

## Cestovní zpráva

# International Data Week 2023: A Festival of Data Salzburg, Rakousko 22. 10. – 26. 10. 2023

### Účastnice:

Petra Černošáková, vedoucí odd. 82 Centrum repozitářů a správy metadat

Hana Heringová, vedoucí ref. 821 – Národní centrum pro perzistentní  
identifikátory

Hana Vyčítalová, koordinátor metadatové podpory, odd. 82 Centrum repozitářů a  
správy metadata



Říjen 2023

## Účel cesty

Účelem cesty byla účast na konferenci International Data Week 2023. Čtyřdenní hybridní konferenci pořádaly organizace The International Science Council's Committee on Data (CODATA), World Data System (WDS) a the Research Data Alliance (RDA), celkem se jí zúčastnilo 702 odborníků z celého světa. Konference zahrnovala setkání RDA 21st Plenary, setkání a prezentace výstupů pracovních a zájmových skupin RDA a přednášky odborné vědecké konference SciDataCon. Cílem cesty bylo dozvědět se o nových trendech v oblasti výzkumných dat a navázat osobní kontakty s odborníky z dalších institucí. Účastnice se zaměřily zejména na přednášky s vazbou na perzistentní identifikátory, práci s metadaty, interoperabilitu a evidenci repozitářů a jejich atributů. Na konferenci byl prezentován poster „National PID Support in the Czech Republic: Where we are headed“ Petrou Černošlávkovou a Hana Heringovou a proběhla osobní účast Hana Vyčítalové na schůzce pracovní skupiny RDA Data Repository Attributes. Kromě toho účastnice měly příležitost s odborníky z jiných zemí prodiskutovat téma základního metadatového modelu pro popis výzkumných dat (plánovaný výstup projektu CARDS) a novinky a zkušenosti v oblasti perzistentních identifikátorů (PID).

## Časový průběh cesty

neděle 22. října: odjezd vlakem z Prahy, přestup v Linzi, příjezd do Salzburgu (19:49)

pondělí 23. – středa 25. října: účast na International Data Week 2023

čtvrtek 26. října: účast na International Data Week 2023, odjezd vlakem do Prahy s přestupem v Linzi (příjezd do Prahy 22:40)

## Průběh navštívené akce/jednání

Celá konference se skládala ze tří akcí: RDA 21st Plenary, setkání a prezentace výstupů pracovních a zájmových skupin RDA a přednášky odborné vědecké konference SciDataCon. Přednášky z RDA 21st Plenary byly společné pro všechny účastníky, ostatní dvě akce měly vždy několik paralelních sekcí. Následuje výčet sekcí, které účastnice navštívily a popis nejvýznamnějších přednášek.

Navštívené sekce:

Pondělí 23. října

- Opening Plenary Session: Welcome to IDW 2023 'Spatial Data Science: Geographic Context Matters
- Building a semantic interoperability framework: towards FAIR mappings and crosswalks
- Data Repository Attributes WG: Recommended Descriptive Attributes for Data Repositories
- PID provision – the current Austrian approach, global challenges and thoughts on future directions
- Software Source Code: Mastering the Art of Research Software Metadata and Metrics
- Initiatives for data and service federation that can connect local and global practices

- Structuring Semantic Information with Respect to Conventional Metadata (Syntactic) Structures
- FAIR Beyond Discoverability: Exploring Technology Approaches and Challenges through DDI-CDI Implementation

#### Úterý 24. října

- Data Discovery Paradigms IG: Ten recommendations for data repositories/catalogues to improve data discoverability
- IDW 2023 Plenary Session 'Inclusivity in Open Science while advancing research assessment and career pathway impact'
- Unleashing the potential of domain specific metadata
- Accelerating research and innovation through data collaboration at national scale
- Research data needs of the Photon and Neutron Science community IG: Tools and Methods working towards the FAIR Facility
- RDA Adoption & Impact Plenary Session
- Data Retention and Disposal infrastructure and policies
- Enhancing Data and Information Discovery and Reuse Across Disciplines - The Roles Of Quality Management In Open Science
- PID Exchange - a curated information resource for PID adopters
- WorldFAIR: Emerging Cross-Domain Perspectives - Envisioning Solutions - the Cross Domain Interoperability Framework (CDIF)
- The Way to FAIR: From the instrument to the researcher -- Calibrating the ecosystem

#### Středa 25. října

- Data Cubes as a tool to organise data in data spaces
- Community-based catalogue of requirements for trustworthy technical repository service providers
- Plenary Session 'Data and global challenges: data, science, trust and policy
- The FAIR hourglass: A framework for the practical implementation of FAIR data and services
- CODATA / WDS Organisational Plenary Sessions
- Why Persistent Identifiers and How Their Connectivity Enhances Research
- Revising the Versioning Principles: The Road to Actionable Recommendations
- Data Management Plans Presentations
- Let's talk about FAIR mappings! Towards common practices for sharing mappings and crosswalks

#### Čtvrtek 26. října

- Sensitive Data IG: Sensitive Data classifications and protocols across regions
- Generalist repository "coopetition" to enhance data sharing and Discovery
- RDA/WDS Certification of Digital Repositories IG: TRUST Principles and Repository Certification Landscape
- Sensitive Data classifications and protocols across regions
- Implementing FAIR Workflows: A Proof of Concept Study in the Field of Consciousness
- Equity and Inclusion: Universal access to PIDs

- Closing Plenary Session 'Ethics, data science, and AI in dialogue with one another and society'

## **Vybrané přednášky se zásadním významem pro NTK a její činnost**

### **Setkání RDA WG Recommended Descriptive Attributes for Data Repositories**

Člen H. Vyčítalová

Obecenstvu bylo představeno 17 atributů doporučených pro popis digitálních repozitářů. Na třech příkladech repozitářů byl popis představen v praxi včetně hodnocení, jak těžké je informace o attributech dohledat.

V diskusi byly probrány možnosti, jak toto doporučení uvést do praxe a jaká bude budoucnost pracovní skupiny (viz sdílené poznámky <https://bit.ly/46II9U4>).

### **FAIR IMPACT (<https://fair-impact.eu>)**

projekt EU, 36 měsíců, začátek červen 2022

V rámci projektu je vytvářen:

- framework pro mapování a crosswalks (FAIR mapping framework)
- pokyny a metodiky, jak správně vytvářet mapování - 1. verze - listopad/prosinec
- strojově použitelný obecný výměnný model pro sdílení mapování
- platforma pro sdílení best practices
- analýza typologií mapování.

Jako významný standard byl zmiňován Datamodel for Simple Standard for Sharing Ontological Mappings (SSSOM) (<https://mapping-commons.github.io/sssom/>).

24. 11. 2023 10:00 – 13:00 CET proběhne workshop Documenting mapping community practices <https://fair-impact.eu/events/fair-impact-events/documenting-mapping-community-practices>

Výstupy z workshopů a dalších akcí jsou v [Zenodu](#) a na [Youtube](#).

### **Upcoming tools for mappings from the FAIRCORE4EOSC**

Projekt FAIRCORE4EOSC (<https://faircore4eossc.eu/>)

Produkt Metadata Schema and Crosswalk Registry

- Vytváří registr pro evidenci metadatových schémat a mapování (crosswalks).
- Handle pro registrovaný obsah, podporuje verzování, každá verze vlastní PID.
- Ve vývoji je tvorba UI pro vytváření crosswalks mezi schématy.
- Podporované formáty CSV, JSON, XML, SKOS, OWL, SHACL (interní formát)
- Autorizace uživatelů, kteří smějí registrovat nové příspěvky.
- Metadata jsou veřejně přístupná vždy, drafty mohou být veřejné nebo soukromé, publikované výstupy veřejné, nesmí se odstraňovat nebo modifikovat.
- Registr podporuje licencování a provenance.

### **Generalist Repository Ecosystem Initiative (GREI)**

<https://datascience.nih.gov/data-ecosystem/generalist-repository-ecosystem-initiative>

GREI repositories - Zenodo, Dryad, Dataverse, Figshare, Mendeley

Iniciativa se snaží o vytvoření společného souboru služeb napříč různými všeobecnými repozitáři, podporu fair data, repository best practises, PIDs, vylepšení vyhledávacích služeb v rámci repozitářů.

Generalist Repository Comparison Chart <https://zenodo.org/records/3946720>

### **Certifikace repozitářů**

Zvedá se administrativní poplatek u Core Trust Seal za žádost o certifikaci – 3000 EUR.

Platnost certifikace je 3 roky, pak je nutné žádat o recertifikaci.

Další možností je Nestor Seal for Trustworthy Digital Archives.

Certifikaci se věnuje jedna ze skupin RDA Certification of Digital Repositories IG

(<https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html>).

### **Domain-specific metadata template**

Na konferenci byla v několika příspěvcích zmíněna důležitost dobře sestavených vkládacích formulářů pro datasey. Šablonám se věnuje projekt CEDAR - Center for Expanded Data Annotations and Retrieval (<https://more.metadatacenter.org/>).

### **Poskytování PIDs - současný rakouský přístup, globální výzvy a úvahy o budoucím směřování**

Bylo představeno rakouské národní konsorcium DataCite, o které je v Rakousku velký zájem a není potřeba jej výrazně inzerovat, jelikož se zájemci hlásí sami. Nyní má konsorcium 21 členů. Rakousko zatím nemá národní strategii pro perzistentní identifikátory (PIDs), každá instituce si musí poradit sama.

O PIDs je třeba uvažovat na lokální, národní i mezinárodní úrovni a brát v úvahu dlouhodobý horizont – kdo za jejich využití bude platit, kdo bude přidělené PIDs udržovat v budoucnu.

Jaká je role výzkumníků? Je možná automatizace přidělování PIDs? V právě vyvíjeném rakouském národním CRIS systému existuje prompt pro výzkumníky k založení ORCID iD apod. Pomohly by jim školicí materiály?

Prezentace „EOSC-A Taskforce - PID policy and implementation“, která se zabývá mapováním PIDs pro EOSC PID Graph.

### **PID Exchange – informační zdroj pro ty, kteří implementují PIDs**

Společná sekce dvou skupin RDA: „Working with PIDs in Tools Working Group“ a „National PID Strategies Interest Group“ (Hana Heringová členkou).

První prezentace shrnula předběžný plán aktivit pro zájmovou skupinu pro národní strategie pro PIDs – mj. dále rozvíjet příklady implementace v různých částech světa, dopracovat seznam doporučení pro tvorbu národní strategie, který vypracovala původní pracovní skupina stejného zaměření a dále rozpracovat „value proposition“ národních strategií.

Skupina pro užívání PIDs v různých nástrojích (tools) prezentovala modelovou integraci identifikátoru pro vzorky IGSN do systému ResearchSpace. Tato modelová integrace mimo jiné umožnila shromáždit poznatky o best practice postupech při zavádění nového typu PID.

Na závěr byla krátce diskutována myšlenka zřízení platformy s informačními zdroji k PIDs – PID Exchange.

## Závěry/shrnutí

Účast na International Data Week byla klíčová pro další rozvoj Národního centra pro perzistentní identifikátory, zejména k získání aktuálních informací o různých perzistentních identifikátorech (RAiD, IGSN, PIDINST). Z hlediska metadatové oblasti byly zásadní informace o nástrojích pro fairifikaci dat, trendech v sémantické interoperabilitě repozitářů výzkumných dat a dalších tématech souvisejících s metadaty a sdílením dat. Účastníci využili také příležitosti s odborníky z jiných zemí prodiskutovat téma základního metadatového modelu pro popis výzkumných dat (plánovaný výstup projektu CARDS) a kritéria pro výběr tohoto modelu. Velmi cenné také bylo získání nových kontaktů a navázání spojení s dalšími odborníky ve zmíněných oblastech.

Další aktivity:

- Prezentace posteru „National PID Support in the Czech Republic: Where we are headed“
- Osobní účast Hana Vyčítalové na schůzce pracovní skupiny RDA Data Repository Attributes
- Komunikace se spoluorganizátory akce PIDfest, spolupořádané Národní technickou knihovnou v roce 2024 a propagace této akce na konferenci IDW 2023
- Účast Petry Černolávkové v panelové diskusi v rámci bloku Accelerating research and innovation through data collaboration at national scale

## Dokumentace

Web konference IDW 2023: <https://internationaldataweek.org/idw2023/>

Web RDA 21st Plenary session: <https://www.rd-alliance.org/rdas-21st-plenary-programme>.

Web SciDataCon: <https://www.scidatacon.org/IDW-2023-Salzburg/programme>

Seznam posterů: <https://www.rd-alliance.org/rdas-21st-plenary-poster-exhibition>

[Pracovní podklady](#) pro schůzku WG Data Repository Attributes

Videozáznamy přednášek jsou dostupné zatím po přihlášení v aplikaci Whova

Poster Národního centra PID uložen v repozitáři NTK: <https://doi.org/10.48813/2vax-5y37>

## Anotace ZC

Účelem cesty byla účast na konferenci International Data Week 2023. Konference zahrnovala RDA 21st Plenary, setkání a prezentace výstupů pracovních a zájmových skupin RDA a přednášky vědecké konference SciDataCon. Účelem cesty bylo zjistit nové trendy v oblasti výzkumných dat, zúčastnit se zejména přednášek s vazbou na perzistentní identifikátory, práci s metadaty, interoperabilitu a evidenci repozitářů a navázat osobní kontakty s odborníky z dalších institucí. Na konferenci byl prezentován poster „National PID Support in the Czech Republic: Where we are headed“ Petrou Černohlávkovou a Hanou Heringovou a proběhla osobní účast Hana Vyčítalové na schůzce pracovní skupiny RDA Data Repository Attributes.