

Czech EOSC, during Czech Presidency!





MUNI CERIT-SC ŠB TECHNICKÁ | IT4INNOVATIONS || UNIVERZITA | NÁRODNÍ SUPERPOČÍTAČOVÉ OSTRAVA | CENTRUM

2 EOSC Symposium (EOSC CZ)

13:50-14:30: Panel Discussion and Q&A

J. Klánová, J. Vondrášek (Environmental and Life sciences), M. Cebecauer (Material Sciences and Technology),

Moderator: Jiří Marek, Open Science Manager, MUNI

P. Straňák, J. Krejčí (Humanities and Social sciences)

D. Hanzlíková (Education and Human Resources),

13:00-13:15: Introduction to the Czech Implementation of EOSC (L. Matyska)

M. Svoboda (Metadata),

13:15-13:50: Czech EOSC Lightening Talks:

Session plan





Steps towards the Czech Implementation of EOSC







....

- Broader context: Open Science principles
 - FAIR (Findable, Accessible, Interoperable, Reusable) data
 - Funding agencies pushing towards openness
 - Open Access
 - Data Management Plans
- Focus on FAIR research data
- EOSC CZ seen as a federation of FAIR data and related services
 - Federation of existing and future data resources (FAIR data repositories)
 - Open ecosystem of services

EOSC CZ Principles



....

- EOSC CZ architecture discussion in 2021
 - Under the auspices of the MEYS
- Result: The Architecture of EOSC Implementation in the Czech Republic
 - National Data Infrastructure (NDI)
- Data-centric infrastructure with 4 key components (pillars)
 - National Metadata Directory (NMA)
 - National Repository Platform (NRP)
 - Thematic repositories
 - Education and training (human resources)

Funding and calls



- EOSC implementation to be supported through structural funds
- Preliminary allocation of 120 M€ for the 2023-2028 period
- Several calls under discussion
- First call already out
 - Projects EOSC CZ (EOSC Secretariat, NMA, training coordination; e-INFRA CZ) and CARDS (Metadata, PIDs; NTK)
- Second call under preparation
 - NRP implementation, call early 2023, allocation around 50 M€
- Third call for the **thematic clusters**, call early 2024, allocation around 40 M€
 - Discussed/prepared within thematic working groups
- Fourth call **Uptake/Upskilling**, call early 2025
 - Initial discussion only

6 EOSC Symposium (EOSC CZ)⊲⊲-

Governance and Preparatory Steps



- Coordination Board at MEYS
 - Chaired by two vice-ministers
- Working groups established since Autumn 2021
 - Open platform, with fluid membership
 - Discussion of basic principles of all aspects of the EOSC CZ implementation
 - What is needed by researchers and their communities
 - And how to best fulfill these expectations
- 4 foundation working groups
 - Metadata, Architecture, Core Services, Training and Education
- 7+1 thematic working groups
 - Defining gradually the thematic clusters

7 EOSC Symposium (EOSC CZ)⊲⊲–

Thematic Working Groups



....

3000

- Bio/Health/Food Jiří Vondrášek (ELIXIR CZ)
- ENVRI Jana Klánová (RECETOX a EIRENE)
- **Physics** Jiří Chudoba (CERN)
- Material Sciences and Technology Marek Cebecauer
- Al and Digital Science Jan Šivic
- Humanities and Arts Jan Hajič (LINDAT/CLARIAH-CZ)
- **Social Sciences** Jindřich Krejčí (ČSDA)

and

• Sensitive Data – Zdenka Dudová (BBMRI CZ)

National Data Infrastructure

3000

Key components

National Metadata Directory (NMA)

- Key component, guaranteeing findability and interoperability
- Interacting directly with all other NDI components
- Centrally managed NDI component
 - Still distributed; e-INFRA CZ responsibility
- National Repository Platform (NRP)
 - NDI backbone, the capacity and basic/common services (AAI, PIDs, DMPs, ...)
 - Defines standards and interfaces for the technical interoperability
 - Will be implemented through a large national consortium (to be coordinated by EOSC Association mandated organization)

National Data Infrastructure

3000

Other parts

Human Resources

- Horizontal Activity
- To be supported through the whole implementation of NDI

• Existing thematic repositories

- Support of on-going activities and repositories
 - Continuity, long term sustainability, ...
- Support of new thematic needs
 - Not new repositories, but with the support of the NRP, its capacity and services
- Interconnection with NMA and NRP
 - Interoperability at the technical and logical (metadata) levels
- Specific care of **sensitive data**

National Data Infrastructure



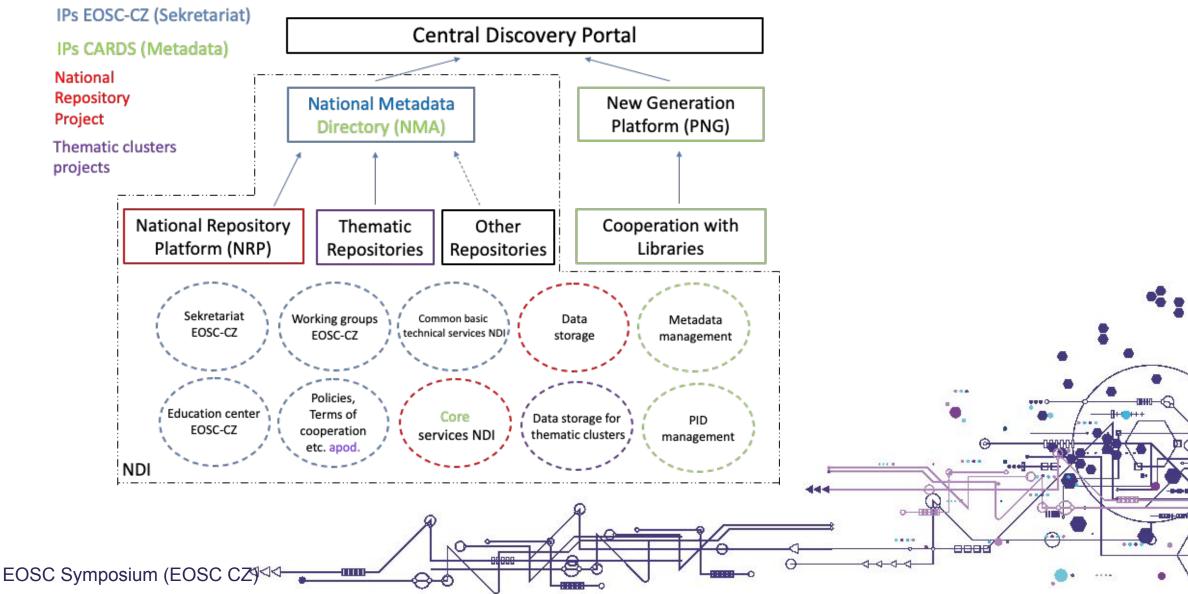
.....

- Complex system with the following aims:
 - Basic infrastructure with sufficient capacity to store research data in Czechia
 - Open set of key and other services for data manipulation
 - National Metadata Directory and National Repository Platform
- With parallel support of human resources
 - Education of (future) scientists, new competencies (curation), data stewards, data scientists
- Not everything built from scratch, the existence of **thematic repositories** accounted for
 - Usually, part of large research infrastructures
 - Universities and other research organizations mostly in planning phase yet
- Data-centric means primary and secondary data, not just data as appendices to publications

11 EOSC Symposium (EOSC CZ)⊲⊲–

Czech EOSC implementation scheme





Summary



- EOSC: move from compute-centric to the data-centric view
 - Strong support from the ministry
 - Opportunity for the e-INFRA CZ reshaping its core business
- EOSC implementation as an opportunity to evolve the national e-infrastructure
 - Collaboration and partnership a key
- Human resources through training and education
 - Universities' involvement a key
- NDI as an open complex environment
 - Capacity, metadata and ontologies, services
 - Full compliance with European level (MVE just a start)
 - Thematic repositories fully integrated into the whole ecosystem
- Also support for sharing and actively working with information about EOSC
 - Secretariat for EOSC CZ, Information and training Portal

13 EOSC Symposium (EOSC CZ)



Thank you for your attention <u>eosc-info@e-infra.cz</u>

More information at <u>e-infra.cz/EOSC</u>

....





Metadata Interoperability in the Czech EOSC

M. Svoboda, P. Černohlávková (National Library of Technology)



B TECHNICKÁ | IT4INNOVATIONS | UNIVERZITA | NÁRODNÍ SUPERPOČÍTAČOVÉ OSTRAVA | CENTRUM

CARDS project (2023 – 2028)

- Czech Academic and Research Discovery Services
 PA 2 Central New Generation Platform (PNG)
 PA 3 Research Data Metadata and Persistent IDentifiers
 Support Metadata WG EOSC-CZ
- People Martin Svoboda, Petra Černohlávková
- Goals, Activities, Outputs
 - management of all types of information resources
 - services based on a central shared catalogue
 - metadata schema services, PIDs management
 - ensuring interoperability between data repositories



50°6'14.083"N, 14°23'26.365"E Národní technická knihovna National Library of Technology



Research Data metadata and PID support (1)



anae

Metadata interoperability support

Dedicated department within NTK

- Support and consultations for research data metadata management
 - universal metadata model for reserach data description (definiton, modification, implementation in National Metadata Directory)
 - metadata workflow settings
 - metadata models for subject clusters
- National Repository Catalogue

EOSC Symposium (EOSC CZ)

- Operational information on repositories, their parameters and tools
- Leadership and coordination of the activities of the EOSC-CZ Metadata Working Group

Research Data metadata and PID support (2)



anae

National Persistent Identifiers Centre

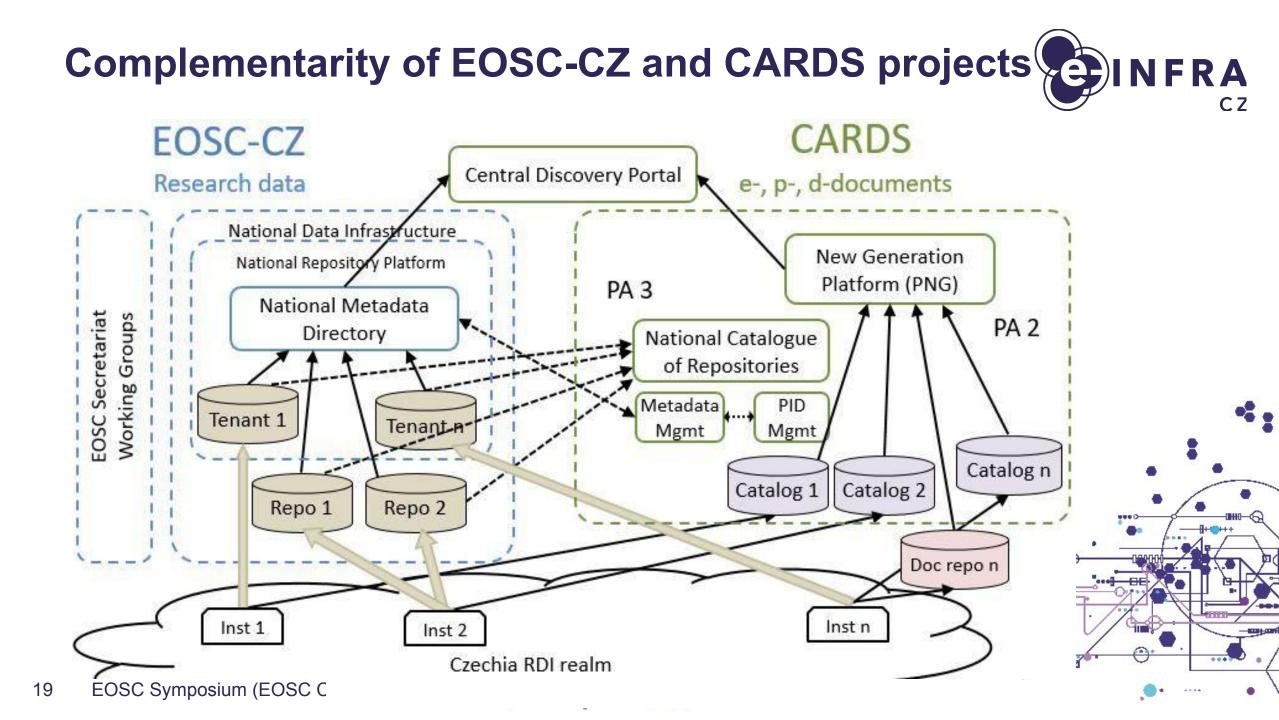
Dedicated department within NTK

- National expert centre for the PIDs
 - administrative, coordination, methodological and financial support
 - DOI consortium
 - support for ORCID consortium
 - upgraded Czech ISSN centre services

• Goals

- facilitate introduction of PIDs, support research institutions consortia
- national access point for PIDs support
- information specialists training

18 EOSC Symposium (EOSC CZ)</



Integration in the Czech EOSC area



TOOP

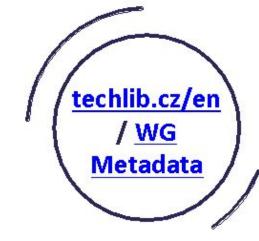
Complementarity of EOSC-CZ and CARDS projects

- National Repository Catalogue
 - directory of relevant resources to link to the NMD
- Support for the National Metadata Directory
 - metadata harvesting into the NMD support
 - implementation of the Universal Metadata Model
 - consultation on metadata issues and their impact on NMD functionality
- Consultations (other metadata and PIDs support)
- Central shared catalogue PNG for digital research output, electronic, digitized and printed resources



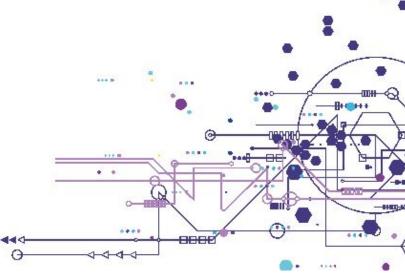
Thank you for your attention <u>ndi-wg-metadata@e-infra.cz</u>





.....

- 60





Data steward? What's a data steward?

Dagmar Hanzlíková



Implementing a new infrastructure



....

30**0**0

What do you need?

- Resources
- People who build the infrastructure
- Clear policies and guidelines
- Trained researchers
 - Researchers need to organise their work differently (FAIR)
- Trained support staff
 - They need support to do that

WG Education & Human Resources



- WG lead: Radka Římanová
- February 2022

EOSC Symposium (EOSC CZ)

- 35 members, 20 institutions
- Aims

24

- Training for researchers
- Training research support staff



CHARLES UNIVERSITY Central Library

TOOP

Trained researchers



-

- What **skills** do researchers need to have?
 - Basic knowledge x subject specific
- How do they **acquire** these skills?



Research Data Management

Introduction

Welcome to the Research Data Management course!

Data-based research is becoming more and more common across a wide range of scientific disciplines. With the growing importance of research data, there is also a growing need to manage data properly throughout the research process. In this course, we will introduce the **basic concepts of research data management** as well as some specific **tools** that can help you with data management.

1111

During this course you will learn

- What are research data
- What is research data management
 What are the benefits of research data management
- What is a data management plan and how to create one
- Useful tips for collecting and processing research

data • What are open research data and how to share your

data

....and a lot more!

