

Výstupy a provazby projektů Open Science

Úvod

Tento dokument shrnuje a dává do souvislosti výstupy projektů z výzev Open Science OP JAK, jmenovitě dvou komplementárních individuálních projektů systémových: IPs EOSC-CZ, IPs CARDS a projektu z výzvy Open Science I: projekt Národní repozitářová platforma pro výzkumná data (dále jen NRP). V rozsahu známém autorům dokumentu také zahrnuje klíčové relevantní milníky velké výzkumné infrastruktury e-INFRA CZ. Dokument předjímá budoucí rozšíření o plánované výstupy aktuálně připravovaného projektu výzvy Open Science II (dále jen OS II).

Kompletní podrobný přehled oficiálních plánovaných výstupů, cílů a produktů projektů IPs EOSC-CZ, IPs CARDS a NRP, jak jsou zachyceny v jejich Chartách, resp. Studii proveditelnosti, zachycuje tabulka, která je přílohou tohoto dokumentu. Tento dokument představuje komentář a „klíč“ k snazšímu pochopení údajů v tabulce, včetně explicitně popsanych vzájemných vztahů a vazeb.

Při práci s tímto dokumentem je třeba akceptovat, že popisované výstupy, cíle, produkty a jejich provazby nejsou statické a definitivní. Detailní naplnění jednotlivých cílů – např. konkrétní forma metadatového modelu NMA, konkrétní vlastnosti první verze systému pro práci s licencemi – bude výsledkem průběžné realizace projektu a jako takové podléhá dynamickému vývoji. Zatímco hlavní cíle a plánované výstupy jsou známy a mají jasnou časovou osu (viz i podklady v příloze), konkrétní plnění jednotlivých cílů je závislé na předchozím krocích a rovněž na vzájemné spolupráci napříč projekty (např. výše zmíněný metadatový model NMA je výstupem projektu IPs CARDS, technická realizace, za kterou odpovídá projekt IPs EOSC-CZ, je pak samozřejmě ovlivněna rozsahem a složitostí výsledného modelu; ten se navíc bude v průběhu realizace projektů dále vyvíjet). Proto, i přes významnou míru koordinace dílčích kroků a akterů, je průběžné upřesňování finální formy jednotlivých výstupů součástí pracovních postupů těchto projektů. Toto je zvláště významné vzhledem k přípravě projektu OS II, kde konkrétní funkce prvních verzí programů, které budou k dispozici v rámci projektu NRP a jsou zpravidla plánovány na rok 2025, budou významně upřesněny v důsledku vzájemné spolupráce při přípravě projektu OS II a upřesňování požadavků, které řešitelé OS II budou mít na vznikající infrastrukturu NRP a její služby (které konkrétní vlastnosti služeb NRP je nezbytné mít již v prvních verzích a které bude možné odložit na pozdější etapy).

V předložené podobě proto dokument představuje konkrétní „snapshot“ k termínu jeho vzniku (začátek září 2024) a bude průběžně aktualizován dle vývoje realizace projektů, potřeb iniciativy EOSC CZ a navazujících aktivit.

Dokument čerpá primárně z těchto podkladů:

- Charta IPs EOSC-CZ
<https://www.eosc.cz/media/3582629/charta-projektu-ips-eosc-cz.pdf>
- Charta IPs CARDS
<https://www.techlib.cz/default/files/download/id/87565/charta-ips-cards.pdf>
- Základní obsah projektu NRP (ze Studie proveditelnosti projektu, autoři měli k dispozici plnou verzi Studie)
https://www.eosc.cz/media/3713976/nrp_zakladni_obsah_projektu_2024.pdf

Národní iniciativa EOSC CZ

Zasazení projektů z výzev OP JAK: Individuální projekty systémové – Výzkum a vývoj I, Open Science I, II a III do iniciativy EOSC CZ, která zastřešuje implementaci EOSC v ČR. Časová osa realizace projektů je na Obr. 1 níže.

- Projekty mají společný cíl: vybudovat Národní datovou infrastrukturu (NDI).
- Projekty jsou součástí iniciativy EOSC CZ.
- Jednotlivé projekty pokrývají vzájemně se doplňující oblasti implementace EOSC v ČR:
 - o IPs EOSC-CZ a IPs CARDS vytváří organizační zázemí pro celou iniciativu EOSC CZ a zajišťují její monitoring;
 - o IPs EOSC-CZ dále implementuje a zprovožňuje klíčové prvky technického řešení, zejména Národní metadatový adresář (NMA) a systém pro velkoobjemové přenosy dat, společně s architekturou autentizační a autorizační infrastruktury EOSC CZ;
 - o projekt NRP zajistí základní úložnou kapacitu NDI, poskytne tři základní repozitářové systémy a vytvoří systém klíčových služeb, nezbytných pro fungování technologického zázemí NDI, a to včetně rozsáhlé metodické podpory budoucích uživatelů; současně vytváří komplexní compliance a kyberbezpečnostní architekturu a implementaci jejich klíčových prvků;
 - o projekt výzvy OS II pak nad tímto základem buduje prostředí péče o konkrétní FAIR datové repozitáře, FAIR data a podle potřeby je rozšiřuje o oborově (doménově) specifické služby a nástroje;
 - o projekt/y výzvy OS III pak napomohou dosud vybudované zázemí a know-how dále využívat a diseminovat do vědecké komunity.



Obrázek 1: Časová osa projektů Iniciativy EOSC CZ

Výstupy jednotlivých projektů

Na základě komplementarity a posloupnosti projektů, jež byly popsány výše, lze jejich cíle a výstupy rozdělit do několika tematických okruhů/logických celků, v nichž se jednotlivé projekty vzájemně doplňují a navazují na sebe. S ohledem na toto dělení budou v následujícím textu blíže rozepsány aktuálně známé cíle, výstupy a produkty všech výše uvedených projektů z výzev OP JAK. Řazení jednotlivých cílů, výstupů a produktů je v rámci každého tematického okruhu chronologické. Na výstupy zde není nahlíženo z pohledu jednotlivých projektů, ale z pohledu celé iniciativy EOSC CZ a její implementace. Pro dohledání přesného znění projektových cílů, výstupů a produktů, přesného data jejich plnění, vazby na odpovídající projekt (IPs EOSC-CZ, IPs CARDS, NRP) a konkrétní klíčovou aktivitu využijte přehledovou tabulku v příloze tohoto dokumentu. Tato tabulka zároveň obsahuje interpretaci projektových výstupů, která má za cíl přiblížit znění

projektového textu do jazyka bližšího vědecké komunitě. Pro detailní popis projektových cílů, výstupů a produktů je pak nezbytné nahlédnout do Chart projektů IPs EOSC-CZ a CARDS a Studie proveditelnosti NRP.

Na konkrétní tematické okruhy současně mapujeme potenciální výstupy připravovaného projektu OS II, které byly zatím identifikovány dle zápisu z Porady ohledně přípravy projektu Open Science II (UK) ze dne 20.8.:

Tvorba nových repozitářů; Inovace repozitářů; Licencování dat; Data management plan – oborová schémata šablon; Metadatová schémata – velmi významně; Šetření mezi uživateli, uživatelskými komunitami (analýzy a definice potřeb); Laboratorní deníky; Vzdělávání, e-learning, helpdesk L1-L3; Budování komunit; Zcela mimo popis výstupů z analýzy vyplývá i to, že pracovní skupiny počítají s tím, že bude využita HW kapacita pro ukládání výzkumných dat v rámci infrastruktury NRP.

Tematické okruhy sdružující výstupy, cíle a produkty uvedených projektů jsou následující:¹

- **Koordinační a základní monitorovací zázemí iniciativy EOSC CZ**

Tento okruh zahrnuje organizačně-komunikační a monitorovací/analytické zázemí pro celou iniciativu EOSC CZ. Tato činnost je od poloviny roku 2023 svěřena dvěma organizačním jednotkám, vytvořeným za účelem koordinace implementace EOSC v ČR a monitorování a koordinace budování NDI – Sekretariátu EOSC CZ a Školicímu centru EOSC CZ. Těm v této činnosti od poloviny roku 2024 pomáhá napříč republikou koordinovaná komunita Data stewardů.

(Adresuje bod Budování komunit.)

Dále tento okruh zahrnuje komunikační kanály budované v rámci IPs EOSC-CZ pro celou iniciativu EOSC CZ, především portál <https://www.eosc.cz> a pravidelný newsletter. Toto zázemí je v současnosti již k dispozici, jeho další rozvoj má spíše charakter postupného vylepšování poskytovaných služeb. Specifickým výstupem je pak dokument s analýzou možností dlouhodobé udržitelnosti iniciativy EOSC resp. obecně pojaté péče o FAIR výzkumná data. Tento dokument, Analýza modelů udržitelnosti NDI po skončení IPs EOSC-CZ, je plánována na druhou polovinu roku 2025 a jeho cílem bude představení a analýza možných modelů budoucího financování a organizace péče, kterou v současné době zajišťují primárně zde zahrnuté projekty, plány velkých výzkumných infrastruktur v oblasti péče o FAIR výzkumná data, vše s přihlédnutím k institucionálním plánům, které budou mimo jiné zahrnuty v návrzích projektů do výzvy OP JAK Výzkumná prostředí.

- **Technické zázemí pro NRP**

Okruh technického zázemí pro NRP zahrnuje hardwarové kapacity, které budou pořízeny v rámci projektu NRP. Tyto hardwarové kapacity jsou reprezentovány základní úložnou vrstvou, která vytvoří prostředí objektových úložišť v technologii Ceph. Ta je dále doplněna hardwarem ve formě bezprostřední výpočetní kapacity (nezbytné pro prohledávání, bezprostřední práci s daty, ...) postavené na prostředí Kubernetes. Vyjmenovaný hardware souhrnně nazvaný jako první verze repozitářové platformy bude k dispozici v druhé polovině roku 2025 (pokud nedojde k problémům

¹ Kurzívou je v následujícím textu uvedeno namapování zatím identifikovaných potenciálních výstupů v rámci přípravy OS II na v textu popsané tematické okruhy. Toto mapování uvádí vzájemné souvislosti potenciálních výstupů OS II s již naplánovanými výstupy, cíli a produkty existujících projektů. Současně je z vazby zřejmé, že upřesnění cílů projektu OS II bude zpětně ovlivňovat realizaci již běžících projektů, především formou zpětné vazby a kontinuálního upřesňování požadavků.

s výběrovými řízeními na nákup hardware) a nabídne uživatelskou kapacitu cca 18 PB. Nad touto základní hardwarovou platformou budou od konce roku 2025 k dispozici tři technologická řešení: repozitářové systémy, které se o základní uživatelskou kapacitu cca 18 PB podělí podle potřeby (nepředpokládáme apriorní alokaci, architektura systému je navržena tak, aby se kapacita v Ceph objektovém úložišti mohla dynamicky poskytovat podle potřeby). Zatímco vytváření nových repozitářů (uživatelskými skupinami, které mohou být podpořeny například z výzvy OS II) plně v prostředí NRP bude možné až od konce roku 2025, ve spolupráci s e-INFRA CZ projekt nabízí možnost pilotního ověření využitelnosti Ceph vrstvy již v současné době.

(Adresuje bod o využití HW: kapacita pro ukládání výzkumných dat v rámci infrastruktury NRP a Tvorba nových repozitářů.)

Dalším prvkem technického zázemí NRP je prostředí pro podporu vysokoobjemových datových přenosů, které je v pilotní verzi již k dispozici pro ověření konceptů; k integraci s repozitářovými platformami – a tedy využitelnosti pro tvorbu repozitářů – dojde ve stejné době, kdy bude prostředí NRP v první verzi k dispozici (konec roku 2025).

- **Propojení na výpočetní prostředí e-INFRA CZ**

Svlastním technickým řešením NDI úzce souvisí problematika propojení s prostředím e-INFRA CZ. Toto propojení musí zajistit možnost zpracování dat ve výpočetních prostředcích e-INFRA CZ (a slouží současně jako model propojení s případně institucionálním výpočetními prostředím, případně s prostředím velkých výzkumných infrastruktur, nejen e-INFRA CZ). V průběhu roku 2025 bude k dispozici architektura vzájemného propojení komponent (NMA, NRP, MetaCentrum, Datová úložiště, Superpočítač IT4Innovations) a budou implementovány první verze automatizace ukládání a přebírání dat, propojení s prostředím SensitiveCloud (ELIXIR CZ a e-INFRA CZ) pro práci s citlivými daty, integrace předávání dat mezi NRP a úložnými kapacitami Datových úložišť e-INFRA CZ.

- **Autentizační a autorizační infrastruktura (AAI)**

Architektura autentizační a autorizační infrastruktury a základní protokoly a API jsou k dispozici již nyní, je tedy možné již při návrhu nových služeb a nástrojů s tímto prostředím pracovat. Celá NDI bude pracovat s jednotnou identitou osob (i případně systémů) a bude k dispozici postupně rostoucí okruh atributů, na jejichž základě bude možné provádět základní autorizační rozhodnutí (řízení přístupu ke službám, repozitářům i datům). Prostor Ceph, Kubernetes i repozitářové platformy, stejně jako všechny služby, vyvíjené v rámci projektů IPs EOSC-CZ, NRP a relevantní služby z IPs CARDS, budou přímo na toto AAI napojeny. Detailnější architektura i implementace pro jemnější podporu řízení přístupu bude k dispozici od začátku roku 2025 (je ovšem třeba očekávat postupný vývoj, i ve vazbě na evropské projekty v této oblasti zejména AAR Tree a EOSC ENTRUST).

V Q3/2025, kdy bude k dispozici kapacita repozitářové platformy, bude možné nově vznikající repozitáře okamžitě připojit na AAI; informace o tom, jak to udělat, bude společně s iniciální architekturou AAI pro NRP dostupná nejpozději od Q1/2025.

(Adresuje bod Inovace repozitářů.)

- **Catch-all repozitář**

V současné době je k dispozici pilotní verze Catch-all repozitáře, určeného pro podporu správy long-tail dat vědecké komunity a obecně podporu pro ukládání vědeckých dat tam, kde zatím nejsou k dispozici vhodnější tematické repozitáře. Pilotní verze byla spuštěna nad hardwarem e-INFRA CZ, technicky je identická s budoucí verzí catch-all repozitáře, který poběží v prostředí NRP; je realizován v repozitářovém systému CESNET Invenio (jeden ze tří podpořených repozitářových systémů v projektu NRP). Jak bylo zmíněno výše, architektura národní

repozitářové platformy nevyžaduje apriorní alokace úložné kapacity konkrétní repositářové platformě, a tedy ani konkrétnímu repositáři – ty mohou růst podle potřeby. Relevantní je pouze celkový součet požadavků na ukládání dat, který v roce 2025 bude možné uspokojit do výše 18 PB přes všechny repositáře (včetně pilotních a catch-all) společně. Catch-all repositář je napojen na AAI a postupně budou přibývat další relevantní služby. Tento repositář tak bude představovat demonstrátor, jak repositáře v NRP mohou být vytvářeny a provozovány.
(Adresuje bod Tvorba nových repositářů.)

- **Pilotní oborové repositáře**

Součástí implementace NRP je také podpora ustavení a provozu 4 pilotních repositářů. Hlavním cílem je získat včasnou a bezprostřední zpětnou vazbu na použitelnost, kvalitu a dostupnost technického zázemí a služeb vyvíjených v rámci NRP. Pilotní repositáře ověří primárně funkcionalitu Ceph vrstvy pro potřeby repositáře, který nepoužije žádný z podpořených repositářových systémů, dále funkcionalitu repositářového systému Invenio (a to jak přenos existujícího repositáře, tak i vytvoření zcela nového) a rovněž ověří možnost implementace dalšího repositářového systému – konkrétně Islandora – v prostředí NRP.

Vývoj pilotních repositářů je již zahájen nebo zahájení proběhne do konce září 2024; pilotní repositáře budou v prostředí NRP integrovány okamžitě po jeho dostupnosti, tj. od Q3/2025. Zkušenost z instalace a následného provozu pilotních repositářů se promítne do dokumentace a metodické podpory všech, kteří budou mít zájem o tvorbu a provoz vlastních repositářů. Tato podpora by měla být k dispozici v průběhu Q4/2025.

(Adresuje body Tvorba nových repositářů a Inovace repositářů.)

- **Metadata**

Metadatový okruh slučuje kompletní zázemí pro tvorbu a správu metadat v rámci NDI, počínaje zřízením pracoviště pro podporu práce s metadaty (v roce 2023), přes tvorbu Národního metadatového adresáře (NMA, začátek roku 2024) harvestujícího metadata ze všech repositářů, které jsou součástí NDI, až po vytvoření nástroje pro správu metadatových profilů včetně kompletní metodiky (v roce 2025). Dílčí součástí tohoto tematického okruhu je postupné společné budování NMA, kdy projekt IPs CARDS dodává základní metadatový model pro NMA, projekt IPs EOSC-CZ dodává technické zázemí NMA a postupně rozšiřuje jeho kapacitu a také zvyšuje robustnost řešení. V současné době existuje pilotní verze NMA, která v tuto chvíli jako metadatový model využívá základ Dacite a která běží na hardware poskytnutém e-INFRA CZ; k přechodu na vlastní hardware projektu IPs EOSC-CZ dojde do konce roku 2024, z pohledu uživatelů by ale tato změna měla být nerozpoznatelná. Každý repositář, podpořený v rámci NDI a propojený na toto prostředí musí být schopen NMA definovaná metadata poskytovat; to ale nijak neomezuje, jaká „interní“/vlastní metadata takový repositář používá (metadatový model NMA nijak neomezuje metadatové modely vlastních repositářů, jen předepisuje, jaká metadata musí každý repositář být schopen poskytnout do NMA). Projekt CARDS ve spolupráci s pracovní skupinou Metadata již nyní poskytuje metodickou podporu pro definici metadatových modelů konkrétních tematických repositářů, odpovědnost za zvolený metadatový model je však na správci příslušného repositáře. Všechny pilotní repositáře budou mít své vlastní metadatové modely a budou sloužit i v tomto aspektu jako příklady pro ostatní, kteří budou chtít zřídit a provozovat další (tematické) repositáře.

Péče o metadata bude dále podpořena nástrojem pro správu metadatových profilů. Do konce roku 2024 vznikne první verze tohoto nástroje, který bude provozován v pilotním režimu tak, aby do Q3/2025 mohla vzniknout doporučená metodika pro správu metadatových modelů, která bude do konce roku 2025 podporovat i hierarchii těchto profilů. Technicky toto bude implementováno v rámci zmíněného nástroje správy profilů. Podobně jako v případě ostatních služeb, i zde je prostor formou spolupráce při přípravě projektu OS II zajistit dopřesnění toho, co

první verze nástroje a návazná metodika budou přesně obsahovat. Jedná se ale pouze o obecné metodiky a podpůrné nástroje, odpovědnost za konkrétní metadatová schémata/profilu je na správcích konkrétních repozitářů.

(Adresuje bod Metadatová schémata – velmi významně.)

- **FAIRifikace**

Okruh FAIRifikace sdružuje vypracování a zveřejnění návrhu postupů k naplňování FAIR principů v repozitářích s nástrojovým zázemím, které bude součástí NRP. Do konce Q1/2025 bude k dispozici Fair Implementation Profile Wizard, který bude prvním krokem k FAIRifikační nástavbě pro ověřování naplňování FAIR principů zabudované automaticky v rámci NRP. Podobně jako ostatní služby a nástroje, úplnější řešení bude k dispozici do konce roku 2025 a aktuálně probíhá dopřeshňování požadavků na očekávání od tohoto nástroje.

- **Podpora péče o Data Management Plány a napojení na repozitářovou platformu**

Tento okruh zahrnuje postupnou integraci nástroje Data Stewardship Wizard (DSW) do prostředí NRP/NDI. Tím vznikne komplexní zázemí, které nejen usnadní tvorbu DMP, ale zejména propojí vytvořené plány s repozitářovou platformou.

Nejdříve dojde k pilotní instalaci tohoto nástroje (v polovině roku 2025). Dále budou integrovány šablony (v polovině roku 2026) a spuštěn nástroj pro vizualizaci procesu správy dat v NRP (konec roku 2026). Bude podpořena i přidělování persistentních identifikátorů samotným plánům správy dat pro jejich snazší identifikaci (v polovině roku 2027). Verze Data Stewardship Wizardu bude integrována do NRP tak, aby bylo možné zavést tzv. actionable DMP, tedy DMP ve strojově čitelné podobě, se kterou bude umět NRP dále pracovat (nastavovat/kontrolovat podmínky a omezení, nastavovat podmínky retence dat apod.). První verze bude k dispozici v polovině roku 2025, ta ale ještě nebude plně strojově proveditelná.

Řešení je univerzální a je postaveno tak, aby nepředjímalo konkrétní oborová schémata šablon, ale spíše aby bylo schopno pracovat s jakýmkoliv šablonami a požadavky, které z tematických skupin (až po správce konkrétních repozitářů) vzejdou.

(Adresuje bod Data management plan – oborová schémata šablon.)

- **Licence**

V rámci tohoto okruhu bude podpořena práce s depozičními licencemi – půjde o nástroj, který nejen umožní správcům repozitářů definovat podmínky přístupů k repozitářům (jak při ukládání, tak při čtení/zpracování dat), ale zejména bude postupně licence převádět do actionable formy, analogicky DSW. To umožní přinejmenším semiautomatické dodržování licencí ekosystémem služeb NRP/NDI (např. licence vyžadující explicitní potvrzení konkrétních licenčních podmínek bude schopna ve spolupráci s AAI zajistit, že uživatel bude při prvním přístupu automaticky přesměrován na formulář s licenčními pravidly a požadavkem potvrzení ze strany uživatele).

V první fázi do poloviny roku 2025 bude k dispozici první verze správy licencí, která by měla zahrnovat kolekci známých licencí a průvodce, který usnadní výběr vhodné licence pro repozitáře i konkrétní datasety. Současně ekosystém NRP/NDI zajistí, že tyto přijaté licence se stanou integrální součástí metadat dataset, uložených v NDI.

(Adresuje bod Licencování dat.)

- **Automatizace sběru dat a metadat z přístrojů**

V tematickém okruhu pro automatizaci sběru dat a metadat z přístrojů bude v první fázi – v Q3/2024 – navržena architektura řešení pro automatizaci sběru dat a metadat z přístrojů do NRP. V dalším etapě, od Q3/2025, dojde k postupné integraci této automatizace do NRP. Nejdříve se bude jednat o automatizaci s jedním ovládacím prvkem (pol. 2026), další fáze bude přes

system pro kontinuální sběr a zpracování dat z webových crawlerů a nástroje pro instalaci dalších komponent až po fázi automatizace s více ovládacími prvky pro automatizaci sběru dat a metadat přímo z vědeckých přístrojů, která bude ve výsledku probíhat napřímo cestou: vědecký přístroj – NRP. Tento okruh lze vnímat jako zázemí, na které mohou přímo navázat další aktivity spojené se sběrem a následnou správou dat, například elektronické laboratorní deníky a jejich implementace v NRP nebo aktivity týkající se data provenance (zde se očekává postupný náběh těchto fází až do roku 2028). Tyto oblasti (laboratorní deníky, provenance, ...) nejsou součástí aktuálně řešených projektů IPs EOSC-CZ, IPs CARDS či NRP, je pouze vytvářeno prostředí pro jejich případnou integraci.

(Adresuje bod Laboratorní deníky.)

- **Kyberbezpečnost a compliance**

V rámci tohoto okruhu budou vytvořeny pravidla ochrany NRP i podmínky připojení do NRP. Na jedné straně bude zajištěné průběžné sledování formou postupného sběru bezpečnostní telemetrie a kontinuálním vyhodnocováním kyberbezpečnostních událostí.

Na straně druhé budou zpracovány podmínky přístupu k NRP, zejména podmínky tvorby repozitářů, požadavky na správce repozitářů, obecné podmínky dodržení API a protokolů (např. mandatorní připojení k AAI). S tím souvisí i základní podmínky pro práci s citlivými daty v prostředí NRP. První verze architektury zajištění kyberbezpečnosti i podmínek compliance bude k dispozici ve druhém pololetí 2025; v oblasti compliance předpokládáme úzkou spolupráci s projektem OS II již ve fázi přípravy tak, aby byly podmínky použití NRP byly v souladu s očekáváním vědecké komunity, která bude realizovat projekt OS II. A současně abychom včas, ještě před odevzdáním projektu OS II identifikovali a vyřešili případné sporné otázky a požadavky, které v rámci NRP nebude možné (např. z právních důvodů) naplnit. Stejně tak očekáváme postupný vývoj těchto podmínek tak, jak se NRP dostane do produkčního stavu a poroste počet uživatelů se specifickými potřebami.

- **Service desk**

Tematický okruh pro service desk zahrnuje pravidla pro připojení služby NRP do provozu service desku a funkční nástroj pro zobrazení stavu služby uživateli. Podpůrné procesy a stejně tak organizace projektových týmů (jak přípravných, tak realizačních) není explicitní součástí cílů a výstupů jednotlivých projektů a je třeba, aby se tyto aktivity a týmy koordinovali na úrovni celé iniciativy EOSC CZ. Pravidla práce service desku budou v nulté verzi k dispozici začátkem roku 2025, plný service desk bude v první verzi provozu k dispozici Q3/2025. Architektura Service desku také explicitně podchytí odpovědnost za L1 až L3 support v rámci NRP/NDI.

(Adresuje bod helpdesk L1-L3.)

- **Mapování vědeckých komunit**

Je plánováno komplexní mapování vědeckých komunit. V první fázi – předpokládáme zahájení do konce roku 2024 -- proběhne plošné (kvantitativní) mapování vědecké komunity v ČR za účelem poznání vědecké komunity, zjišťování jejich potřeb a měření dopadu iniciativy EOSC CZ. Ve vazbě na průběh a zejména výsledky tohoto mapování počítáme s podobným rozsáhlým mapováním v roce 2027 (aby výstupy byly k dispozici v roce 2028) a případně i menším „průběžným“ mapováním koncem roku 2025 či začátkem 2026. Podrobnosti a rozsah dalších mapování bude upraven v rámci realizace projektu NRP ve vazbě na výsledky a průběh prvního mapování. Výsledek prvního kola bude zmapování současného stavu (před implementací NRP) a další sběr potřeb a požadavků vědeckých komunit (v dosud nerealizovaném rozsahu).

(Adresuje bod Šetření mezi uživateli, uživatelskými komunitami (analýzy a definice potřeb.)

- **Správa a vyhledávání dat**

Okruh správy a vyhledávání dat slučuje technické zázemí pro správu perzistentních identifikátorů, které bylo vytvořeno na počátku roku Q1/2024 s přehledem všech repozitářů začleněných do NDI reprezentovaným vznikem Národního katalogu repozitářů (Q2/2025) a s pořízením a pilotním provozem Platformy nové generace postupně fungující od roku 2025 dále S přesahem do okruhu Iniciativa EOSC CZ sem lze zařadit i zřízení centra pro podporu perzistentních identifikátorů při NTK (www.identifikatory.cz) v roce 2023.

- **Vývoj služeb NRP**

Jak píšeme výše, aktuální prostředí implementace NDI (EOSC CZ) je vysoce flexibilní a počítá s tím, že v následujících letech bude třeba připravit nové služby, některé ze současných se stanou nepodstatnými apod. Z tohoto důvodu je jedním z klíčových výstupů Roadmapy služeb NRP. Z nulté verze této roadmapy, která byla připravena v rámci práce Pracovní skupiny EOSC CZ, vznikla množina služeb, které jsou aktuálně podpořeny implementací projektu NRP. Roadmapa bude každoročně aktualizována, nejbližší aktualizace je plánována na Q1/2025. Roadmapa již nyní zahrnuje služby poskytované všemi projekty, souvisejícími s implementací EOSC v ČR a očekáváme, že do ní budou zintegrovány i služby, které budou podpořeny projektem OS II.

Závěr

V dokumentu jsou popsány hlavní oblasti, v nichž jsou plánovány konkrétní výstupy projektů, které tvoří zázemí implementace EOSC v ČR. Popis je záměrně vysokoúrovňový, detaily jednotlivých cílů a výstupů jsou v dispozici v příloze.

Jak je z přehledu zřejmé, komplexní první vlna technických výstupů projektů vyvrcholí ve druhé polovině roku 2025, kdy bude k dispozici prostředí NRP s kapacitou cca 18 PB uživatelských dat, třemi podpořenými repozitářovými systémy, čtyřmi pilotními a jedním catch-all repozitářem. Současně budou v provozu v první verzi všechny služby; přesné vlastnosti této první verze služeb jsou součástí realizace projektů (zejména NRP), podrobnější specifikaci lze očekávat kolem přelomu let 2024 a 2025.

Do poloviny roku 2025 bude rovněž k dispozici architektura a první verze implementace kyberbezpečnostního zajištění celého prostředí a rovněž compliance dokumenty, definující podmínky využití NRP a jejích služeb.

Současně bude k dispozici dokumentace, začnou vznikat školicí materiály a budou k dispozici metodici a další členové širší podpory efektivního využití celého prostředí NDI.

Vzhledem k dynamickému vývoji implementace EOSC v EU (např. relativně nová koncepce EOSC Federace a EOSC uzlů) jsou i národní projekty IPs EOSC-CZ, IPs CARDS a NRP připraveny reagovat na měnící se vnější podmínky a současně i na měnící se požadavky a potřeby uživatelů (lze oprávněně očekávat, že po zprovoznění NRP v roce 2025 a následném zahájení realizace projektu OS II v roce 2025 získá vědecká komunita v ČR lepší představu potenciálu NRP a přijde s dalšími požadavky. Ty pak najdou cestu do aktualizace Roadmapy služeb a prostředí NDI a mohou vést k úpravám současných plánů realizace EOSC v ČR; vždy s cílem maximalizovat efektivitu vynakládaných prostředků a zajistit podporu rozvoje české vědy.

Z těchto důvodů plánujeme pracovat s tímto dokumentem jako s „živým materiálem“, který bude průběžně aktualizován a zpřesňován jak podle aktuálního stavu realizace již existujících projektů, tak i ve vazbě na vyvíjející se očekávání pozitivního dopadu implementace iniciativy EOOSC nejen v ČR.

V Brně 4.9.2024

Michaela Capandová, Jiří Marek,
Matej Antol, Luděk Matyska