or other material, using molecular biophysics methods, such as Microscale Thermophoresis (MST), Biolayer interferometry (BLI), Surface Plasmon Resonance (SPR) and others.

MBDB development is supported by the project **MOSBRI** - Molecular Scale Biophysics Research Infrastructure of the European Commission, no. 101004806

Currently supported techniques

MST
Microscale thermophoresis

BLI
Bio-layer interferometry

SPR
Surface plasmon resonance

ITC
Isothermal titration calorimetry



Co-funded by
the European Union



National Czech
Programme



Národní metadatový adresář

[← ZPĚT NA VÝSLEDKY HLEDÁNÍ](#)

Hledat



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 speech corpus acoustic model

Alternativní identifikátory

ID ZCU_CZ_ATC**HANDLE**<http://hdl.handle.net/11858/00-097C-0000-0001-CCA1-0>

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

APA



Identifikátory CZ



Persistent Identifiers

[Persistent Identifiers](#) ▾ [Services](#) ▾ [About us](#) [News](#)



[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.

**ORCID iD for
researchers**

DOI for objects

ISBN for books

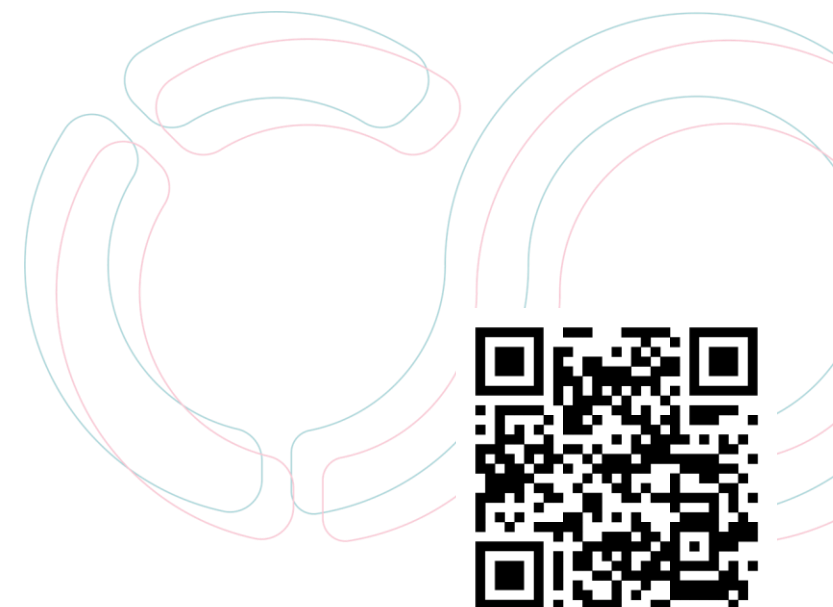
ISSN for periodicals

**ISMN for notated
music**

**ROR for
organizations**

IGSN for samples

Other PIDs



SensitiveCloud

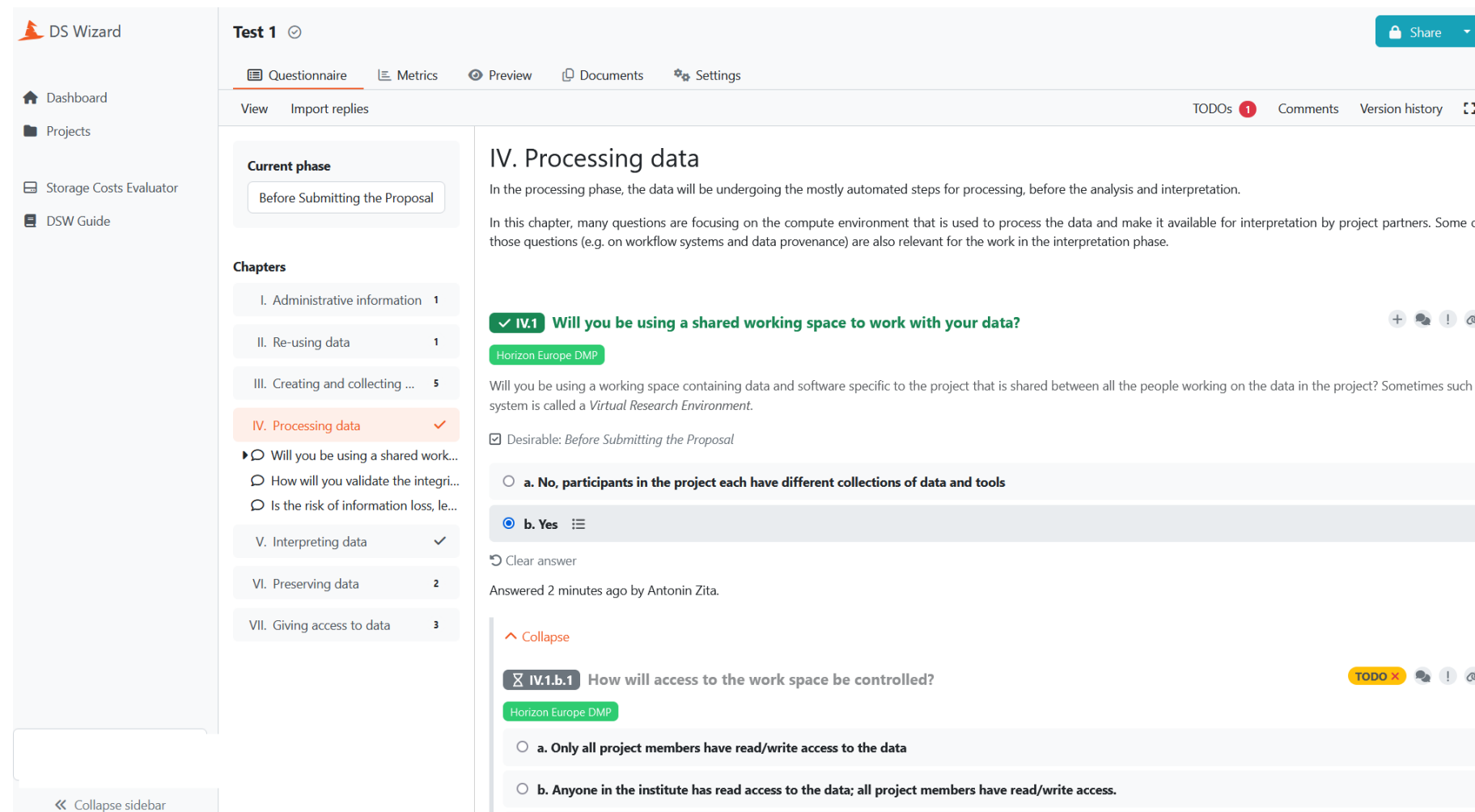
Prostředí pro zpracování citlivých dat

- Virtuální plocha
- Výpočetní zdroje
- Bezpečné aplikace
- Ukládání, sdílení a spolupráce
- VPN, Kubernetes



Data Stewardship Wizard

Nástroj na plánování správy
výzkumných dat



The screenshot shows the 'Data Stewardship Wizard' interface. On the left is a sidebar with navigation links: 'DS Wizard', 'Dashboard', 'Projects', 'Storage Costs Evaluator', and 'DSW Guide'. The main content area is titled 'Test 1' and includes a 'Share' button. Below the title is a navigation bar with 'Questionnaire', 'Metrics', 'Preview', 'Documents', and 'Settings'. The 'Questionnaire' section is active, showing a 'Current phase' of 'Before Submitting the Proposal'. A list of chapters is on the left, with 'IV. Processing data' selected. The main content area displays the 'IV. Processing data' section, which includes a description of the processing phase and a list of questions. The first question is 'IV.1 Will you be using a shared working space to work with your data?'. It has two options: 'a. No, participants in the project each have different collections of data and tools' and 'b. Yes'. The second question is 'IV.1.b.1 How will access to the work space be controlled?'. It has two options: 'a. Only all project members have read/write access to the data' and 'b. Anyone in the institute has read access to the data; all project members have read/write access.'.

DS Wizard

Test 1

Share

Questionnaire Metrics Preview Documents Settings

View Import replies

TODOs 1 Comments Version history

Current phase

Before Submitting the Proposal

Chapters

- I. Administrative information 1
- II. Re-using data 1
- III. Creating and collecting ... 5
- IV. Processing data ✓**
- V. Interpreting data ✓
- VI. Preserving data 2
- VII. Giving access to data 3

IV. Processing data

In the processing phase, the data will be undergoing the mostly automated steps for processing, before the analysis and interpretation.

In this chapter, many questions are focusing on the compute environment that is used to process the data and make it available for interpretation by project partners. Some of those questions (e.g. on workflow systems and data provenance) are also relevant for the work in the interpretation phase.

✓ IV.1 Will you be using a shared working space to work with your data?

Horizon Europe DMP

Will you be using a working space containing data and software specific to the project that is shared between all the people working on the data in the project? Sometimes such a system is called a *Virtual Research Environment*.

☒ Desirable: Before Submitting the Proposal

☐ a. No, participants in the project each have different collections of data and tools

☒ b. Yes

Clear answer

Answered 2 minutes ago by Antonin Zita.

^ Collapse

⌵ IV.1.b.1 How will access to the work space be controlled?

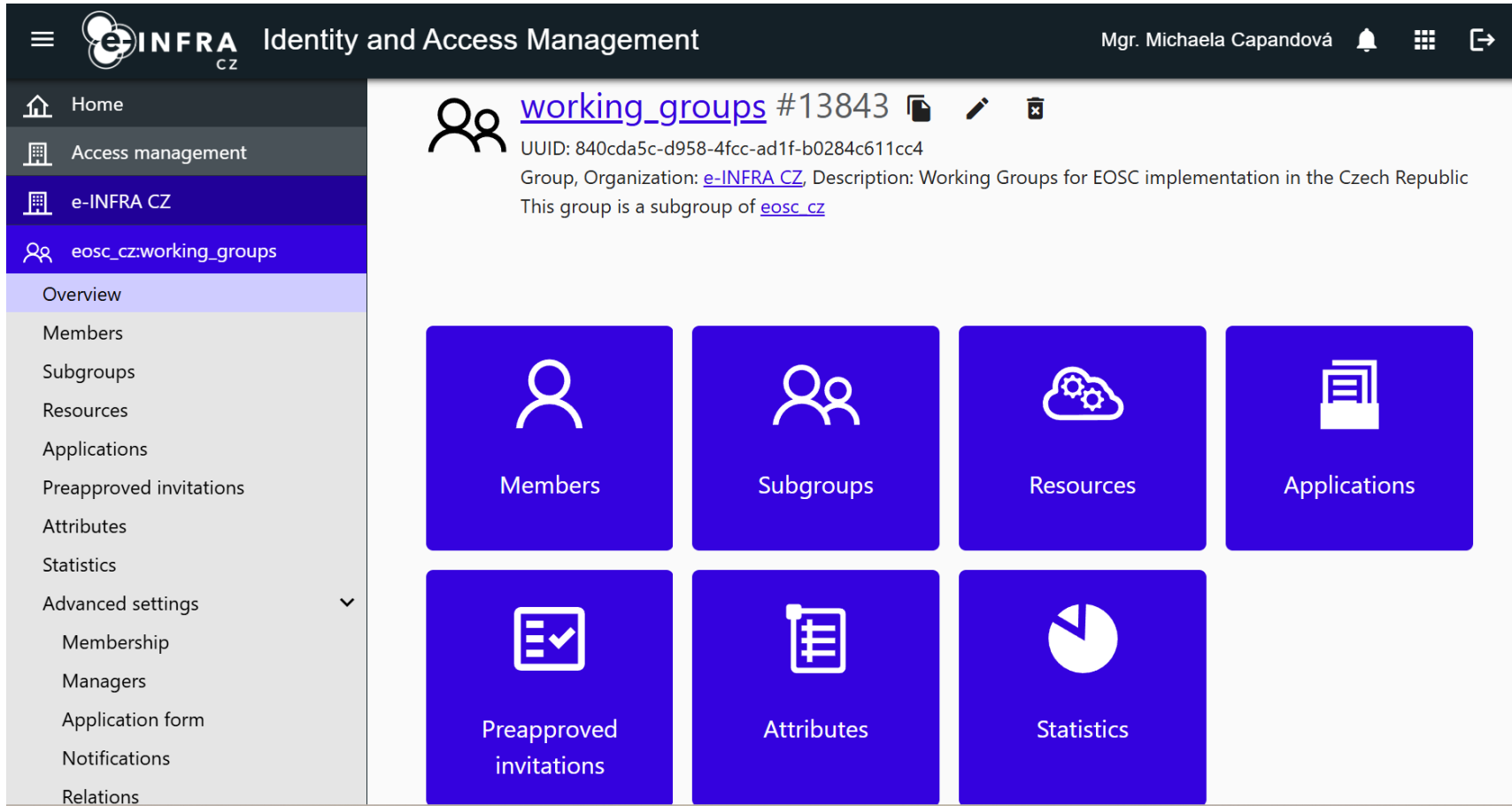
Horizon Europe DMP

☐ a. Only all project members have read/write access to the data

☐ b. Anyone in the institute has read access to the data; all project members have read/write access.



Autentizační a autorizační infrastruktura



The screenshot displays the 'e-INFRA Identity and Access Management' interface. The left sidebar contains a navigation menu with the following items: Home, Access management, e-INFRA CZ, eossc_cz:working_groups (selected), Overview, Members, Subgroups, Resources, Applications, Preapproved invitations, Attributes, Statistics, Advanced settings (with a dropdown arrow), Membership, Managers, Application form, Notifications, and Relations. The main content area shows details for the 'working_groups' group (#13843). It includes a UUID (840cda5c-d958-4fcc-ad1f-b0284c611cc4), the group name 'working_groups', and a description: 'Group, Organization: e-INFRA CZ, Description: Working Groups for EOSC implementation in the Czech Republic'. It also states 'This group is a subgroup of eossc_cz'. Below this information is a grid of seven blue tiles with white icons and labels: Members, Subgroups, Resources, Applications, Preapproved invitations, Attributes, and Statistics.

Umožňuje uživatelům z různých institucí jednoduchý přístup k datům a službám.

- Řízení přístupu a identit
- Řízení skupin a jejich rolí
- Přidělování práv

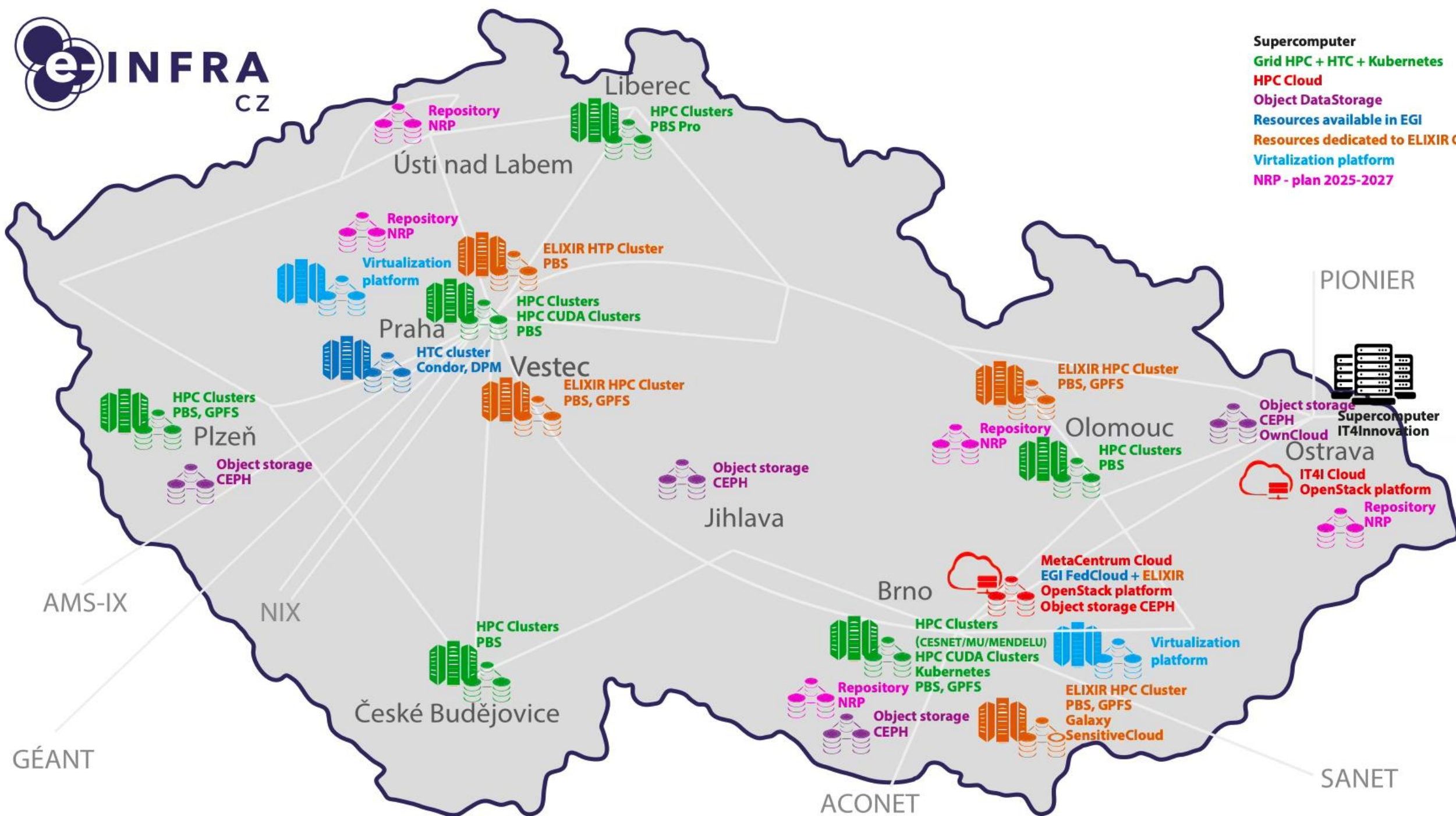


Výpočetní možnosti e-INFRA CZ

Gridová infrastruktura MetaCentrum (CESNET)

a

Superpočítačové centrum IT4Innovations





INTERACTIVE APPLICATIONS

INTERACTIVE APPLICATIONS

- Matlab
- Gromacs
- AlphaFold
- CLCbio
- RStudio
- Ansys ...

Konkrétní aplikace ovládaná uživatelem interaktivně a v grafickém (web-based) prostředí.



PLATFORMS

PLATFORMS

- Jupyter Notebooks (OpenPBS, K8s)
- Galaxy (OpenPBS)
- OnDemand (OpenPBS)

Okamžitě dostupná web-based služba/platforma umožňující spustit požadované aplikace.



INFRASTRUCTURE

INFRASTRUCTURE

- OpenPBS for HPC
- Kubernetes (K8s) / Rancher
- OpenStack Cloud / Sensitive Cloud

Výpočetní infrastruktura řídící přidělování HW a SW jednotlivým uživatelským úlohám.



HARDWARE

HARDWARE

- MetaCentrum: 47 tis. CPUs, 462 GPUs
- IT4I SuperComputers: 30 tis. CPUs, 16,7 PFlops

Fyzický HW (GPU, CPU, RAM, disková pole) rozdělovaný mezi jednotlivé uživatelské úlohy.

Podpora práce s daty

Data stewardi všude

Agenda datastewarda:

- Tvorba **data management plánu** a jeho pravidelná aktualizace.
- Správa výzkumných dat v souladu s **FAIR** principy.
- Výběr vhodných datových a metadatových schémat.
- Zajištění **kvality dat** a metadat.
- Dodržování politik institucí a oborových standardů (**Data Compliance**).
- Zaručení ochrany dat (**Data Security**).
- Výběr vhodného datového repozitáře či úložiště.
- Podpora při otevřeném sdílení dat "**as open as possible, as closed as necessary**".



Komunita data stewardů

- Podpora sdílení znalostí v oblasti správy výzkumných dat.
- Členové (už přes 300 lidí) si vyměňují **tipy a konzultují** se zkušenými kolegy.
- Komunita na **Discordu**.
- **Metodické materiály a workshopy** jako užitečné nástroje pro účastníky.
- **Testování a hodnocení** Národní repozitářové platformy (NRP), které pomáhá jejímu zlepšení – již brzy.
- **Spolupráce** zajišťuje přístup k aktuálním znalostem a osvědčeným postupům.
- **Otevřeno** všem zájemcům o správu výzkumných dat.



Data stewardi a EOSC CZ

Manuál začínajícího data stewarda

Autoři

Dagmar Hanzlíková (Univerzita Karlova)

<https://orcid.org/0000-0001-9287-399X>

Matyáš Hiřman (Univerzita Karlova)

<https://orcid.org/0000-0003-0153-9414>

**FAIR příručka pro data
steward komunitu v ČR**

Pracovní nabídky

Manažer/ka otevřené vědy odborný pracovník

Zaměstnání na plný úvazek, nástup od 1. 4.
2025 nebo po dohodě

[VÍCE O POZICI](#)

Specialist(k)a na Open Science (Data Steward)

IT4Innovations hledá nového/novou
kolegu/kolegyni, který/á bude pracovat na
pozici Specialist(k)a na Open Science (Data
Steward)

[VÍCE O POZICI](#)

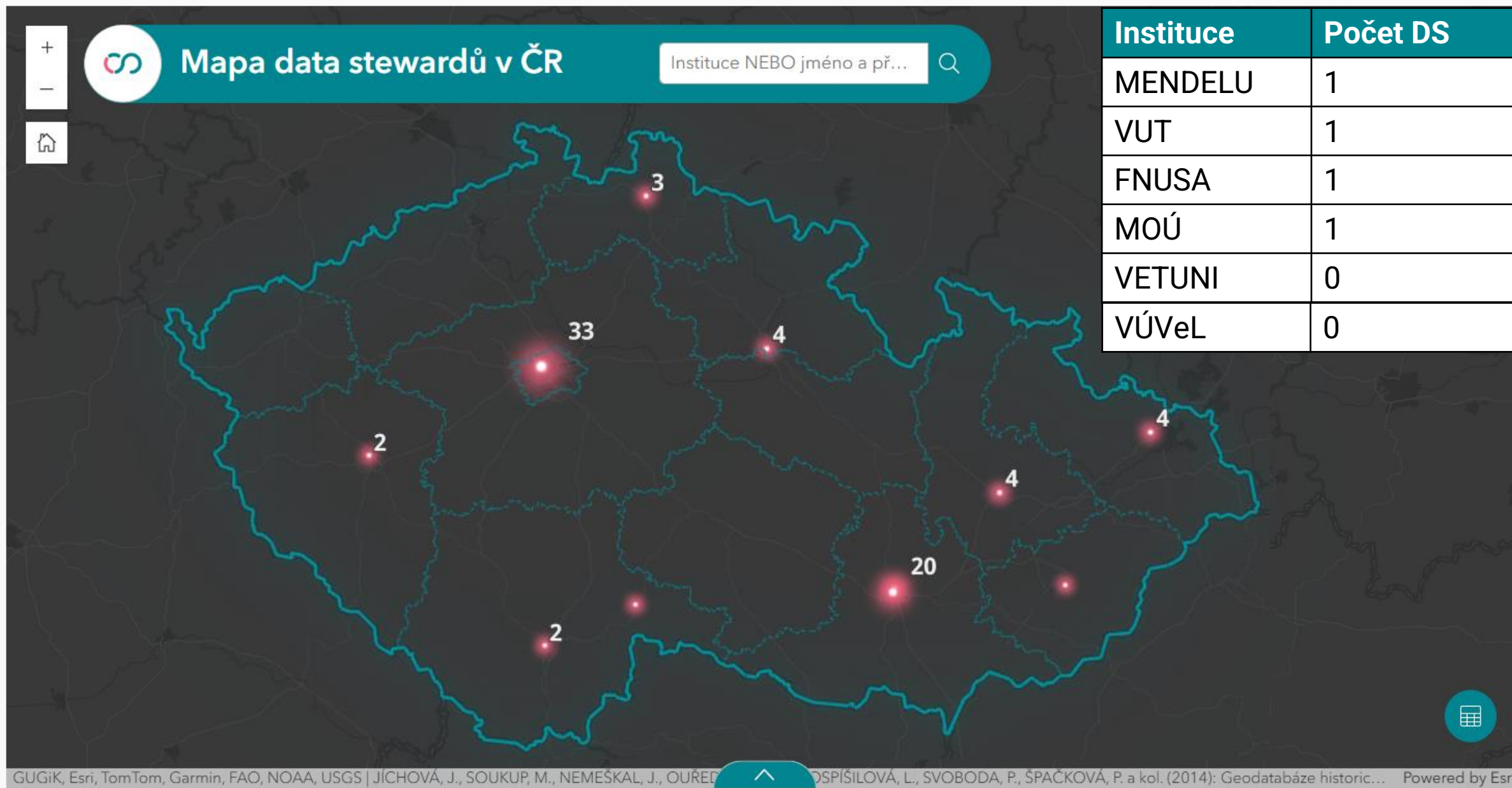
Data steward do oddělení Open Science (Knihovna AV ČR)

Nástup: možný od 1. března 2025 nebo po
dohodě

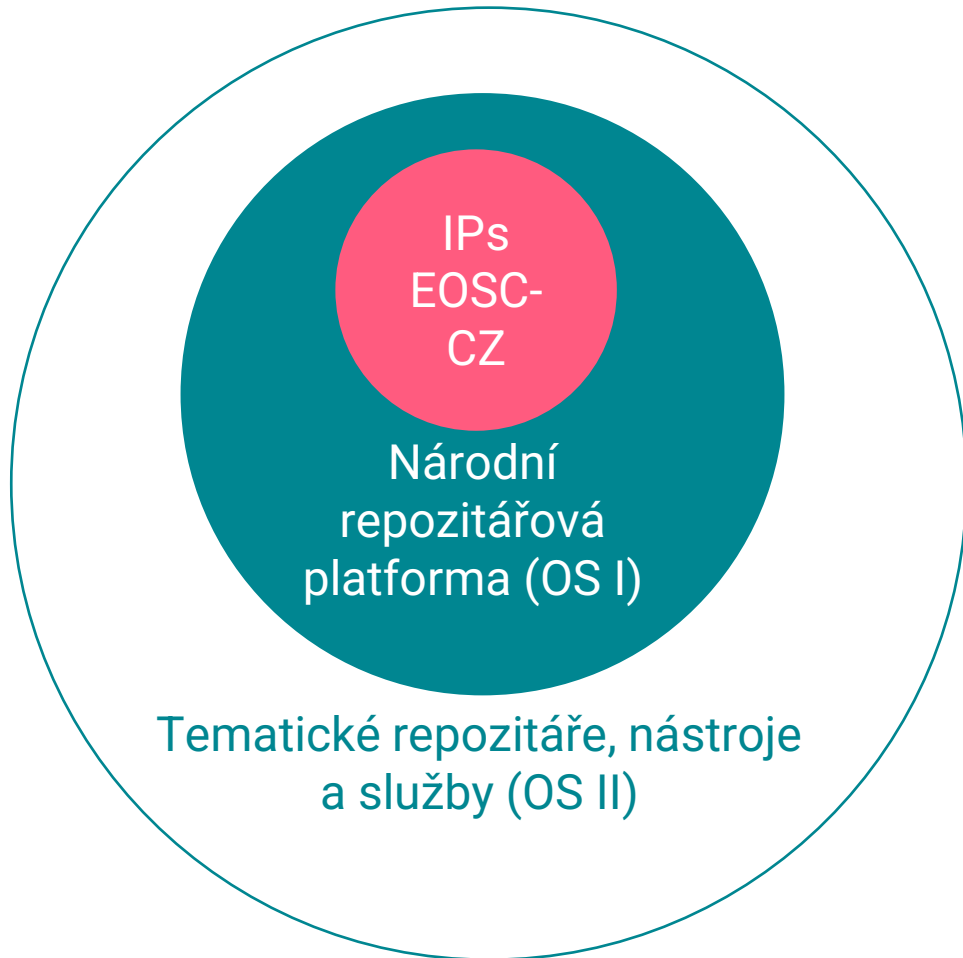
[VÍCE O POZICI](#)

**Objevte svět data stewardů:
Seriál o klíčové roli v moderním
výzkumu**





Projekty iniciativy EOSC CZ



IPs EOSC-CZ (od r. 2023) – Složky EOSC implementace v CZ

- Organizační (**Sekretariát**) – <https://www.eosc.cz/en/secretariat>
- Technická (**Národní metadatový adresář**) – <https://nma.eosc.cz/>
- Vzdělávací (**Školicí centrum**) – <https://www.eosc.cz/en/training-centre>

Národní repozitářová platforma (OS I, od r. 2024) – “železo”

- **Repozitářové systémy** (Dspace, Cesnet Invenio, ASEP-ARL) (50+ PB kapacita)
- **Pilotní repozitáře**
- **Základní služby** (PIDs, DSW, licence, ...)
- **Compliance** a UX (kyberbezpečnost, ServiceDesk, ...)
- **Školení** – technické aspekty

OS II (od r. 2025) – “oborově specifické”

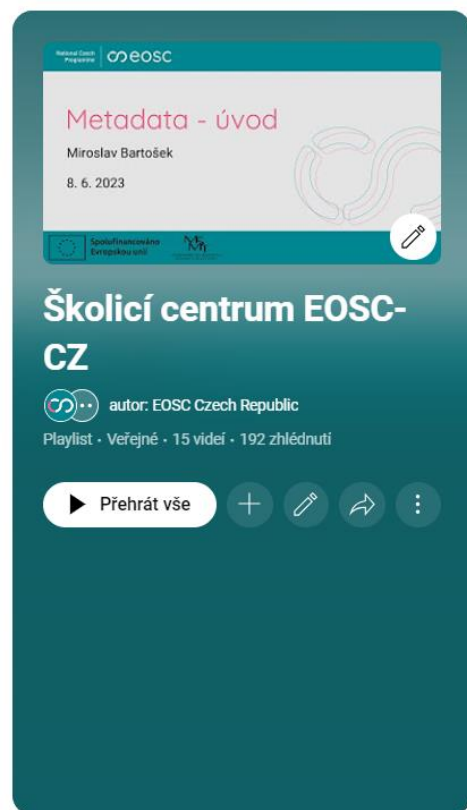
- **Na základě vstupů od 8 tematických pracovních skupin**
 - Bio/Zdraví/Potraviny, Materiálové vědy, AI & ML, Sociální vědy, Fyzika, Humanitní vědy a umění, Enviro, Citlivá data
- www.eosc.cz/pracovni-skupiny

OS III (od r. 2026)



Vzdělávání

Databáze školení na YouTube a webu



Metadata - úvod

EOSC Czech Republic • 450 zhlédnutí • před 1 rokem



Úvod do metadat II - metadeta v kontextu NRP

EOSC Czech Republic • 221 zhlédnutí • před 1 rokem



Úvod do organizace znalostí: řízené slovníky, ontologie

EOSC Czech Republic • 166 zhlédnutí • před 1 rokem



Systémy pro Národní repozitářovou platformu (NRP)

EOSC Czech Republic • 254 zhlédnutí • před 10 měsíci



Pilotování Data Stewardsh

EOSC Czech Republic • 119 zhlédnutí



Školení podle životního cyklu dat



Přehledové a úvodní kurzy



Plánování a příprava dat



Sběr dat



Zpracování dat



Analýza dat



Ukládání dat



Sdílení dat



Znovuvyužití dat

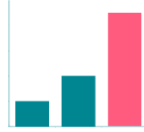


Ostatní



Data veřejného sektoru vs. Výzkumná data

Data veřejného sektoru



Jasná hierarchie metadat



Řízený přístup k datům částečně digitalizován



Primárně Open Data



DCAT-AP CZ

Výzkumná data



Oborové (komunitní) standardy a více federováno



Řízený přístup k datům digitalizován

FAIR

Primárně FAIR Data



DCAT-AP + DataCite



Využívají Open Source repozitáře



Dělá/dělal na tom „Klímek“, „Dvořák“ nebo „Med“... ☺

Co bude následovat

V tomto roce...



Infrastruktura

- Instalace **prvního hardwaru pro repozitáře**
- Vznik **prvních a pilotních repozitářů**
- Národní katalog repozitářů



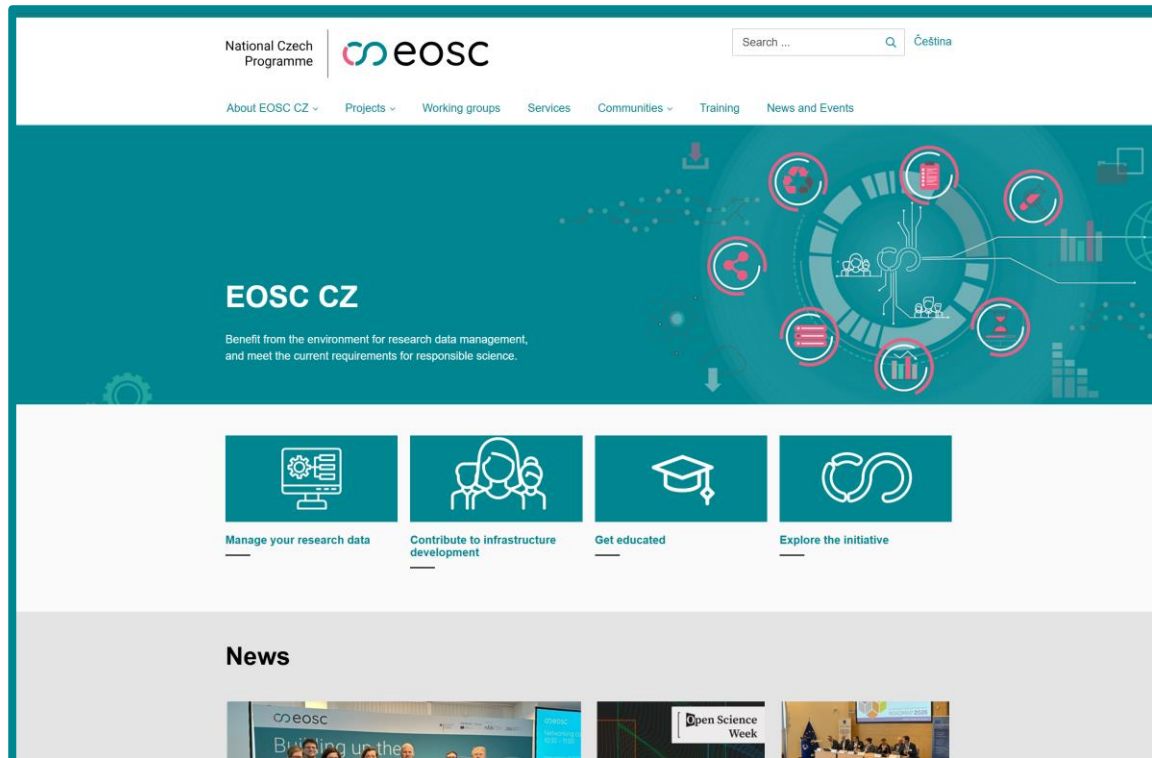
Služby

- První verze všech hlavních základních služeb:
 - NRP DSW (konec března 2025)
 - Izolované instance DWS (konec srpna 2025)
 - Catch-all repozitář v druhé verzi (konec roku 2025)
 - Nástroje pro stažení dat z NRP (konec roku 2025)
 - Propoj na výpočet v e-INFRA (konec roku 2025)



Události

- National Tripartite Event
- Letní škola pro data stewardy
- ... školení, workshopy, semináře, konference, ...
- Open Science Workshop 2025



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

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

Matej Antol, Jiří Marek, Michaela Capandová, Jaroslav Juráček, 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—EOOSC, EOOSC 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 (EOOSC) [3], [4] is an international initiative that builds on the Open Science principles. EOOSC 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. EOOSC CZ – INFRASTRUCTURE AND SERVICES FOR FAIR RESEARCH DATA

The establishment of fundamental principles for the Czech national EOOSC implementation took place in 2021, resulting in the document called *Architecture of EOOSC implementation in the Czech Republic* [6]. The document represents the official start of the EOOSC 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

of Sciences and Research Infrastructures. Examples are environments such as LINDAT/CLARIAH-CZ [13] for natural language processing, Czech-BioImaging [14] for biological and medical imaging or EIRENE RI [15] for human exposome. 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 EOOSC CZ's success. Since 2021, as a reaction to the EOOSC CZ Architecture document, 12 EOOSC CZ working groups [22] have been established through a self-organizing community effort. These groups will be operational during the entire EOOSC 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

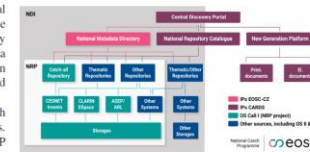


Fig. 1. NDI and NRP blueprint with five abstraction layers. Bottom-up: hardware infrastructure dedicated across Czechia; three initial repository systems – CESNET Invenio [8], CLARIN-DSpace [9] and ASEP/ARL [10]; specific domain and other repositories, metadata directory and on top, Central Discovery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

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

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 EOOSC CZ OUTCOMES

The NDI's ecosystem of services will be offered to the whole research community regardless of their active participation in the EOOSC CZ initiative. The EOOSC 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 EOOSC CZ initiative.

ACKNOWLEDGMENTS

The EOOSC 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 Slanáková, Petra Čermohávková, Martin Svoboda, Miroslav Bartošek, David Antoš and Michal Ruzicka.

APPENDIX: FINANCIAL SUPPORT FOR EOOSC IN CZECHIA
Czech Ministry of Education, Youth and Sports (MEYS) supports the EOOSC CZ initiative [24] via two systemic projects and three open science calls:

- Individual Systemic Project (IPs) EOOSC-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 mil. 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] Munafò, M., Nosek, B., 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/stia-mat/>
- [6] https://www.mstt.cz/uploads/311/Architektura_implementation_EOOSC_v_CK.pdf
- [7] <https://www.eosc.cz/en/>
- [8] <https://purlhub.com/CESNET/>
- [9] <https://purlhub.com/dal/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.czech-bioimaging.cz/>
- [15] <https://www.eirene-ri.eu/>
- [16] <https://perun-sas.org/>
- [17] <https://dsi-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.data-cz.eu/en/statisticky-a-analyzy/seznam-operaci-tjrepceni>



Matej Antol is the principal project manager of the IPs EOOSC-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 EOOSC CZ initiative and head of the EOOSC 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 EOOSC 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 Juráček is the secretary to the EOOSC CZ Working Groups 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 EOOSC-CZ steering committee.

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

Důležité odkazy a kontakty

Veškeré informace

eosc.cz/cz



Newsletter iniciativy EOSC CZ



- Máte nějaké otázky?
 - info@eosc.cz
- Máte nápad na přednášku nebo školení?
 - events@eosc.cz
- Chcete kontaktovat naše komunikátory?
 - pr@eosc.cz



@EOSC Czech Republic



@eosccz.bsky.social

Děkuji Vám za pozornost.



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

MUNI
ICS

cesnet
.....

VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA

IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM