

Technická architektura NRP

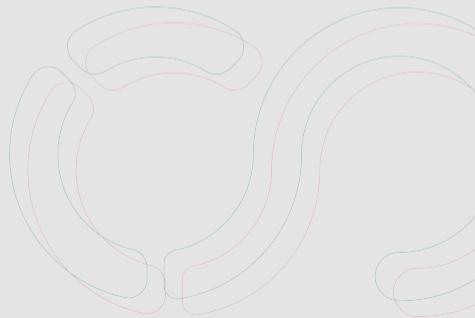
David Antoš

CESNET

22. 11. 2023



Spolufinancováno
Evropskou unií



Národní repozitářová platforma

- technická architektura NRP
- návaznost na pomocné systémy
- nestrukturovaná úložiště v národní e-infrastruktuře
 - plus další služby pro kontext

Repozitář v kontextu EOSC (CZ)

- úložiště dat s metadaty
- s podporou pro FAIR principy
- webové rozhraní a API pro strojový přístup
- organizačně: s odpovědností za uložená data
- certifikovatelný na CTS
 - srovnej <https://www.clarin.eu/content/checklist-clarin-b-centres>
- má obsahovat „citovatelné datové sady/záznamy“
 - a zajišťovat jejich konzistenci/neměnnost
 - uvnitř jsou samozřejmě i rozpracované
- repozitář je technické, personální a procesní řešení pro dlouhodobé ukládání a zpřístupňování citovatelných digitálních objektů

Národní datová infrastruktura

- se skládá zejména z následujících komponent
 - Národní repozitářové platforma
 - Národní metadatový adresář
 - Národní katalog repozitářů
 - úložiště v infrastruktuře
 - pomocné systémy

Repozitář v NRP

- repozitář/instance/tenant
 - standardní repozitář
 - určený specifickou konfigurací
 - URL
 - vizuální podoba (skin)
 - popis v něm používaných metadatových modelů
 - podpora pro specifická metadata (např. vizualizace)
 - s určeným řízením přístupu
 - navázaný na další systémy v NDI
 - zejména exportuje metadata, je registrován, ...
 - nepoužití dostupné komponenty má být řádně zdůvodněno
- repozitář typicky oborový nebo institucionální
 - kdo chce repozitář v NRP, musí se podílet na jeho nastavení

NRP

- implementace založena na dostupných nástrojích
- distribuovaná, multitechnologická infrastruktura
- s vrstvou pro vytváření repozitářů prostou konfigurací
 - nikoli tedy plnou separovanou instalací(!)
- NRP musí řešit i pravidelné aktualizace
 - zejména bezpečnostní
- implementace NRP musí zahrnovat
 - schopnost vytváření repozitářů na žádost
 - jednoduše, co nejvíc automaticky
 - (to neznamená na kliknutí uživatele)
 - dokumentaci
 - pokročilou podporu

Kdo je uživatel

- koncový uživatel repozitáře
 - ty řeší správce repozitáře
- správce konkrétního repozitáře
 - to je pro nás partner
- rámcově odpovídá rozdělení virtuální organizace
 - se členy a správci

Co může být implementací NRP

- „něco jako repozitář“
- v návaznosti na okolní systémy (za chvíli)
- potřebuje tým, který zajišťuje jeho provoz
- může/nemusí mít vlastní HW
 - nicméně hodila by se kompatibilní storage vrstva
- v základu DSpace, Invenio, ARL
 - plus storage S3 a deployment v Kubernetes

Prostředí a okolní systémy I

- Národní metadatový adresář
 - provozován projektem EOOSC CZ
 - metodická podpora v EOOSC CZ+CARDS
 - NRP řeší napojení a exporty metadat do NMA
- Technická implementace vydávání PID
 - metodická podpora v CARDS
 - konsorcia pro členství v poskytovatelích PIDů v CARDS
 - NRP řeší vydávání PIDů vlastním ukládaným objektům
- Správa uživatelů
 - primárně v rámci e-INFRA, kapacita pro specifické problémy alokována v EOOSC CZ
 - specifika pro NRP?

Prostředí a okolní systémy II

- Nástroje pro přenosy dat
 - vývoj a centrální komponenty v EOSC CZ
 - NRP řeší konektory z jednotlivých systémů a produkční provoz
- Nástroje pro monitoring infrastruktury
 - podobně
- Obecná úložiště (zejména u výpočetních zdrojů)
 - pořizovaná z e-INFRA, zaměřena na rozpracovaná data
- Dlouhodobé binárně spolehlivé archivy
 - systém k dispozici
 - NRP přebere jejich vývoj a provoz
- Národní katalog repozitářů
 - v CARDS

Další implementace repozitářových systémů

- musí stále splňovat znaky repozitáře
- nutný tým správců s expertízou k provozu
- možnost vytváření instancí na žádost uživatelské skupiny
 - včetně pokročilé podpory správců takového repozitáře
- návaznost na další systémy platformy
- závazek udržitelnosti

**Napište nám
info@einfra.cz**