

Open Science II

Zpráva o realizovaných aktivitách za rok 2025



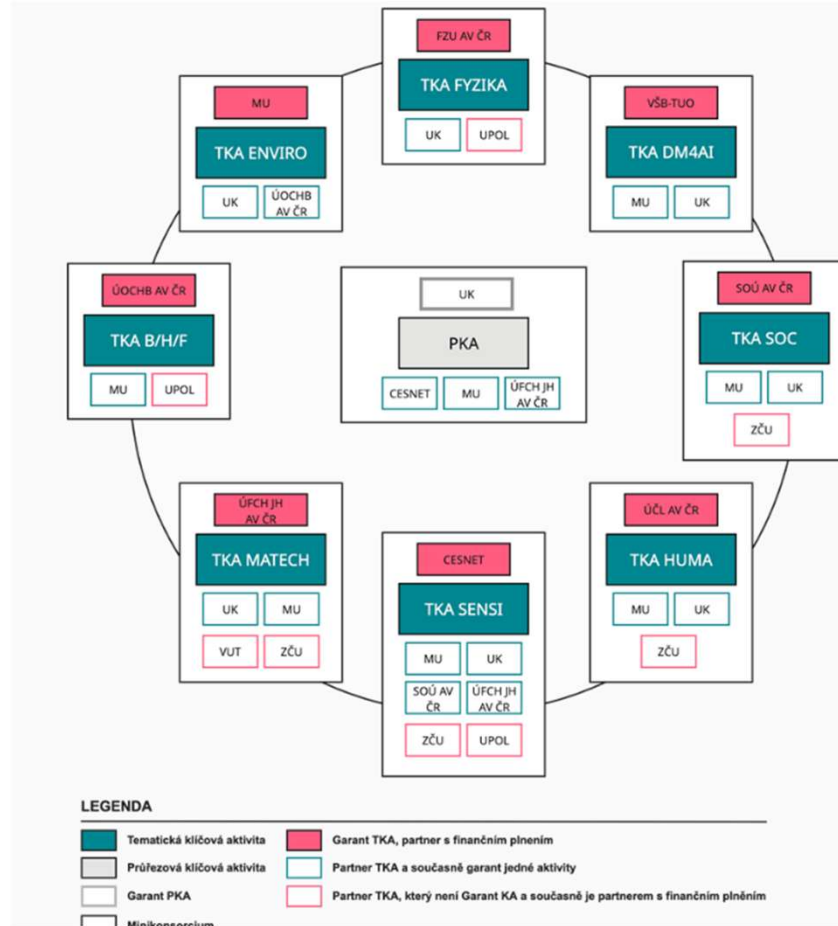
Základní parametry projektu

- Koordinace: Univerzita Karlova
- Partneři: 11 (AV ČR, VŠ, z.s.p.o.)
- Realizace: 1. 10. 2025 – 31. 12. 2028
- Celkový schválený rozpočet: 857 195 386,51 Kč
 - Investice: 16 874 818,15 Kč
 - Osobní náklady: 532 012 767,61 Kč
 - Cestovné: 5 547 837,89 Kč
 - Služby: 5 902 742,66 Kč
 - Minizáměry: 171 439 000,00 Kč
 - Ostatní neinvestice: 125 418 220,20 Kč