How does it work?

Each chapter introduces one topic of research data management and offers basic **definitions** and **description**, supplemented by multimedia, **useful links** and an overview of **recommended literature**. At the end of each chapter, there is a short **quiz** to test your newly acquired knowledge. In order to be able to complete the quizzes, you need to enrol in the course (cogwheel in the top right comer > find me in this course).



Open Course Data Stewardship

The online course in Data Stewardship provides an introduction to research data management. Managing research data in accordance with good practices is an easily transferable and sought-after skill both within and outside of academia. In many areas of the non-academic sphere data management is vital, and competencies and skills within this topic are therefore in high demand.

0-

....



Trained support staff

....

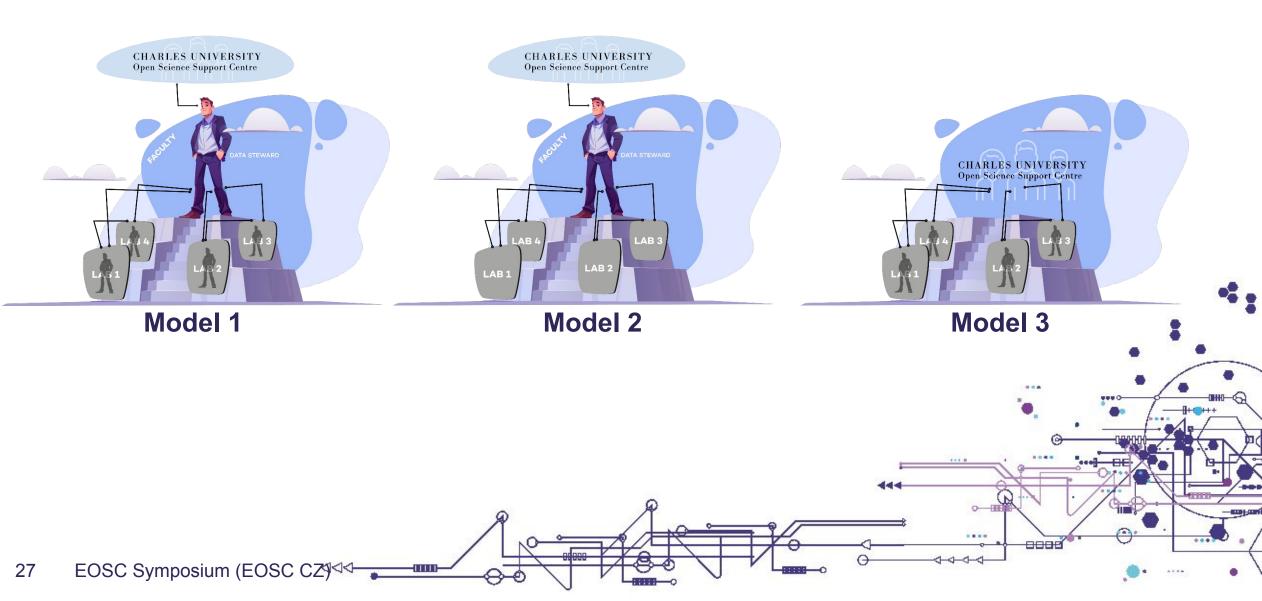
0000

Data Steward

- New type of job position in the Czech Republic
 - Job profile
- How to implement the position at an institution (money matters)
 - Administrative or research position?
 - Project-based or institutional?
- 3 models of implementation

Models for implementing data stewards





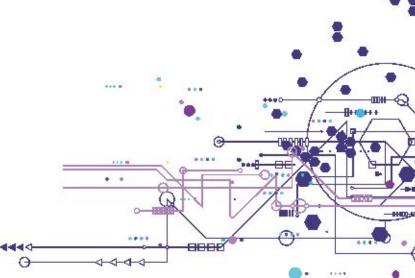


CHARLES UNIVERSITY Central Library

Thank you for your attention ndi-wg-education@e-infra.cz









Czech EOSC, during Czech Presidency!

Environmental Health Sciences

Jana Klánová

EOSC Symposium 2022, Prague, 14.11.2022

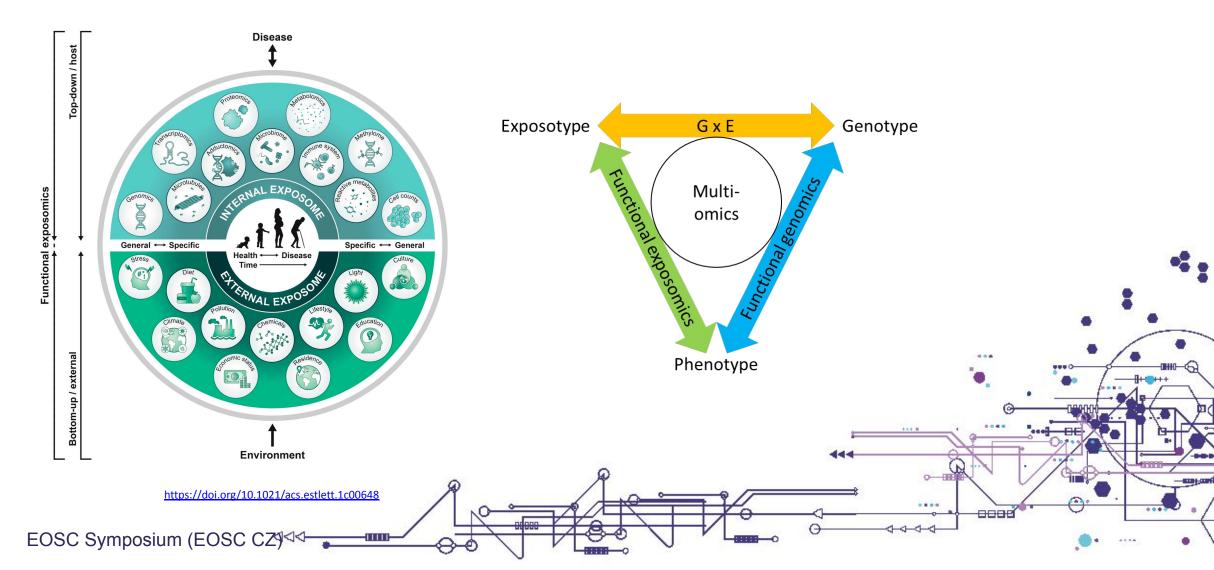
B TECHNICKÁ | IT4INNOVATIONS | UNIVERZITA | NÁRODNÍ SUPERPOČÍTAČOVÉ OSTRAVA | CENTRUM

Introduction

30



WG ENV EOSC-CZ: Linking environment and health



Objectives



TOOP

- Instrumental for implementation of the Green Deal (Zero pollution ambition, Climate) and European Chemical Strategy
- Meta data harmonisation accross the research domains to enable federated analyses (to implement human exposome concept)
- Alignment / synergies with on-going European initiatives (Partnerships (PARC, Agriculture of data), Ris (EIRENE), EHEN, EURION)
- Implementation of FAIR standards

EOSC Symposium (EOSC CZ)

• Education and training

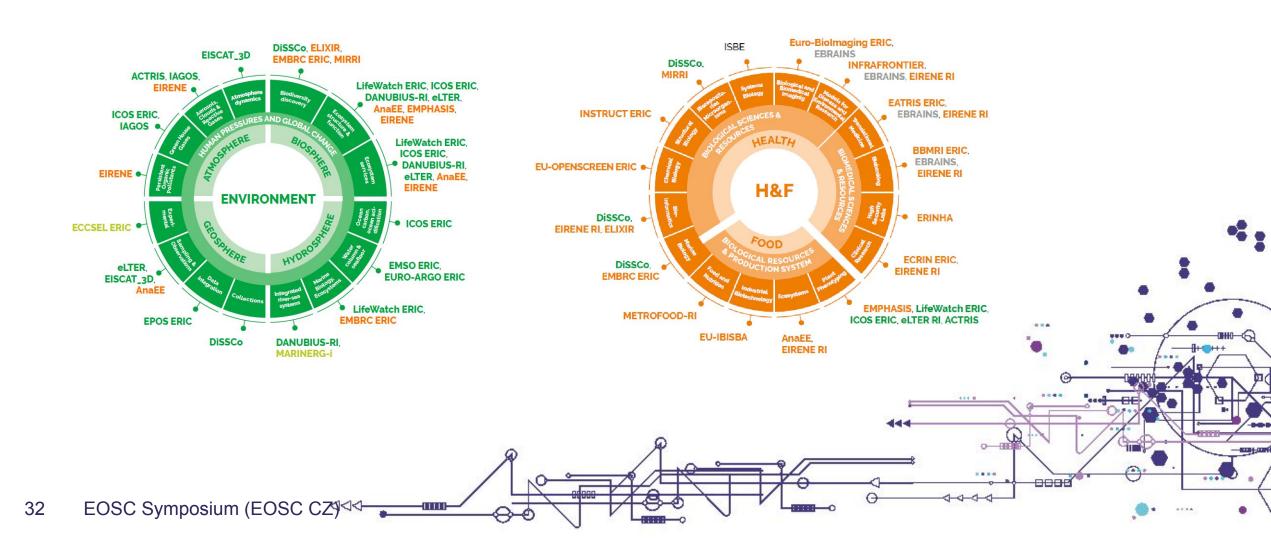
31

 Collaboration accros the domains (EIRENE, ICOS, ACTRIS, DANUBIUS, BBMRI, EATRIS, ECTRIN, ELIXIR).

Challenges and synergies



Environment, Life Sciences, Health and Food, Social, and e-RIs



Interdisciplinary domain of Environmental Health Sciences works with data from

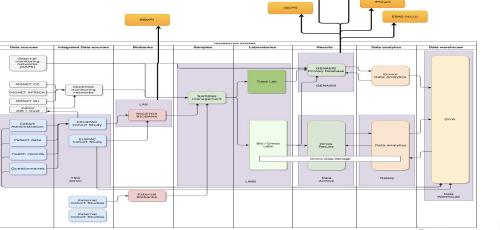
- environmental monitoring networks and field studies (air, water, soil, food, indoor environment),
- longitudinal population studies linking environmental data to socio-economic, lifestyle, health, etc.,
- health registers,
- sample collections and biobanks,

Data repositories include, for instance,

- geo-coded data from on-line measurements,
- results from laboratory analyses of biotic and abiotic samples or toxicological testing.
- data from questionnaires and health examinstions,
- sensitive personal data.

33 EOSC Symposium (EOSC CZ)</

Data types





Participating Institutions and Infrastructures



anae

MUNI | RECETOX

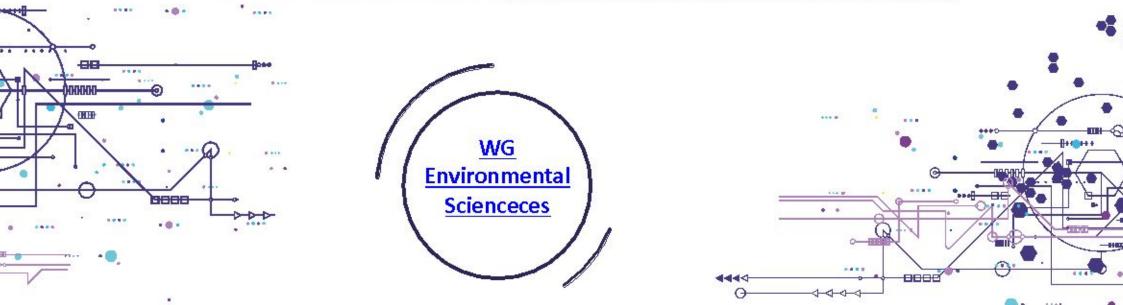
- Institutions:
 - Masaryk University, Charles University, South-Bohemian University, Global Change Reseach Institute, J. Heyrovský Institute of Physical Chemistry, Institute of Chemical Processes, Czech Hydrometeorological Institute
- Research Infrastructures:
 - RECETOX RI (EIRENE, ACTRIS, BBMRI)
 - ACTRIS CZ (ACTRIS RI)
 - CzechCOS (ICOS ERIC, ANEEE, EUFAR AISBL, ACTRIS, DANUBIUS, ELTER)
 - CENAKVA (DANUBIUS-RI)
 - DiSSCo
 - NanoEnviCz

34 EOSC Symposium (EOSC CZ∮</

MUNI RECETOX

Thank you for your attention ndi-wg-enviro@e-infra.cz







Czech EOSC, during Czech Presidency!

A.Strachotová, (J. Vondrášek)



Introduction



3000

• Work Group **Bio-Health-Food**



- Towards Open Science in Bio-Health-Food Data space
- Towards FAIR User Friendly Environment
- Health-care begins with Food and Bio Informatics

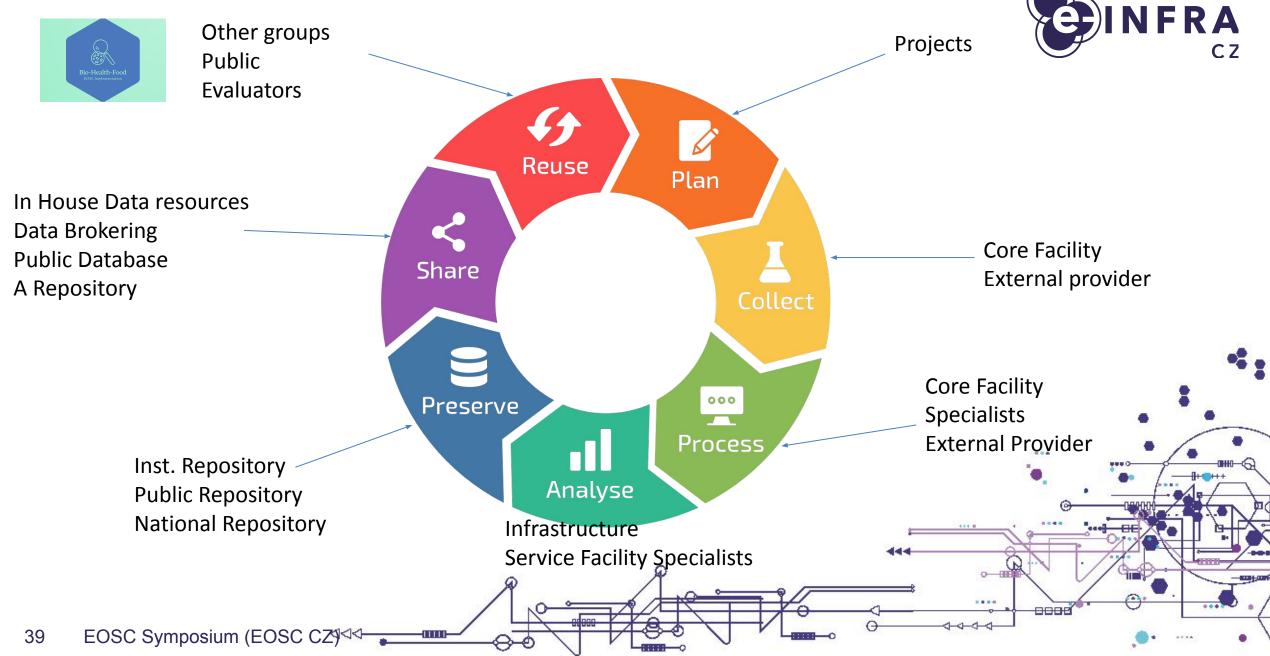
37 EOSC Symposium (EOSC CZ∮⊲⊲ _

Participating Institutions and Infrastructures



- ÚOCHB (IOCB) Prague
- MUNI CEITEC, Faculty of Informatics Brno
- FNUSA Brno
- METROFOOD-CZ (ČZU) Prague
- ÚMG (IMG) AS CR Prague
- UPOL Olomouc
- UK LFHK, PřF, 1.LF (Ústav lékařské biochemie a laboratorní diagnostiky), IMCF -Prague
- CZ-OPENSCREEN (UMG), Prague
- BIOCEV Prague
- VFN Hradec Kralove
- CCP (UMG), Prague

Data Life Cycle and Open Acess (Institutional Point of View) in the WG



Disciplinary heterogeneity and intersection with other working groups - Sensitive data



TOOP

- Supporting links to the decisions of the "Open Access Data Commission "
- Support for DMP creation.
- Support for DMP specific templates/checklists for sensitive data.
- Support for DMPing, i.e. data management planning, i.e. the process as such, not just the creation of DMP as a document.
- Methodological support for managing sensitive data.
- Suitable technical means for data storage and processing
- Secure storage appropriate certification, security policy, etc., which may be referenced from the DMP to be used.

ELIXIR contribution to the EOSC

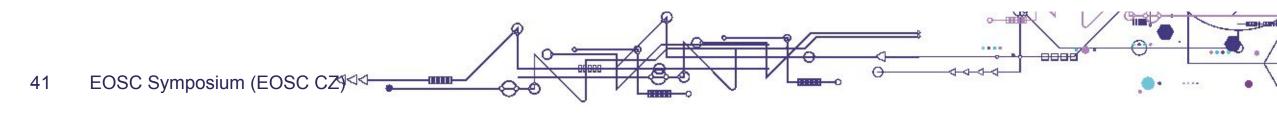


GUIDING PRINCIPLES ON FAIR DATA MANAGEMENT IN LIFE SCIENCES

SEPTEMBER 11, 2017

The ELIXIR hub has published a set of guiding principles on FAIR (Findable, Accessible, Interoperable and Reusable) Data Management.

- Open sharing of research data is a core principle for publicly-funded research and ELIXIR encourages all funders to adopt Open Data mandates.
- Data Management is crucial part of good scientific practice and research excellence.
- Whenever possible, biological research data should be submitted to the recommended community deposition databases.
- All data submitted to Open Data archives must be annotated in accordance with community-defined standards.
- ELIXIR Nodes are the national implementation of a harmonised FAIR Data Management programme for the life sciences.
- FAIR data management requires professional skills and adequate resources.
- Good research data management requires appropriate funding for data infrastructures.

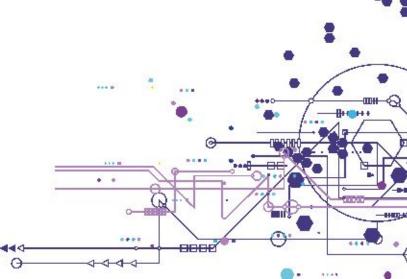




Thank you for your attention ndi-wg-bio-health-food@e-infra.cz









NAMIDAS - decentralized data repositories for active communities

Marek Cebecauer

J. Heyrovsky Institute of Physical Chemistry of the Czech Academy of Sciences

EOSC Symposium 2022, Prague, 14.11.2022

VŠB TECHNICKÁ | IT4INNOVATIONS ||| UNIVERZITA | NÁRODNÍ SUPERPOČÍTAČOVÉ OSTRAVA | CENTRUM

Introduction



WG Material Sciences and Engineering/Technologies EOSC-CZ

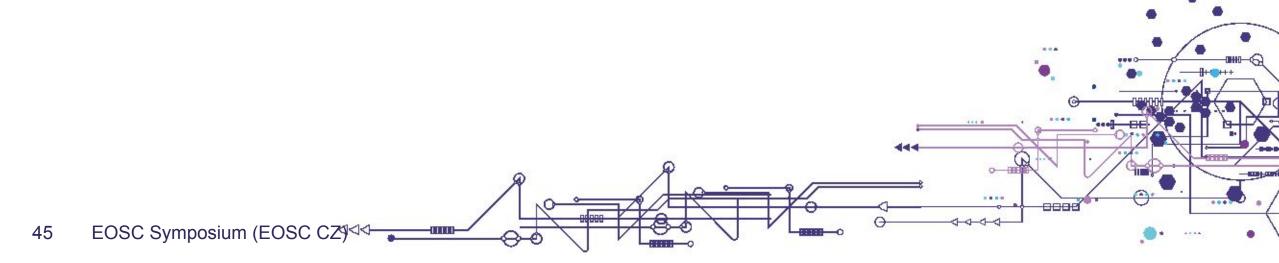
- 26 members, 14 universities and research institutes from all corners of the Czech Republic, 4 RIs
- Experimental and computational work
- The aim is to map needs of the Czech MSE community for storage and sharing of FAIR^t data, to design domain-specific EOSC-CZ repository and to get involved in the development of data management standards with similar activities around the world.





Inspired by Petr Čermák, MGML RI / MFF UK Praha

- FAIR data principles: <u>https://www.go-fair.org/fair-principles/</u>
- FAIR findable, accessible, interoperable, reusable
- What is missing?







....

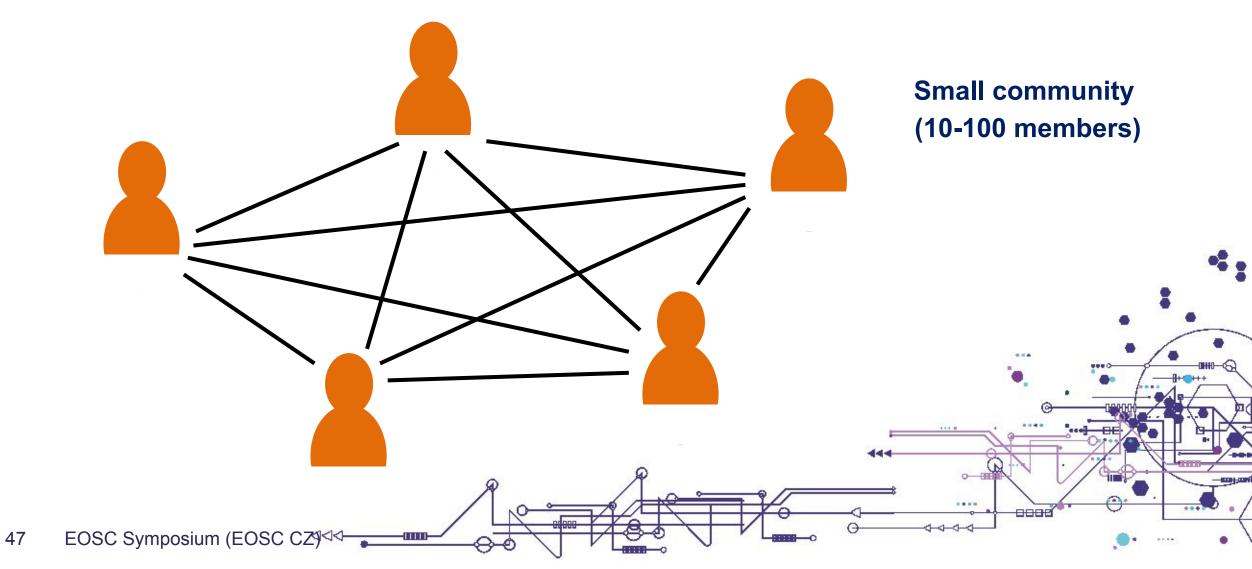
anae

Inspired by Petr Čermák, MGML RI / MFF UK Praha

- FAIR data principles: <u>https://www.go-fair.org/fair-principles/</u>
- FAIR findable, accessible, interoperable, reusable
- What is missing?
- FAIR^t = FAIR + trustworthy
- We suggest that data in our repository should be FAIR^t or at least provide information about their verification state.

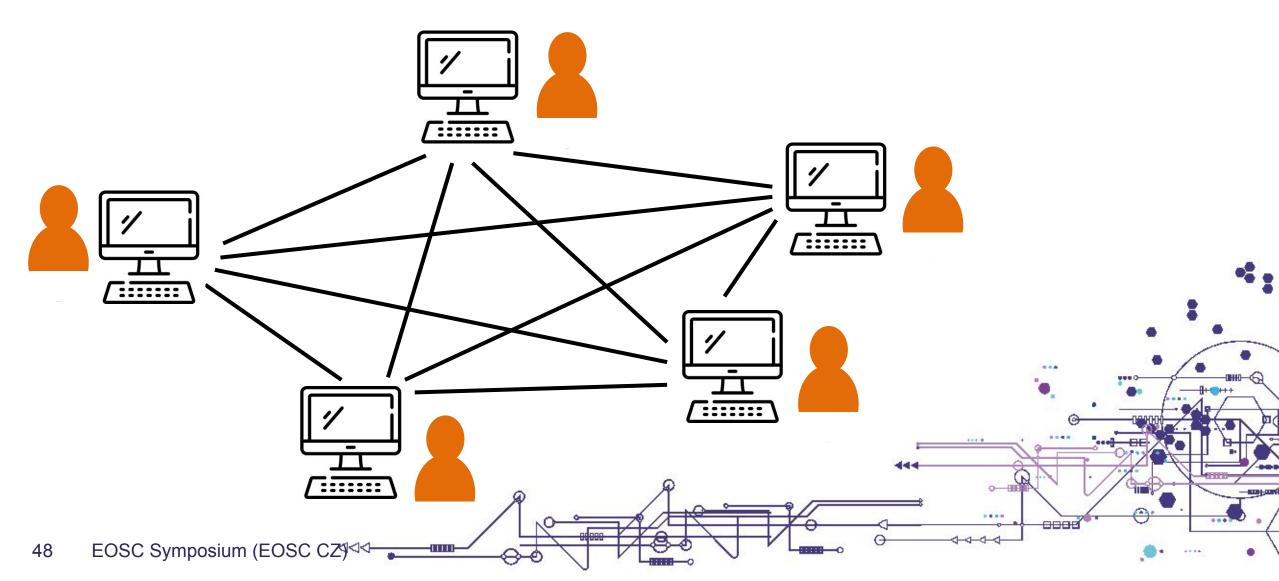
... with partners from Prague, Brno, Vienna, Ghent, Heidelberg, ...





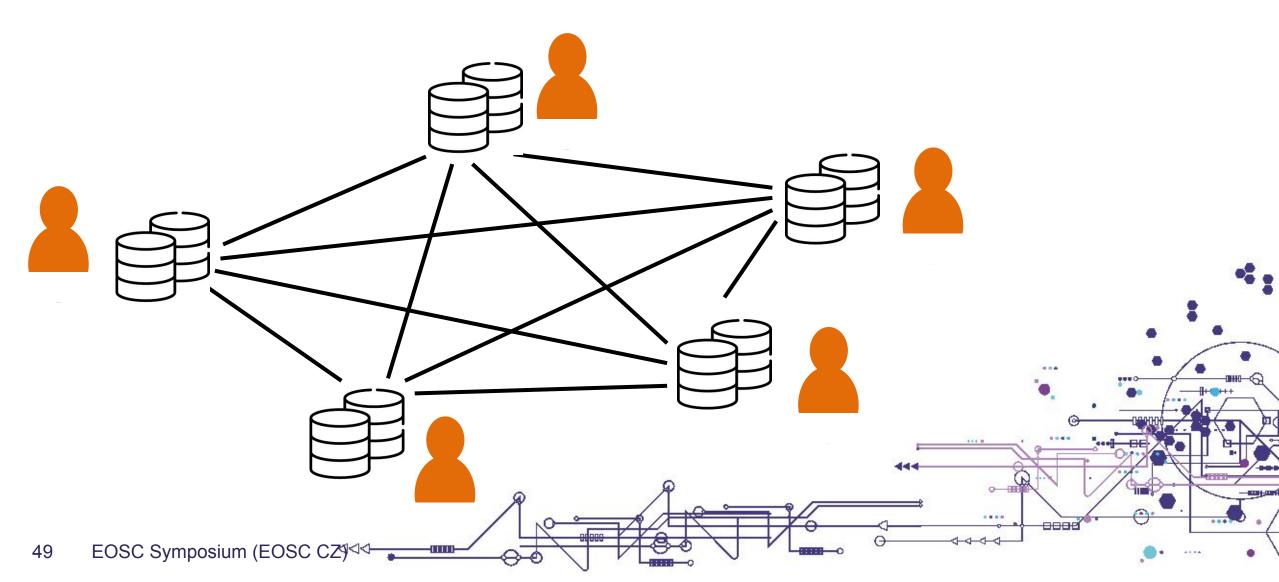


Sharing 'hot' data – decentralised repository





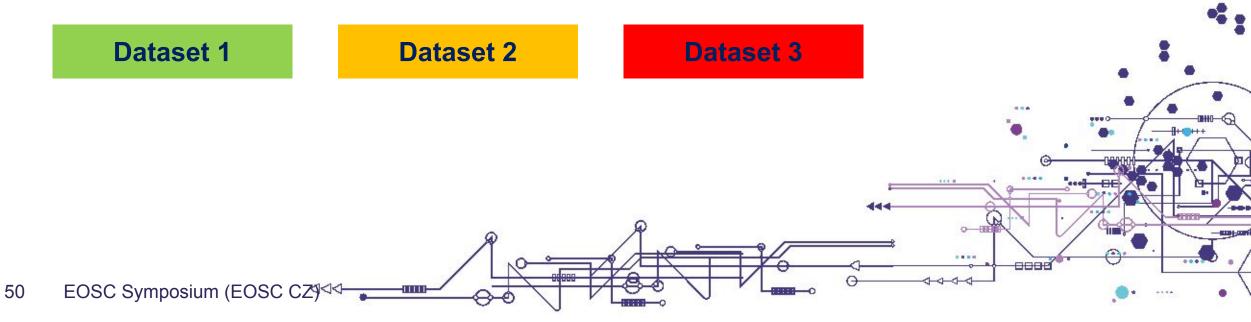
Sharing 'hot' data – decentralised repository





Decentralised community repositories with web-based architecture

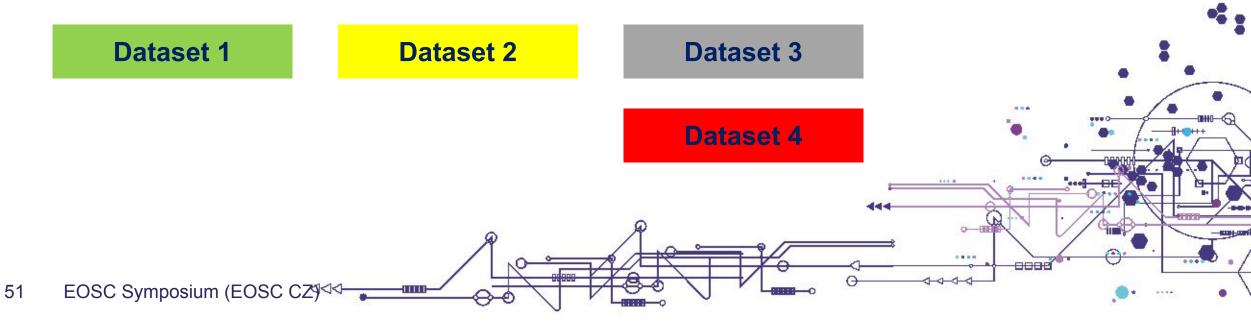
- Linked Data principles and metadata models
- Solid Pod and Solid principles
- Social network-like technology and principles
- Data verification within a small community to achieve trustworthiness (FAIR^t)
- Colour-coded data verification state:





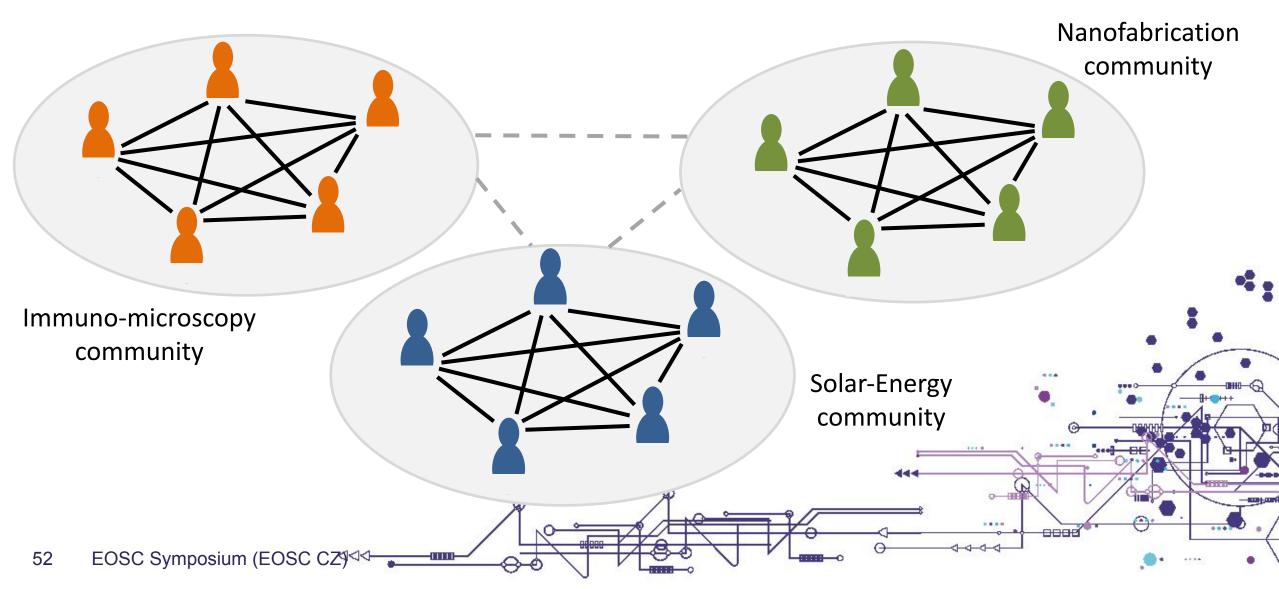
Decentralised community repositories with web-based architecture

- Linked Data principles and metadata models
- Solid Pod and Solid principles
- Social network-like technology and principles
- Data verification within a small community to achieve trustworthiness (FAIR^t)
- Colour-coded data verification state:





Interoperability of decentralised repositories



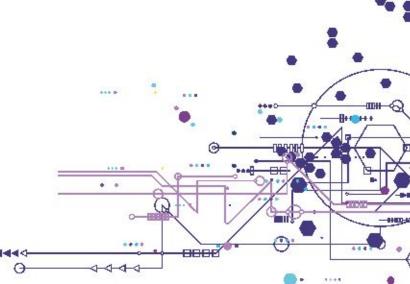




Thank you for your attention ndi-wg-matech@e-infra.cz

More information at <u>https://namidas.eu</u> (from Dec 2022)







LINDAT Services for EOSC

Pavel Straňák



Introduction



EOSC-CZ WG Humanities and the Arts

- FAIRification of data landscape in Arts and Humanities
- Data management
- Training
- Technical solutions
 - LINDAT language technologies as web services (lindat.cz/services)
 - machine translation, named entities, keywords, ... via REST API
 - A key solution is a repository system with comprehensive support of data licensing scenarios:
 - CLARIN-DSpace

55 EOSC Symposium (EOSC CZ)

CLARIN-DSpace for EOSC



anae

Development

- 13 years of development and maintenance experience
- Originally adaptations of DSpace for CLARIN centre: LINDAT-DSpace
- Substantial improvements targeted at FAIR sharing of datasets
- ... while keeping DSpace ease of deployment
- more deployments within CLARIN RI (and some outside):
 - CLARIN-DSpace
- collaborative development at <u>https://github.com/ufal/clarin-dspace</u>
 - most development still by the LINDAT team
 - lately substantial participation of Oxford University
 - an upgrade to the latest DSpace 7 code base
 - testing in December 2022
 - deployment in 2023

56 EOSC Symposium (EOSC CZ∮</

Any license (including custom licenses and signing)

- More metadata:
 - CMDI (Component Metadata Infrastructure)
 - BibTeX + more (for direct citation of datasets)
 - OpenAIRE integration
 - Clarivate DataCitationIndex, OLAC ... metadata
- Control Panel additions (backups, AAI debugging, health checks ...)
- UI makeover (search-centric, no hierarchies, submission: auto-suggest)
- Statistics (Matomo, UI integration + standalone application)
- Versioning via DC relations (data, new submission, UI, search)

57 EOSC Symposium (EOSC CZ)

CLARIN-DSpace for EOSC

Main Features

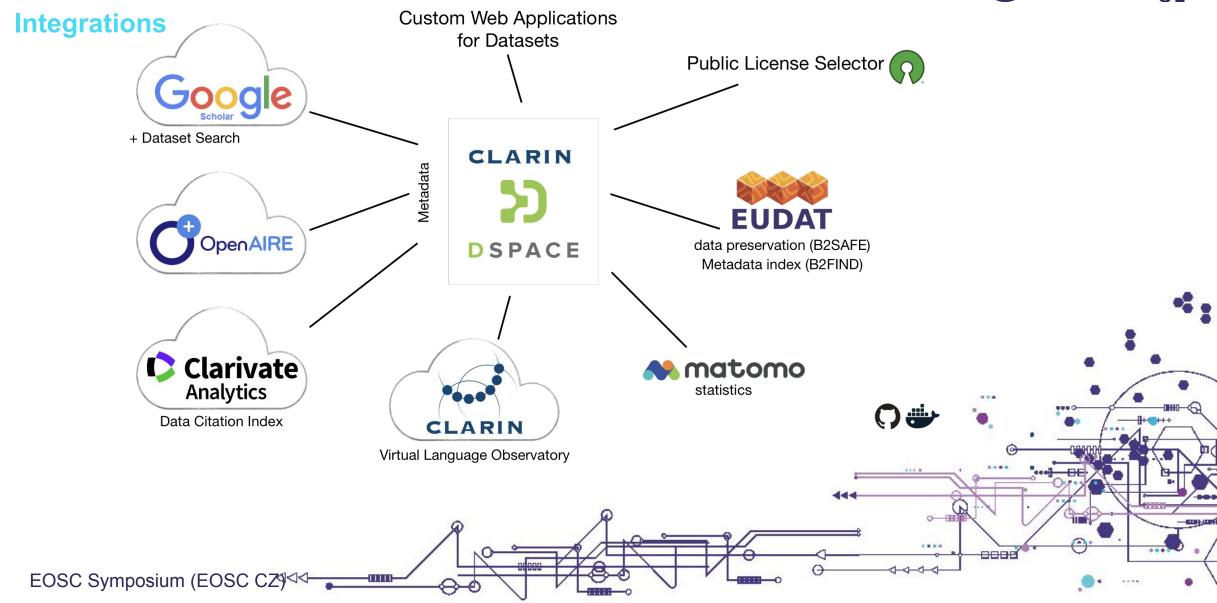
- EPIC Handles (European PID Consortium)
- Licensing framework:





CLARIN-DSpace for EOSC





- CLARIN-IS (Árni Magnússon Institute for Icelandic Studies)
- + several more (Czechia, Spain, Brasil)

+ virtualisation platform in Czech EOSC?

EOSC Symposium (EOSC CZ∮< 59

CLARIN-DSpace for EOSC

Deployments

- LINDAT/CLARIAH-CZ (Prague, Czechia)
- Oxford Text Archive (University of Oxford, UK)
- CLARINO (Bergen University, Norway)
- CLARIN.SI (Jožef Stefan Institute, Slovenia)
- CLARIN-PL (Politechnika Wrocławska, Poland)
- CLARIN-IT (ERCC and CNR, Italy)
- CLARIN-LT (Vytautas Magnus University, Lithuania)
- SWE-CLARIN (Språkbanken Text, Gothenburg)





....







Czech EOSC, during Czech Presidency!

cesne[•]



EOSC Symposium 2022, Prague, 14.11.2022

VŠB TECHNICKÁ | IT4INNOVATIONS ||||| UNIVERZITA | NÁRODNÍ SUPERPOČÍT OSTRAVA | CENTRUM

Introduction

Jindřich Krejčí

Institute of Sociology of the Czech Academy of Sciences Czech Social Science Data Archive (ČSDA) - CESSDA Service Provider CESSDA ERIC Widening & Outreach WG







- WG EOSC-CZ on the Social Sciences
- Sociology, economics, political science, social geography, psychology, law...
- Landscape analysis; EOSC relevant procedures, standards & tools; issues of multidisciplinary research; data management; training

62 EOSC Symposium (EOSC CZ)

Domain data archives - added value

- Why Open Access to Research Data?
 - Data reuse data in the scientific knowledge cycle
 - Enahancing quality of research incl. its transparency
 - Data storage & protection
 - Permanent curration
 - Ways to access the data
 - Understandable data
- Domain specific knowledge:

Data acquisition, data management, data organisation, reach metadata, tools, technologies, standards, context of other data, training

=> FAIR Data

63 EOSC Symposium (EOSC CZ)

vs. Data repository as a black box



Social science data services



TOOP

- Roper Center (USA 1957), ZA Zentralarchiv f
 ür Empirische Sozialforschung (Germany 1960, now GESIS), ICPSR - Inter-university Consortium for Political and Social Research (USA 1962), Steinmetz Archive (Netherlands 1964)...
- Czechoslovakia/Czech Republic: First efforts in 1967; Sociological Data Archive in 1998
 -> Czech Social Science Data Archive
- Data library vs. data archive: Archivation of virtual data is a permanent process
- Created by researchers in response to their needs
 => data ready to reuse; building data sharing culture
- Data sharing needed for comparative research (time, cross-national)
- Longitudinal/international comparative research programmes since 1970s/1980s

Landscape CZ Social Science Data

Universities, Czech Academy of Sciences, public research institutes

• data producers & data users

Official statistics Administrative data "New data types"



Other disciplines, multidisciplinary research

Commercial sector data for public use

Research Infrastructures: ČSDA (Czech Social Science Data Archive), ESS- CZ (European Social Survey), SHARE-CZ (Survey of Health, Ageing and Retirement in Europe), GGP-CZ (Generations & Gender Programme); ISSP (International Social Science Programme), EVS (European Values Study) ...





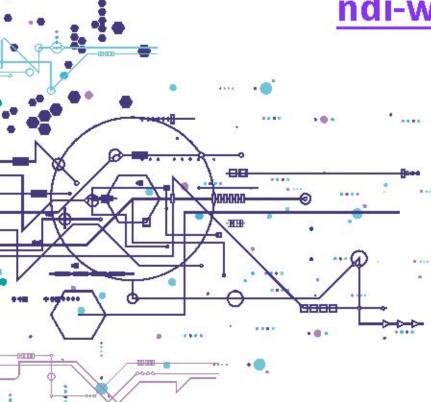
Expectations from EOSC

- Upgraded data sharing culture
- Political and economic aspects: speeding up long-lasting processes, sustainability
- Interoperability, complementarity, cooperation economy of scale
- Widening the perimeter of data; multidisciplinary environment
- Risks: black box instead of understandable data ready for reuse
 = no benefits, just costs => priorities: data policies, FAIR data principles, strengthening of the data sharing culture

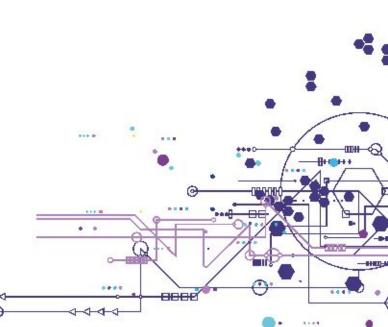
66 EOSC Symposium (EOSC CZ)



Thank you for your attention ndi-wg-social-sciences@e-infra.cz









Czech EOSC, during Czech Presidency! – Panel Discussion

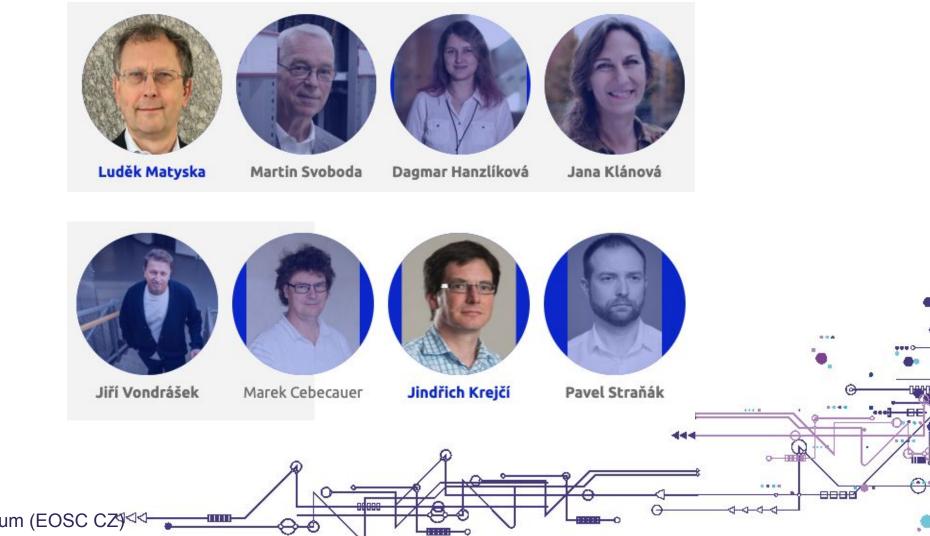
L. Matyska, M. Svoboda, D. Hanzlíková, J. Klánová, A. Strachotová, M. Cebecauer, P. Straňák, J. Krejčí

EOSC Symposium 2022, Prague, 14.11.2022

B TECHNICKÁ | IT4INNOVATIONS | UNIVERZITA | NÁRODNÍ SUPERPOČÍTAČOVÉ OSTRAVA | CENTRUM

Panelists





69 EOSC Symposium (EOSC CZ∮⊲⊲ ____



Thank you for your attention <u>eosc-info@e-infra.cz</u>

More information at <u>e-infra.cz/EOSC</u>

....

