



cerit
scientific cloud

cesnet
.....

VŠB TECHNICKÁ UNIVERZITA OSTRAVA | IT4INNOVATIONS NÁRODNÍ SUPERPOČÍTAČOVÉ CENTRUM



EOSC Implementační architektura v ČR

Luděk Matyska

ředitel ÚVT MU, člen Řídící rady VI e-INFRA CZ, ředitel CERIT-SC

3. listopadu 2021



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

European Open Science Cloud

- Federace (web) FAIR dat a souvisejících služeb
 - Federace existujících i budoucích datových zdrojů
 - Otevřený ekosystém souvisejících služeb
- Budování Evropského výzkumného datového prostředí
 - European Research Data Commons
 - Ukládání a zpřístupnění FAIR dat (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable)
- Velmi obecné formulace
 - Příležitost pro státy a jejich instituce definovat konkrétní naplnění
 - V každém případě musí obsahovat prostředí pro práci s daty

Implementace EOSC v ČR

- Součást rozvoje Open Science
 - Dva z 8 pilířů: FAIR data a vlastní EOSC
- Vytvoření Národní datové infrastruktury
 - Ekosystém FAIR dat a souvisejících služeb
- První fáze (od konce minulého roku) koordinována MŠMT
 - EOSC Platforma
 - Diskuse vhodného uchopení implementace EOSC v ČR
- Výsledkem [Architektura implementace EOSC v ČR](#) (červen 2021)
 - Pracovní dokument, podklad pro další kroky

- Prostředí pro ukládání, zpřístupnění a další práci s FAIR daty v ČR
 - Technické i procesní/operační standardy
 - Technická interoperabilita
- Nestaví se na zcela zelené louce
 - Existující repozitáře a prostředí pro ukládání dat
 - Plány pro institucionální repozitáře
 - Plány pro národní repozitář
- Federace (mozaika, patchwork) konkrétních systémů
 - Národní repozitářová platforma
 - Repozitáře oborově-vědních clusterů
 - Další repozitáře
- Základní metadatový adresář

Národní repozitářová platforma

- Základ Národní datové infrastruktury
- Vycházející z principu cloudové platformy
 - „Univerzální“, nikoliv však homogenní centrální prostředí
 - Podpora tvorby „tenantů“
 - Vlastní instance, např. institucionální repozitář
 - Silná podpora pro řízení přístupu (data nejsou automaticky „Open“)
 - **One approach does not fit all**
- Součástí i standardy pro manipulaci s daty
 - Nutná podmínka propojeného prostředí NDI
- Data potřebují také zázemí pro jejich zpracování
 - Koordinace s e-INFRA CZ, ko-budování výpočetních kapacit

Repozitáře oborově-vědních clusterů

- Řada subjektů (zejména velké výzkumné infrastruktury) již dlouhá léta buduje vlastní systémy ukládání a zpřístupnění dat
- Ty je nutné propojit do ekosystému Národní datové infrastruktury
 - Při zachování jejich specifických vlastností a potřeb
 - Přirozené rozšíření mozaiky národní repozitářové platformy
- Nezbytné aktivity v rámci implementace EOSC v ČR
 - Podpora zpřístupnění i mimo původní oblast zájmu
 - Napojení (prostřednictvím standardů a API) na národní repozitářovou platformu
 - Poskytnutí metadat do základního metadatového adresáře
 - Případné rozšíření kapacit a služeb

Základní metadatový adresář

- Klíčová složka Národní datové infrastruktury
- Primární vstupní místo
 - Zajišťuje „F“ ve FAIR
 - Obsahuje základní metainformace o všech datech v rámci NDI
 - Podrobnější metainformace v rámci vlastních tenantů či samostatných repozitářů
- Zastřešuje všechny repozitáře
 - Povinné plnění metadaty
- Přímou propojeno s národní repozitářovou platformou
 - Podpora automatického generování základních metadat pro uložená data
- Silné řízení přístupu
 - Ani metadata nejsou automaticky Open

Aktuální etapa implementace EOSC v ČR

- K dispozici základní Architektura implementace EOSC v ČR
- Ustaveny základní (průřezové) pracovní skupiny
 - **Architektura** (David Antoš, CESNET)
 - **CORE služby** (Michal Růžička, MU)
 - **Metadatový adresář** (Petra Černohlávková, NTK)
 - **Vzdělávání** (Radka Římanová, UK)
- Web pro EOSC – <https://e-infra.cz/eosc>
- Do 15. října proběhl první sběr zájmu o zapojení do pracovních skupin
 - Aktuálně se chystají první schůzky (CORE služby dnes odpoledne zde v NTK)
- Příprava pracovních skupin oborově-vědních clusterů
- Aktivace všech významných subjektů v rámci ČR

Další kroky

- Koordinovaná práce průřezových pracovních skupin
 - Technický „výbor“, tvořený vedoucími a dalšími osobnostmi
 - Podmínky pro zapojení dalších členů
- Očekávané výstupy
 - Definice základních vlastností a služeb NDI
 - Cesta k (de-facto) povinným standardům
 - Příprava aktivit do projektu budování národní repozitářové platformy
 - Jádru NDI a jeho služby
 - Metadatové schéma a základní principy plnění a práce s ním
 - Kritické „pojítka“ celé NDI
- Inicializace pracovních skupin oborově-vědních clusterů

Očekávaná finanční podpora

- Primárně prostřednictvím OP JAK
 - Klíčová výzva Open Science v rámci OP JAK
- Jednotlivé součásti
 - Vytvoření a provoz Národní repozitářové platformy
 - Širší spolupráce klíčových institucí, koordinační role e-INFRA CZ
 - Vytvoření a provoz základního metadatového adresáře
 - Oborově-vědní clusterů a jejich začlenění do NDI
 - Vytvoření a podpora odborného a organizačního zázemí NDI
 - Sekretariát EOSC ČR
 - Cílená podpora odborného a kvalifikačního růstu lidí
 - Datoví kurátoři

- Implementace EOSC v ČR jako příležitost pro
 - Konsolidaci a FAIR zpřístupnění dat produkovaných v ČR
 - Shodu na základních principech přístupu a práce s daty
 - S přihlédnutím ke specifickým požadavkům (např. práce s citlivými daty)
 - Plány práce s daty a podpora jejich realizace
 - Meziinstitucionální a meziinfrastrukturní spolupráce
 - Odstranění nebo alespoň snížení bariér pro (mezioborové) využívání dat
 - Systémovou podporu FAIRifikace
 - Ocenění (karierní růst) kvalitní péče o FAIR data
- Lépe spravovaná a FAIR zpřístupněná data jako základ budoucího rozvoje otevřené excelentní vědy v ČR i mimo ni



Děkuji za pozornost

e-infra.cz