Partneři projektu



Partneři projektu



KA1 Řízení projektu

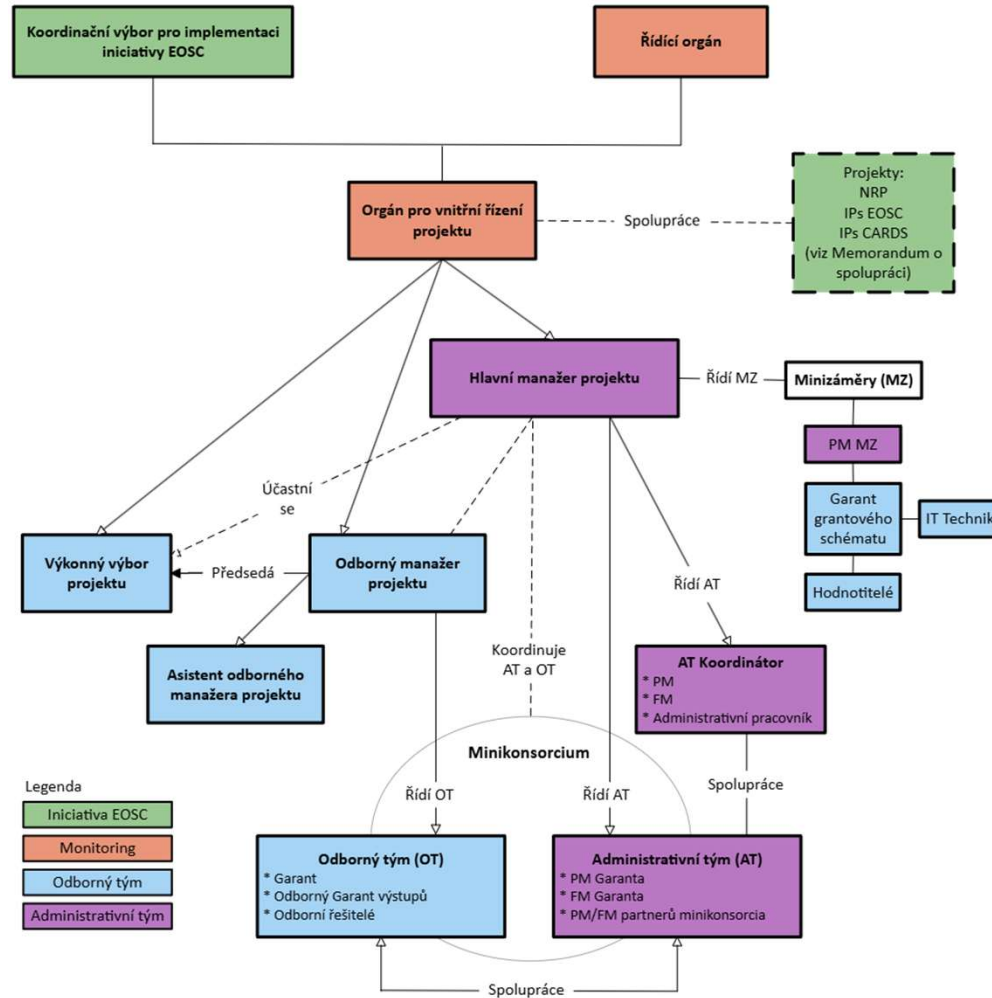


Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Řízení projektu



Přehled způsobilých výdajů/nákladů partnerů s finančním příspěvkem

Instituce	Rozpočet CELKEM	INVESTICE	NEINVESTICE
Univerzita Karlova	294 995 483,27 Kč	8 187 301,65 Kč	286 808 181,62
CESNET, zájmové sdružení právnických osob	60 231 596,60 Kč	2 500 000,00 Kč	57 731 596,60
Fyzikální ústav AV ČR, v. v. i.	62 924 239,76 Kč	0,00 Kč	62 924 239,76
Masarykova univerzita	161 597 983,21 Kč	5 323 576,50 Kč	156 274 406,71
Sociologický ústav AV ČR, v. v. i.	30 607 459,07 Kč	0,00 Kč	30 607 459,07
Univerzita Palackého v Olomouci	42 689 752,20 Kč	0,00 Kč	42 689 752,20
Ústav fyzikální chemie J. Heyrovského AV ČR, v. v. i.	39 699 342,18 Kč	863 940,00 Kč	38 835 402,18
Ústav organické chemie a biochemie AV ČR, v. v. i.	24 522 501,74 Kč	0,00 Kč	24 522 501,74
Ústav pro českou literaturu AV ČR, v. v. i.	24 338 146,01 Kč	0,00 Kč	24 338 146,01
Vysoká škola báňská - Technická univerzita Ostrava	39 434 695,77 Kč	0,00 Kč	39 434 695,77
Vysoké učení technické v Brně	22 926 057,67 Kč	0,00 Kč	22 926 057,67
Západočeská univerzita v Plzni	53 228 129,03 Kč	0,00 Kč	53 228 129,03
CELKEM	857 195 386,51 Kč	16 874 818,15 Kč	840 320 568,36 Kč

Interní činnost

Typ komunikace	Zodpovědná osoba	Výstup/Účel	Frekvence
Schůzky admin tým KA	HMP	Administrativní požadavky na realizaci projektu	Týdně
Schůzky OT	OMP	Stav odborné stránky projektu	Měsíčně
Jednání VVP	OMP	Setkání všech partnerů projektu	Měsíčně
Jednání OVŘP	Pověřený zástupce UK	Pravidelné zasedání zástupců OVŘP	1x za 6 měsíců
Podklady pro odevzdání ZoR/ŽoP	AT + OT	Pravidelný reporting	1x za 6 měsíců
Nepodstatné ŽoZ	AT	Změnová řízení	1x za 6 měsíců
Podstatné ŽoZ	AT + OT	Změnová řízení	Dle potřeby
Setkání koordinátora s garanty KA	HMP	Sdílení zkušeností	1x za 6 měsíců

- Hlavní manažer projektu (HMP) – Ing. Michaela Hynková
- Odborný manažer projektu (OMP) - Mgr. Jan Tuček, Ph.D.
- Pověřený zástupce UK pro Orgán pro vnitřní řízení projektu (OVŘP) - prof. PhDr. Ladislav Křišťoufek, Ph.D.
- Administrativní tým (AT)
- Odborný tým (OT)
- Výkonný výbor projektu (VVP)

Klíčové aktivity

KA2 Tematického clusteru Bio/zdraví/potraviny

KA 3 Tematického clusteru Materiálové vědy a technologie

KA 4 Tematického clusteru Data management pro umělou inteligenci a strojové učení

KA 5 Tematického clusteru Sociální vědy

KA 6 Tematického clusteru Fyzikální vědy

KA 7 Tematického clusteru Humanitní vědy a umění

KA 8 Tematického clusteru Environmentální vědy

KA 9 Tematického clusteru Citlivá data

KA 10 Přířezová témata projektu

KA 11 Minizáměry

Základní cíle roku 2025

- Realizace pilotního kola minizáměrů v rámci přípravy projektu
- Odevzdání projektu do 06/2025
- Zahájení realizace projektu 1. 10. 2025
- Podpis partnerské smlouvy
- Postupné vybudování a propojení administrativního a realizačního týmu
- Zahájení aktivit dle harmonogramu
- Prohlubování spolupráce s Iniciativou EOSC včetně synergických projektů

KA2 Tematického clusteru Bio/zdraví/potraviny

Garant KA: prof. RNDr. Jiří Vondrášek, CSc.



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

- **Národní omický repozitář (OmiCZ):** provedena dekompozice architektury na jednotlivé funkční komponenty, probíhá prvotní průzkum možných technických implementací těchto komponent. Dokončen konceptuální model metadatového modelu, probíhají práce na detailní specifikaci jednotlivých metadatových entit a jejich vztahů. Harmonizace technických přístupů a interoperability s existující národní infrastrukturou je zajišťována formou konzultací se zástupci NRP.
- **Propojení repozitáře OmiCZ na FEGA, Beacon a ClinData:** práce na finalizaci implementace českého FEGA uzlu, včetně identifikace technických a procesních požadavků pro napojení repozitáře OmiCZ. Probíhá analýza možností integrace s mezinárodní FEGA sítí s využitím zkušeností získaných v rámci spolupráce s ELIXIR Switzerland.

Realizace hlavních aktivit v roce 2025

- **Repozitář strukturních simulačních dat (BioSimCZ):** dokumentace návrhu architektury prototypu repozitáře (včetně návrhu architektury nástroje pro vyčítání metadat a metadatového schématu).
- **Repozitář pro lidská i animální obrazová a fyziologická multimodální data (Imaging repozitář):** probíhá implementace jádra repozitáře, veškeré technické aspekty jsou průběžně konzultovány s experty na platformu Invenio z týmu NRP. Proběhla konzultace zaměřená na budoucí napojení repozitáře na evropskou infrastrukturu EBRAINS.
- **Metadatový model:** připraven návrh kontrolního seznamu pro FAIR regulovaná klinická výzkumná data. Tento návrh bude sloužit jako jeden z řídicích dokumentů pro vývoj FAIRifikované metodologie a metadatového modelu.

KA 3 Tematického clusteru Materiálové vědy a technologie

Garant KA: Mgr. Marek Cebecauer, Ph.D.



Spolufinancováno
Evropskou unií



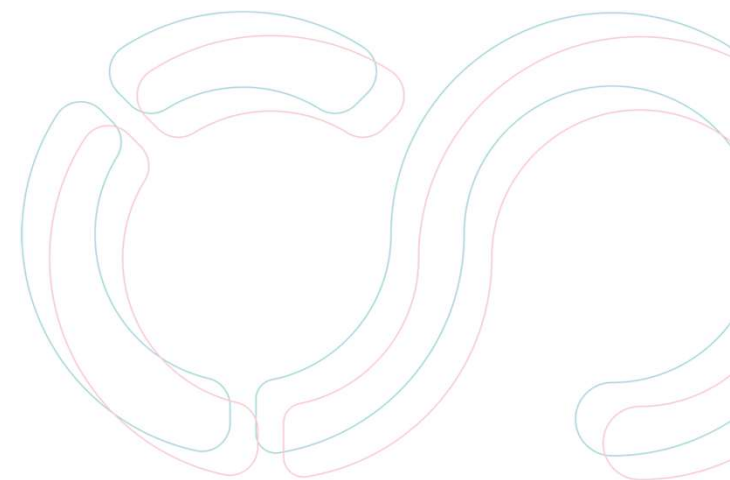
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

- **Repozitář DANTEc:** příprava základního metadatového modelu DANTEcore na bázi centrálního modelu CCMM (NRP) pro kolekci UNITECH v repozitáři DANTEc a jeho harmonizace s clustery FYZIKA a B/H/F; příprava submission workflow pro repozitář DANTEc a v návaznosti na to první návrh schvalovacích a kurátorských procesů; připraven první návrh základního dokumentu (Policies and Guidelines) pro repozitář DANTEc a první návrh depoziční smlouvy mezi správcem repozitáře DANTEc a vkladateli. V návaznosti na průběh vývoje softwarového řešení Invenio-CESNET připraven návrh revize technického řešení repozitáře DANTEc.

Realizace hlavních aktivit v roce 2025

- **Integrace ELN Kadi4Mat s repositářem DANTEc:** návrh softwarového řešení požadavků správců výzkumných infrastruktur na správu uživatelů, přístrojů a produkovaných dat a analýza prostředí Invenio-CESNET pro přímé vkládání záznamů z elektronického laboratorního deníku ELN Kadi4Mat.



KA 4 Tematického clusteru Data management pro umělou inteligenci a strojové učení

Garant KA: Ing. Jan Martinovič, Ph.D.



Spolufinancováno
Evropskou unií



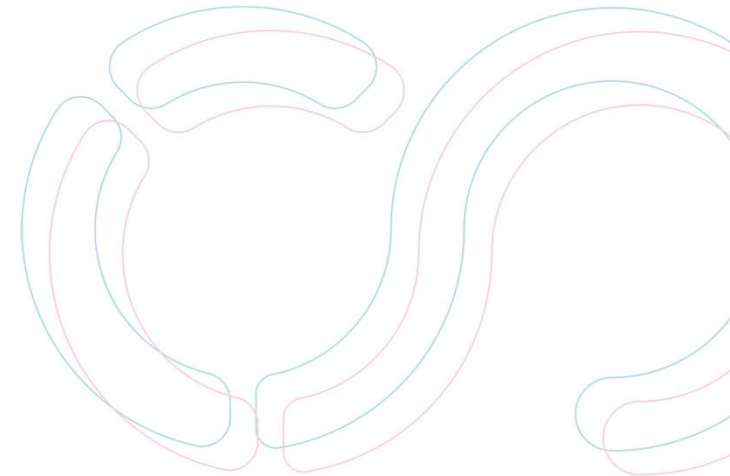
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

- **Inicializace tvorby repozitáře DM4AI** prostřednictvím návrhu základních vlastností a účelu navrhovaného repozitáře (metadatový model, přístupová a datová politika, služby apod.).
- **Zahájení oficiálního procesu konzultací** k založení repozitáře DM4AI (DSpace) prostřednictvím formální žádosti podané v listopadu 2025 (vyplnění vstupního dokumentu: Účel a základní vlastnosti navrhovaného repozitáře).
- **Technický meeting k systému LEXIS** a možnostem jeho integrace, se zaměřením na technická, datová a procesní propojení s připravovaným repozitářem.

Realizace hlavních aktivit v roce 2025

- **Inicializace tvorby datové politiky** pro práci s citlivými údaji – série odborných meetingů s pracovní skupinou SENSI (Citlivá data).
- **Realizace VZ Prostředí pro omezenou interakci s AI modely v repozitáři.** Jedná se o dodávku tří systémů pro benchmarkování a inferenci velkých jazykových modelů na alternativních platformách bez klasického GPU typu AMD MI210 nebo Nvidia H100.



KA 5 Tematického clusteru Sociální vědy

Garant KA: Mgr. Ilona Trtíková, Ph.D.



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

- **Inovace repozitáře ČSDA a FAIRifikace dat** probíhala zejména systematizace a revize dat a metadat v repozitáři ČSDA, včetně čištění metadat a kontroly souladu s FAIR principy. Byly zahájeny přípravné práce na depozici dat, probíhaly konzultace k metodám ukládání dat a metadatovým standardům.
- **Napojení repozitáře ČSDA do NRP a jeho rozvoj a provoz:** probíhají analytické a koncepční práce zaměřené na napojení repozitáře ČSDA na NRP
- **Platforma Atlasobyvatelstva.cz pro specializované mapy:** práce na návrhu platformy pro zpřístupnění a vizualizaci populačních a prostorových dat, včetně technického řešení a vazby na další datové služby projektu. Bylo založeno S3 úložiště na CESNETu, kde budou uloženy mapy a data.

Realizace hlavních aktivit v roce 2025

- **Analýza možností postupu kódování citlivého textu a jeho následného zveřejnění bez citlivých údajů:** analytické práce zaměřené na identifikaci metod kódování a úprav citlivých textových dat tak, aby bylo možné jejich bezpečné zpřístupnění bez ohrožení ochrany citlivých údajů.
- **Inovovaný repozitář DataHub:** analytické a návrhové práce zaměřené na inovaci repozitáře DataHub, s důrazem na jeho technickou architekturu, optimalizaci metadatových modelů, implementaci a správu persistentních identifikátorů a zajištění interoperability s dalšími komponentami NRP.
- **Metodika/standard práce s daty z experimentů z oblasti ekonomie:** mapování specifik experimentálních ekonomických dat, jejich struktury a požadavků na ukládání, dokumentaci a zpřístupňování.

KA 6 Tematického clusteru Fyzikální vědy

Garant KA: RNDr. Antonín Fejfar, CSc.



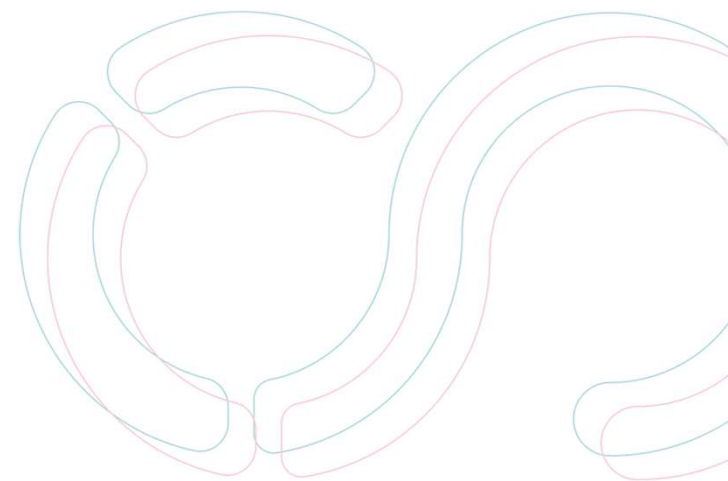
Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

- **Tvorba nového repozitáře Fyzika a fairifikace dat:** příprava architektury repozitáře, průběžné konzultace s metodiky repozitářových systémů, testování přenosů velkoobjemových dat (řádu 10 TBy) z CERN do prototypu repozitáře národní datové infrastruktury a na sladění metadat s CCMM.



KA 7 Tematického clusteru Humanitní vědy a umění

Garant KA: Mgr. Vojtěch Malínek, Ph.D.



Spolufinancováno
Evropskou unií



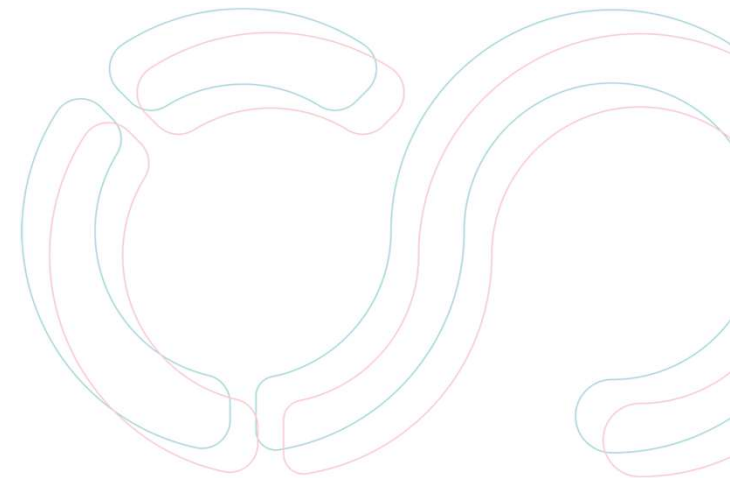
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

- **Repozitář Digitalia MUNI ARTS:** zahájeny a metodicky ukotveny práce na integraci infrastruktury Digitalia MUNI ARTS do NRP prostřednictvím standardizovaných AIP balíčků (Archival Information Packages). Na základě analýzy technického stavu repozitáře Islandora a jeho datových struktur bylo rozhodnuto o dvoufázovém řešení, které umožňuje v první fázi dosáhnout plné archivovatelnosti bez ztráty datové integrity a současně vytvořit základ pro pozdější sémantickou interoperabilitu. Byl navržen mechanismus uložení komplexní strukturované JSON reprezentace jako mediálního objektu do repozitářového jádra Fedora systému Islandora, k němu je připojen Dublin Core záznam (DC XML) prostřednictvím rozhraní OAI-PMH, který zajišťuje základní interoperabilitu a automatizované předávání metadat.

Realizace hlavních aktivit v roce 2025

- **Inovace repozitáře ArchaeoVault:** ustavení role digitálního kurátora infrastruktury a jeho seznámení s technickým a datovým řešením pilotního repozitáře; revize dosud zachycených požadavků archeologické komunity a jejich vztahu k existujícím metadatovým profilům a datovým strukturám repozitáře.



KA 8 Tematického clusteru Environmentální vědy

Garant KA: prof. RNDr. Jana Klánová, Ph.D.



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

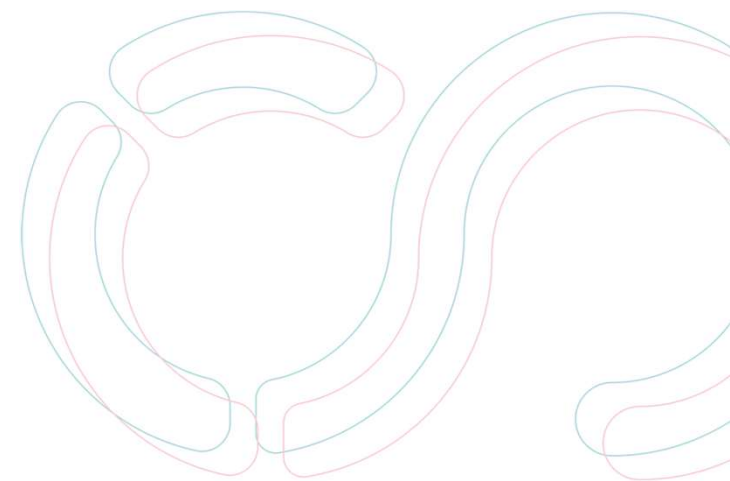
- **Inovace a rozšíření repozitáře GENASIS:** vyhodnocení repozitáře z pohledu FAIR principů; analýza požadavků a doporučení pro repozitáře v prostředí NRP, včetně analýzy řešení pro přiřazování perzistentních identifikátorů k datovým objektům či extraktům; analýza současné struktury metadat a procesů používaných při ukládání dat, a to i v kontextu standardů vytvořených v partnerství HEU PARC (Partnership for the Assessment of Risks from Chemicals). Na základě této analýzy byl připraven koncept aktualizovaného metadatového modelu, který umožní jak ukládání dat z nových matic i emergentních polutantů, tak i jednodušší a interoperabilní popis dat; harmonizace metadatového modelu a řízených slovníků pro modul chemických látek ve vnějším prostředí v GENASIS.

Realizace hlavních aktivit v roce 2025

- **Repozitář pro ukládání dat z necílených hmotnostně spektroskopických analýz pro účely hodnocení lidského expozomu:** analýza a specifikace potřeb a datového prostředí výzkumné komunity pro data z necílených hmotnostně spektrometrických analýz pro účely hodnocení lidského expozomu; analýza pracovního postupu (workflow) pro zpracování a dlouhodobé a bezpečné ukládání FAIR metabolomických (meta)dat z necílených analýz, včetně nároků na úložné a výpočetní kapacity; analýza a tvorba návrhu technického řešení repozitáře; analýza existujících metadatových modelů a analýza potřeb komunity pro jejich rozšíření s důrazem na možnost jejich znovuvyužití a interoperabilitu.

Realizace hlavních aktivit v roce 2025

- **Repozitář pro ukládání toxikologických a ekotoxikologických dat:** analýza (ve spolupráci s evropským Partnerstvím pro analýzu chemických rizik) potřeb a datového prostředí mezinárodní toxikologické komunity, na základě které bude vytvořen návrh struktury metadatového modelu a repozitáře pro ukládání toxikologických a ekotoxikologických dat.



Realizace hlavních aktivit v roce 2025

- **Nové úložiště umožňující propojení geokódovaných dat z různých domén:** analýza potřeb výzkumné komunity v oblasti propojování geokódovaných dat z různých oblastí a dostupných řešení pro přípravu nového repozitáře a jeho metadatového modelu; analýza relevantních a dostupných mapových a dalších prostorových datových zdrojů s potenciálem dlouhodobého využití pro hodnocení environmentálních expozic; zahájení analýzy dostupnosti a využitelnosti vybraných národních a regionálních datových zdrojů, zejména dat ze Sčítání lidí domů a bytů (SLDB) z Českého statistického ústavu (ČSÚ) s potenciálem jejich propojení s environmentálními vrstvami.

Realizace hlavních aktivit v roce 2025

- **Repozitář pro referenční obrazová data flory a rostlinných společenstev, Repozitář pro genetický biomonitoring a genetická data, Repozitář pro zoologické sbírky:** zahájeny práce na FAIRifikaci a konsolidaci dat; personální zajištění IT pozice správce repositářů a nasazení AI nástrojů; mapování potřeb v oblasti metadatových modelů pro jednotlivé typy vědeckých dat.
- **Vývoj a pilotní zavedení služeb pro analýzu enviromentálních a biodiverzitních dat:** rozšířena funkcionality stávajících pipeline o práci s původními daty ve formátu RAW pro přístroje Thermo Scientific (formát snižuje náklady na úložný prostor a umožňuje bezztrátovou archivaci); testování pokročilých algoritmů pro redukci šumu a zlepšení identifikace píků.

KA 9 Tematického clusteru Citlivá data

Garant KA: Mgr. Zdenka Dudová, Ph.D.



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

Práce byly zahájeny v následujících oblastech:

- Analýza správné praxe pro sdílení citlivých dat z kolaborativního výzkumu
- Metodika FAIRifikace citlivých dat v NDI
- Softwarový modul pro kontrolu a zajištění kvality dat repositářů, vč. metodické příručky QA/QC dat v repositářích
- Služba řízeného zabezpečeného procesu příjmu/výdeje dat do/z prostředí pro zpracování citlivých dat
- Rozhraní pro koncové uživatele pro interaktivní/dávkové zpracování citlivých dat v zabezpečeném prostředí
- Služba pro autorizaci dávkového zpracování citlivých dat

KA 10 Průřezová témata projektu

Garant KA: Mgr. Kateřina Zvoníková



Spolufinancováno
Evropskou unií

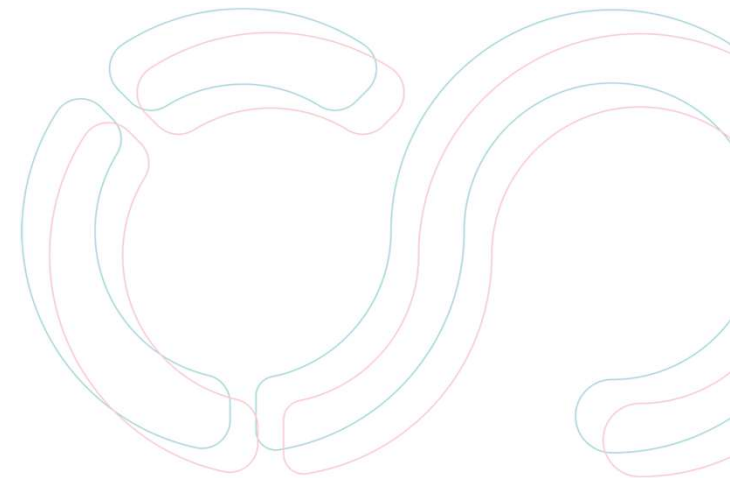


MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

Velké jazykové modely/LLM

- Zahájeny práce na sběru požadavků a návrhu architektury pro využití velkých jazykových modelů.
- Příprava podkladů pro realizaci VZ na pořízení HW prostředků pro LLM.



Realizace hlavních aktivit v roce 2025

Kyberbezpečnost

- Zahájení sběru a analýzy požadavků.
- Byli kontaktováni vývojáři NRP, správci datových úložišť a zástupci bezpečnostního týmu. Proběhly první schůze, na kterých byl zjišťován současný a plánovaný stav NRP infrastruktury, požadavky na bezpečnost a monitoring infrastruktury a jejích služeb, diskutována byla možná řešení.

Realizace hlavních aktivit v roce 2025

Compliance

- Zahájena byla analýza právních předpisů a problematiky autorství a licenčních podmínek s cílem identifikovat požadavky na plánované výsledky, způsob práce s dostupnými nástroji a knihovnamí umělé inteligence a možnosti využití existujícího softwaru pro budování AI asistenta i pro účely zpracování dat pro zajištění kyberbezpečnosti.
- Zahájena byla podpora v podobě právní analýzy a odborných právních konzultací jako podpora řešitelských týmům podaktivit Kyberbezpečnost a Velké jazykové modely.

Realizace hlavních aktivit v roce 2025

Elektronické laboratorní deníky pro NRP

Proběhla první fáze analýzy řešení PKA ELN:

- Mapování dostupných možností nasazení podporovaných ELN na Kubernetes
- Byl zvolen systém na týmovou tvorbu a publikaci metodiky nasazení národních instancí ELN
- Proběhla analytická fáze implementace publikace záznamů z ELN do repozitářů NDI/NRP



KA 11 Minizáměry

PM: Ing. Monika Keřková, MPA



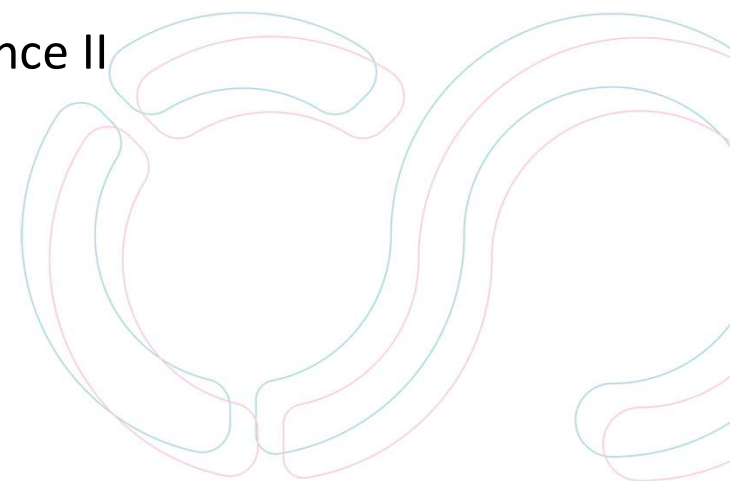
Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Realizace hlavních aktivit v roce 2025

- Realizace pilotního kola minizáměrů v rámci přípravy projektu
- Vyhodnocení procesu podávání Žádostí o projekt a procesu hodnocení
- Identifikace problematických bodů v metodikách a procesech
- Příprava a podání ŽoZ
- Příprava 1. a 2. Výzvy ke spolupráci - Minizáměry Open Science II



Děkujeme za pozornost!

E: eosc2@ruk.cuni.cz

W: <https://www.eosc.cz/projekty/open-science-ii>



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY