

Seminář

Projekt NRP pro členy PS EOOSC CZ
(představení projektu a diskuse)

online



Spolufinancováno
Evropskou unií



Anotace

Seminář tematicky navazuje na přednášku „Systémy pro Národní repozitářovou platformu (NRP)“. Tento seminář proběhne ve formátu interaktivní diskuse s odborníky, nikoliv jako standardní přednáška.

Cílem semináře je:

- Poskytnout členům pracovních skupin pro implementaci EOSC v ČR informace o projektu NRP z výzvy Open Science I. Účastníci semináře mohou očekávat informace o struktuře projektu i diskusi o službách, které jsou v rámci projektu NRP obsaženy.
- Otevřít prostor pro diskusi nad přesahy a návaznostmi obsahu jednotlivých projektů, jež přispívají k budování Národní datové infrastruktury (NDI) v rámci celého ekosystému EOSC CZ.

Rámcový program

Luděk Matyska, hlavní manažer projektu a vedoucí KA1;
ředitel CERIT-SC na ÚVT MU

- Představení rámce projektu NRP a jeho klíčových prvků
- Diskuse

David Antoš, vedoucí oddělení DÚ na CESNET, vedoucí PS
ANDI EOSC CZ, vedoucí KA3

- Vývoj a provoz NRP
- Diskuse

Michal Růžička, specialista pro FAIR data z CERIT-SC ÚVT
MU, vedoucí PS ZS EOSC CZ, vedoucí KA5 NRP

- Služby v NRP
- Diskuse

Anastasia Avdeeva, projektová manažerka v Ústřední
knihovně UK, vedoucí KA7 NRP

- Školení a osvěta v oblasti funkcionalit a služeb NRP
- Diskuse

Jan Kolouch, metodik kybernetické bezpečnosti CESNET,
vedoucí KA6 NRP

- Kyberbezpečnost a podpůrné služby
- Diskuse

Ilona Trtíková, Český sociálněvědní datový archiv SOÚ AV
ČR, CESNET

- Podpora FAIRifikace výzkumných dat
- Diskuse

Výzva OP JAK Open Science I

- Cílem výzvy je podpořit implementaci iniciativy European Open Science Cloud (dále též „EOSC“) v ČR v souladu s koncepčním dokumentem „*Architektura implementace EOSC v ČR*“.
- Konkrétně výzva podporuje **vybudování distribuované Národní repozitářové platformy (dále též „NRP“)** jako klíčového stavebního prvku Národní datové infrastruktury (dále též „NDI“), která vytvoří základní prostředí v ČR pro správu výzkumných dat, jejich ukládání, sdílení a práci s nimi v souladu s principy FAIR tak, aby byla vyhledatelná, dostupná, interoperabilní a znovu využitelná.
- **NRP zajistí kromě technického řešení i portfolio klíčových služeb a taktéž vzdělávání uživatelů, datových odborníků a dalších relevantních pracovníků ve výzkumných organizacích.**

Národní repozitářová platforma pro výzkumná data (projekt NRP)

- Realizace: 2024 – 2028
- Koordinátor: sdružení CESNET
 - Hlavní projektový manažer: Luděk Matyska
- Partneři (9+1):
 - Univerzity: *MU, VŠB-TUO, UK, ČVUT, UPOL*
 - AV ČR: Heyrovského ústav, BTÚ, KNAV, SSČ
 - NTK (bez finančního příspěvku)
- Rozpočet: 1,3 miliardy korun
 - Cca 260 milionů ročně
 - Celkem 400 milionů hardware a související náklady (energie)

Interní struktura projektu

(klíčové aktivity)

- KA1 (Luděk Matyska, CESNET):
Řízení projektu
- KA2 (David Antoš, CESNET):
Vývoj a provoz NRP
- KA3 (bude vybrán, MU): **Pilotní repozitáře**
- KA4 (Slávek Licehammer, CESNET & MU): **Integrace služeb NRP**
- KA5 (Michal Růžička, MU):
Vývoj a provoz nových služeb NRP
- KA6 (Jan Kolouch, CESNET):
Kyberbezpečnost a podpůrné služby
- KA7 (Anastasia Avdeeva, UK):
Školení a osvěta v oblasti funkcionalit a služeb NRP

Cíle projektu

- Vybudovat technické a procesní zázemí NRP
 - Dostatečná kapacita
 - Vysoké kyberbezpečnostní standardy
 - Shoda (compliance) s relevantními právními předpisy
- Nabídnout široce použitelný systém řízení přístupu
 - Společně s e-INFRA CZ
 - K dispozici bude i mimo NRP
- Vytvořit a provozovat ekosystém základních a pokročilých služeb
 - Důraz na uživatelskou použitelnost
- Propojit NRP s dalšími systémy
 - Výpočetní a analytické kapacity
 - Přístroje vytvářející velké objemy dat
- Zajistit metodické a dokumentové zázemí
 - Metodici, dokumentace
 - Školení
 - Komunita data stewardů

Přínosy projektu

- Dostatečná kapacita a softwarové zázemí
 - 50+PB úložného prostoru
 - Objektové úložiště a Kubernetes pro běh služeb
- Tři podporované repozitářové systémy
 - CLARIN DSpace, Invenio, ASEP
 - Procesy a metodická podpora pro tvorbu a správu repozitářů
 - Catch all repozitář
- Napojení na NMA
 - Předávání základních metadat
- Nástroje řízení přístupu k repozitářům a datům v nich uložených
- Vysoké kyberbezpečnostní standardy
 - Podpora práce s citlivými daty
- Systémová podpora plánování správy výzkumných dat
 - "Actionable DMP"
 - DSW a jeho rozšíření
- FAIRifikace dat
- Správa licenčních modelů
- Klasifikace velkých přístrojů, nástroje a metodika propojení a přenosu dat do NRP
- Napojení na výpočetní kapacity
 - e-INFRA CZ, obecně i další prostředky
- Podpořená komunita data stewardů
- Školicí materiály, dokumentace

Repozitáře přímo v NRP

- **Catch-all repozitář (CESNET)**
 - First and last resort
 - Měl by sloužit pro uložení dat, pro která (ještě) není k dispozici jiné řešení
 - Využívá repozitářový systém Invenio
- **4 pilotní repozitáře**
 - Early adopters
 - Zpětná vazba pro vývoj NRP a ekosystému služeb
 - Podpořeno vytvoření nebo přenos repozitáře do NRP, integrace se službami
 - Obsahová podpora repozitáře či péče o data mimo rozsah podpory NRP
- **Repozitář pro biologická zobrazovací data (MU, CzechBioImaging)**
 - Nový systém, plánuje využití repozitářového systému Invenio
- **Národní úložiště pro biodiverzitní data (UK, PŘF)**
 - Vlastní repozitářový systém
- **Molecular Biophysics Database (BTÚ)**
 - Existující systém, plánuje využití repozitářového systému Invenio
- **HUME Lab (MU, FF)**
 - Pilotní nasazení repozitářového systému Islandora

NRP v čase, shrnutí

- První pilotní plně použitelná verze k dispozici v průběhu roku 2025
 - Omezená kapacita i rozsah/parametry některých služeb
 - Catch all repozitář v pre-pilotní verzi již začátkem 2025
 - Kyberbezpečnostní zajištění, podpora práce s citlivými daty
- 2026 sběr zkušeností a zpětné vazby
 - Druhé velké dotazníkové šetření
- Plný rozsah kapacity i služeb v roce 2027
- Od roku 2025 bude k dispozici pokročilé prostředí pro tvorbu repozitářů a ukládání dat do nich
 - Vlastní tvorba a odpovědnost za repozitáře mimo rozsah NRP
- Přístupy k repozitářům i datům bude možné velmi citlivě řídit
- Bude k dispozici zázemí pro správu dat, metadatových a licenčních modelů
- NRP bude propojeno s výpočetním prostředím e-INFRA CZ
- Bude k dispozici odpovídající metodická podpora, dokumentace i školicí materiály

Vývoj a provoz NRP (David Antoš)

- KA2 vybuduje Národní repozitářovou platformu
 - Pořídí hardware, uživatelská kapacita cca 50 PB, geografická distribuce
 - Alespoň dvě místa pro citlivá data
- Bude provozovat tři implementace repozitářových systémů
 - CESNET Invenio, CLARIN DSpace, ASEP/ARL
 - Které bude poskytovat „jako službu“ (tedy i pro OS II)
 - Partnerem je správce repozitáře
- Pro jiné implementace repozitářů nebo repozitářových systémů bude provozovat
 - Prostředí pro běh aplikací v Kubernetes a úložiště S3
 - Opět „jako službu“
- Pokročilá podpora
- Rozvíjí základní funkce těchto nástrojů
 - Vývoj v KA2 vs. ostatní KA

Služby v NRP 1/2 (Michal Růžička)

- **Cíl: Pomoci uživatelům plnit FAIR principy**
- Obecně zaměřením na strojovou zpracovatelnost, integrovatelnost a automatizaci
- Nástroje pro práci s metadatovými profily v repozitářích NRP
 - Podpora pro snadné napojení na Národní katalog repozitářů
 - Usnadnění práce s oborovými metadaty
- Plnění FAIR atributů pro datasey v NRP repozitářích
 - Podpora FAIR vestavěná do NRP
- Metriky – informační nástroj s tipy pro zlepšení FAIRness
 - Předpokládá se spolupráce s oborovými skupinami v OS II pro oborová specifika
- Technická podpora pro vznik komunitních standardů v podobě FAIR Implementation Profiles (FIPs)
- Podpora plánování správy daty
 - Průvodce pro uživatele: Doporučování vhodných repozitářů, (meta)datových standardů apod. při zohlednění citlivosti dat apod.
 - Podpora procesu data management planningu
 - DMP dokument „vedlejší efekt“
 - API, integrovatelnost

Služby v NRP 2/2 (Michal Růžička)

- Automatizace sběru dat a metadat
 - Přímé ukládání dat z vědeckých přístrojů do datových úložišť NRP
 - Včetně agregace nebo dělení měření apod.
 - Následné další využití a organizace dat v rámci NDI
 - Včetně vytěžování metadat
 - PID pro zajištění jednoznačnosti
- Integrace výpočetních workflows
 - Napojení NRP repozitářů jako datového zdroje nebo úložiště pro výpočetní workflows v e-INFRA CZ
 - Podpora pro Galaxy, Lexis, Jupyter notebooky
- Průběžně udržovaná roadmapa služeb NRP

Služby v NRP – licence (Michal Růžička)

- Nástroj pro správce repozitářů v NRP
 - Definice licencí, které jejich repozitáře podporují
- Podpora pro uživatele repozitářů
 - Průvodce výběrem licence
 - Vodítka pro výběr vhodné licence pro jejich data v daném repozitáři
- Kontrola dodržování licence při uložení dat v daném repozitáři
 - Vazba na řízení přístupu k datům
 - Podpis smlouvy on-line nebo proces individuální žádosti o přístup
- Vazba na další prvky NRP – např. služby plánování práce s daty apod.
 - Strojově zpracovatelná informace o licenci

Školení a osvěta v oblasti funkcionalit a služeb NRP (Anastasia Avdeeva)

- Prohlubování znalostí a dovedností pro užívání NRP/NDI
- Cílová skupina:
 - Zaměstnanci datové podpory institucí (data steward, data kurátor, IT, management vědy, ...)
- Základní koncept – “training trainers”
- KA7.1 Metodické materiály pro software/služby NRP
 - usnadní práci s repositářovými systémy NRP
 - v editovatelných verzích
- KA7.2 Komunita data stewardů
 - zpětná vazba z užití metodických materiálů NRP
- KA7.3 Konzultační služby a školicí služby pokročilé podpory NRP
 - helpdesk pokročilé podpory (level 2)

Kyberbezpečnost a podpůrné služby (Jan Kolouch)

- **KA 6.1 - Zabezpečení NRP a zajištění kyberbezpečnostního dohledu**
 - Zajištění fyzické bezpečnosti
 - Zajištění organizační bezpečnosti (dokumentace aj.)
 - Soulad s NIS2, GDPR, DSA aj.
 - Zajištění provozní bezpečnosti.
 - pravidla pro správce a uživatele
 - bezpečnostní a provozní dohled
 - reaktivní bezpečnost
 - proaktivní bezpečnost
 - Vybudování systému centrálního log managementu.
 - Vzdělávání v oblasti kybernetické bezpečnosti
- **KA 6.2 - Compliance a řízení ISMS i s péčí o citlivá data**
 - **Systém řízení, jehož účelem je soulad s pravidly** (legislativní, smluvní závazky, interní)
 - **Zajištění důvěrnosti, dostupnosti a integrity veškerých dat a informací v projektu NRP**
 - Nastavení standardů pro uživatele, technologie a procesy.
 - specifická pravidla a standardy pro zvláštní kategorie dat
 - soulad projektových činností a výstupů projektu s právním předpisy, podmínkami projektu a závaznými standardy,
 - **ochrana dat a informací před neoprávněným přístupem,**
 - **ochrana osobních údajů** a soukromí uživatelů (GDPR aj.),
 - **transparentní komunikace služeb směrem k uživatelům.**

• KA 6.3 - Prouživatelský přístup

- Zvyšování kvality ekosystému NRP/NDI z uživatelského pohledu
 - Naplňování potřeb cílových skupin:
 - koncových uživatelů,
 - správců služeb a repositářů
 - Použitelnost prostředí z uživatelského pohledu
- Mapování potřeb vědecké komunity
 - dotazníky, rozhovory, přirozená komunikace
- Uživatelské testování NRP/NDI
- Navrhování a testování změn + pomoc při implementaci
- Zjištění z mapování potřeb uživatelů
- Monitoring ekosystému z uživatelského hlediska
- Případná další doporučení

• KA 6.4 - Uživatelská podpora a ServiceDesk

- Uživatelská podpora a Service Desk (24/7/365)
 - přijímání, zpracování a řešení uživatelských požadavků
 - přijímání hlášení o incidentech
 - monitoring stavu služeb jak směrem k uživateli, tak k vlastníkovvi služby
- Reporting dostupnosti služeb
- Self-Service portal
- podpora data stewardům aj.

Podpora FAIRifikace výzkumných dat (Ilona Trtíková)

- Cílem poskytnout uživatelům/správcům repozitářů informaci o míře plnění principů FAIR
 - pro jednotlivé datasety ukládané v rámci NRP
 - zpětná vazba, související doporučení
- Předpokládaná úzká spolupráce s oborovým skupinami
- Technická příprava pro komunitní plánování naplňování FAIR principů – **FAIR Implementation Profile Wizard:**
 - volba standardů, technologií, nástrojů a postupů pro jednotlivé komunity/obory
 - metodika GO FAIR

Děkujeme za pozornost

www.eosc.cz



Spolufinancováno
Evropskou unií



MUNI
ICS

cesnet

VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA

IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM