

Pracovní skupina Základní služby

Pracovní verze k 22. 9. 2021; vypracoval Michal Růžička (ruzicka@ics.muni.cz) a kol.

1 Úvod

Jedním z hlavních pilířů implementace European Open Science Cloud (EOSC) v ČR je vybudování Národní repozitářové platformy (NRP) pro vědecká data. Jedná se o samostatnou část Národní datové infrastruktury (NDI). Ta má sloužit jako jeden ze základních prvků infrastruktury pro vědecká data v ČR, spolu s prostředím pro ukládání nestrukturovaných dat, výpočetním prostředím a prostředím pro spolupráci.

Pracovní skupina Základní služby (PS ZS) má za cíl definovat základní služby NRP pro efektivní práci s vědeckými daty v ČR. Návrh vychází zejména z potřeb podpory práce s daty definovanými v rámci EOSC-Core v [EOSC Strategic Research and Innovation Agenda \(SRIA\)](#). Využity byly také další zdroje, např.:

- [EOSC Task Forces](#) zaváděné na evropské úrovni,
- RDM témata identifikovaná [WG-RDM.cz](#),
- témata z diskusí kolem pracovní skupiny k ESOC na MŠMT.

Cílem skupiny je analyzovat soudobé prostředí, standardy a trendy v každodenní práci s vědeckými daty v ČR, EU a ve světě a na základě toho připravit standardy, politiky, doporučení a návrhy technických řešení Základních služeb NRP, které budou podporovat principy [FAIR Data](#) (Findable, Accessible, Interoperable, and Reusable Data). Týkat se mohou např. persistentních identifikátorů (PID) a práce s nimi, podpory vyhledávání a zpřístupňování dat, správy uživatelů a identity (Authentication and Authorization Infrastructure [AAI]), podpory plánování správy vědeckých data (Data Management Plan[ning], DMP), FAIRifikace výzkumných dat apod.

Funkční a kvalitativní požadavky na NRP, definované i ve spolupráci s ostatními pracovními skupinami, budou převážně implementovány pracovní skupinou Architektura národní datové infrastruktury (PS ANDI). Důležitým úkolem pracovní skupiny bude také expertní podpora v oblasti standardů práce s výzkumnými daty v ČR všem zájemcům. Doporučení PS ZS budou pro infrastrukturu závazná po schválení steering committee.

2 Cíle

Pracovní skupina se bude zabývat definicí základních služeb NRP pro podporu práce s (FAIR) daty:

- Služby pro podporu práce s persistentními identifikátory (PID),
- požadavky na služby řízení přístupu a autorizaci uživatelů pro podporu jejich spolupráce přes NRP,
- služby pro podporu plánování správy výzkumných dat,
- služby pro FAIRifikaci výzkumných dat apod.

Konkrétněji se bude PS ZS zabývat např.:

- Požadavky na kvalitu služeb (dostupnost, rychlost, odezva, ...) v NRP.
- Spolupráce na definici funkčních požadavků pro technickou implementaci v NRP s PS ANDI.
- Práce s persistentními identifikátory (PID):
 - Výběr a doporučení konkrétního standardu nebo standardů pro PID.

- Bude zohledněna kompatibilita s [EOSC PID Policy and Architecture](#).
 - Primárně pro potřeby NRP, ale použitelné i pro „externí“ zájemce o PID.
- Definice požadavků na práci s PID v rámci NRP (včetně např. možností strojového volání/zpracování, integrace s data managementem apod.).
- Dlouhodobá podpora/rozvoj PID služeb v rámci NRP (kontinuální podpora pro přidávání nových typů entit, např. softwarové kontejnery, workflows, ..., sledování a adaptace nových technologií).
- Řízení přístupu/autorizace uživatelů pro podporu spolupráce:
 - Definice požadavků na autentizaci uživatelů (např. požadavky na integraci s existujícími národními a mezinárodními systémy AAI v akademické sféře).
 - Definice požadavků na řízení přístupu k datům i metadatům v NRP (např. požadavky na funkce kontroly vlastníka nad sdílením dat).
- Plánování správy výzkumných dat:
 - Definice politik práce s daty v rámci NRP pro různé oblasti (např. PID, zálohování, bezpečnost, ...).
 - Definice služeb pro podporu plánování správy výzkumných dat (např. podpora tvorby DMP, podpora měření FAIR atributů dat v NRP apod.).
- Služby pro FAIRifikaci výzkumných dat:
 - Analýza možností plnění/podpory FAIR principů v rámci NRP (např. dohledatelnost dat přes Metadatový adresář, přidělování PID různým typům entit, doporučené datové formáty apod.).
 - Na základě toho doporučení politik, standardů apod. + doporučení k implementaci v NRP.
 - Definice/doporučení FAIR metrik.

V budoucnu je možné rozšiřovat činnosti PS ZS o další oblasti dle aktuální potřeby.

3 Pracovní postupy

Pracovní skupina je otevřena zájemcům o členství se zkušenostmi v oblasti práce s vědeckými daty, počet členů pracovní skupiny není omezen. Aktivně spolupracuje s relevantními projekty v ČR, EU i ve světě. V základu volně navazuje na [aktivity](#) otevřené skupiny [WG-RDM.cz](#). Cílem je získat expertní skupinu reprezentující v maximální míře subjekty pracující s výzkumnými daty v ČR.

Členové pracovní skupiny se dle své odbornosti a zájmů organizují do podskupin, které pracují na jednotlivých aktivitách a výstupech pracovní skupiny směřující k dosažení stanovených cílů.

Výstupy pracovní skupiny potvrzené steering committee budou závazné pro implementaci infrastruktury a všechny zájemce o využívání jejích služeb.

3.1 Předpokládání členové

Osoby se zkušenostmi se správou výzkumných dat na větší akademické instituci – plnění požadavků výzkumníků a výzkumných skupin na práci s daty:

- Ukládání a aktivní zpracování dat (i větších objemů),
- řízení přístupu k datům,
- sdílení s jinými skupinami,
- veřejné publikování zpracovaných dat,
- datové kurátorství – plánování správy dat, metadatový popis, volba vhodných formátů / datové konverze, opatřování PID.