



Příloha č. 1.1 Smlouvy - Datové sady

Portál EDP, včetně datového skladu a vizualizace, bude tvořit zásadní složku **systému monitoringu Národní RIS3 strategie včetně její krajské dimenze**. Je připravován s ohledem na jeho hlavní účel, a to je **měření a hodnocení výkonnosti RIS3 strategie v ČR v období 2021 – 2027**.

Bližší informace k nastavení monitoringu Národní RIS3: Kap. 4.1, 4.2 a 5.4 dokumentu https://www.mpo.cz/assets/cz/podnikani/ris3-strategie/dokumenty/2021/1/A_RIS3-Strategie.pdf

Bližší informace k monitorovacím indikátorům a financování RIS3 strategií: https://www.mpo.cz/assets/cz/podnikani/ris3-strategie/dokumenty/2021/1/D_Priloha-3-Indikatory-a-Financovani_1.pdf

Smyslem zakázky je posílení informačních a datových zdrojů pro zefektivnění analytických prací a zkvalitnění entrepreneurial discovery process, a to prostřednictvím rozvoje dílčích datových systémů, agregace dezintegrovaných statistických systémů dat do webového portálu, využití foresightu a dalších informačních, datových a analytických zdrojů.

V rámci zakázky nedojde k rozvoji již existujících stávajících datových zdrojů, ale pouze k jejich využití v rámci integrace a scelování datových sad.

1. Dlouhodobá vize datového skladu

Splnění účelu datového skladu Portálu předpokládá nastavení jeho vnitřní struktury včetně funkcionalit jako systému datových sad a informačních zdrojů, který bude umožňovat interpretaci vývoje naplňování cílů RIS3, konkrétně pak:

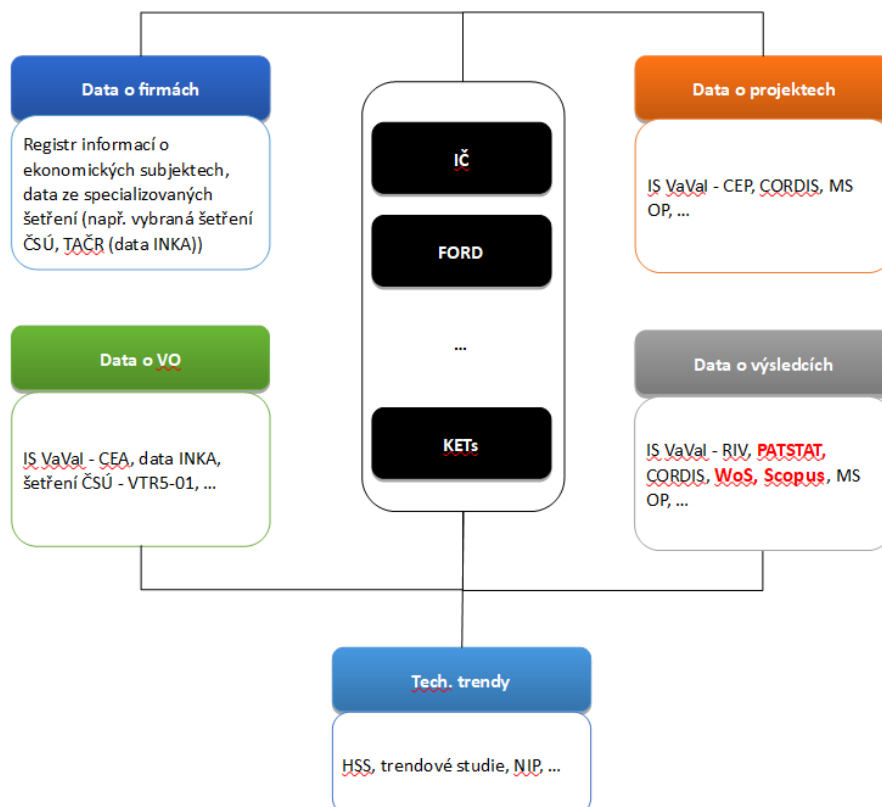
- dlouhodobé strategické vize
- horizontálních priorit – strategických a specifických cílů
- vývoje strategických směrů inteligentní specializace (domén specializace) v oblasti orientovaného a aplikovaného výzkumu a inovací,

a to jak u jednotlivých ukazatelů a cílů, tak v jejich vzájemných souvislostech (kde je relevantní). Budou vyjasňovat a znázorňovat účel, vývoj a fungování strategie a umožňovat její srozumitelnost a využitelnost pro odbornou i širší veřejnost.

V rámci zakázky půjde o vybudování základní struktury datového skladu s omezeným počtem indikátorů. Ve fázi rozvoje bude zadavatel vlastními zdroji postupně doplňovat další datové sady, poskytovatel zajistí jejich technické napojení. Výhledově (mimo tuto zakázku) je počítáno s napojením jednotlivých datových zdrojů do EDP portálu pomocí API rozhraní a integrací dat v potřebném formátu z vybraných datových zdrojů. Cílem je část těchto úkonů co nejvíce automatizovat, aby data byla ve správném formátu, který bude umožňovat třídění, integraci, odstranění duplicit atp. Část výstupů bude muset být integrována, očištěna a tříděna ručně za využití pracovníků zadavatele. Dojde i k vytvoření převodníků dat nekompatibilních formátů dat různých poskytovatelů.



Architektura datového skladu bude umožňovat po zanesení úvodní sady monitorovacích indikátorů další doplňování dat tím způsobem, že bude možné zakomponování parametru (parametrů), který/ktelé bude/budou umožňovat vzájemné propojení jednotlivých dat a vytváření individuálních datových souborů (např. ID projektu, CZ-NACE kód nebo IČ ekonomického subjektu). Zakomponování těchto "převodníků" (černé bloky) znázorňuje níže uvedené ilustrativní znázornění:



Legenda: IČ - identifikační číslo ekonomického subjektu, FORD (Field of Research and Development - číselné označení výzkumného směru), KETs – Key Enabling Technologies (klíčové slovo), ostatní zkratky jsou zkratky různých datových zdrojů v rámci jednotlivých kategorií

2. Požadavky na datový sklad a vizualizaci

2.1 Mezi hlavní požadavky na datový sklad patří:

- Ukládání dat, jejich třídění, vyhledávání a vzájemné propojování.
- Škálovatelnost dat.
- Vytěžování dat, tvorbu specializovaných datových sad podle zadaných kritérií (filtrů, dotazů).



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



- Nad datovým skladem vizualizaci klíčových/zastřešujících dat formou tematických dashboardů znázorňujících širší sadu indikátorů, obsahující grafy (sloupčové, koláčové, mapové kartogramy, bodové).
- Stažení dat z dashboardů do formátu .csv, .xlsx.
- Přípravu pro možnost vytvoření integračního API rozhraní, aby bylo možné datové zdroje automaticky stahovat od třetích stran.

2.2 Typy dat a informací, které budou ukládány do databáze/datového skladu

- Data z vybraných dostupných databází či specializovaných šetření (vkládané „ručně“ či přes API rozhraní) a jejich vizualizace.
- Datové vrstvy pro přípravu vizualizací (např. mapové vrstvy CzechInvestu a TAČR); data budou uložena mimo datový sklad Portálu EDP.
- Textová data – textové databáze - zdroj pro fulltextové vyhledávání.

3. Etapy tvorby datového skladu Portálu

3.1 První etapa: Vytvoření datového skladu a BI nástroje (3. etapa podle harmonogramu VZ)

- Příprava datového skladu v termínu do 30. 4. 2022.
- Nahrání indikátorů vize (viz Příloha 1.2 Datové sady - tabulka) - přípravu dat, jejich čištění a finální podobu zajišťuje zadavatel, poskytovatel zajistí technické napojení.
- Tvorba jednoho tematického dashboardu (viz Příloha 1.2 Datové sady - tabulka):
 - dashboard s indikátory vize Národní RIS3 strategie - poskytovatel zrealizuje dashboard, na kterém budou po dohodě se zadavatelem vizualizovány první 4 vybrané indikátory z Tabulky 1 této Přílohy.
- Poskytovatel technicky dimenzuje kapacitu datového skladu tak, že počítá s přípravou na doplnění dalších indikátorů (cca 200 ukazatelů v členění dle krajů, odvětví ekonomických činností v časových řadách posledních cca 25 let, tj. cca 200 MB dat) pro jejich zakomponování do nových dashboardů (s cílem možnosti vlastní administrace volby indikátorů pro zadavatele, přípravu dat v rámci fáze rozvoje zajistí zadavatel).

3.2 Druhá etapa: Provoz, správa a rozvoj Portálu EDP (5. etapa podle harmonogramu VZ)

- Rozvoj datového skladu v termínu do 30. 11. 2022.
- Nahrání dalších relevantních indikátorů, která jsou v době přípravy této zadávací dokumentace k EDP portálu ve fázi tvorby či testování (především data ke strategickým a specifickým cílům z Monitorovacího systému MS2021+, krajská data, data ze specializovaných šetření k výzkumným infrastrukturám a klastrům, data Horizon Scanning apod.) v rozsahu cca 200 ukazatelů v členění dle krajů, odvětví ekonomických činností v časových řadách posledních cca 25 let, tj. cca 200 MB dat.



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



- Přípravu dat, jejich čištění a finální podobu zajišťuje zadavatel, poskytovatel zajistí pouze jejich technické napojení a zaškolení vybraných zástupců zadavatele v oblasti importu dat.

Tabulka 1: Indikátory vize Národní RIS3 strategie

	<i>Zdroj</i>	<i>Územní členění</i>	<i>Měřicí jednotka</i>	<i>Frekvence vyhodnocení</i>	<i>Formát</i>	<i>Dostupnost zdarma/ano/ne</i>		<i>Napojení přes API / vkládání ručně</i>
The Global Competitiveness Index 4.0	World Economic Forum	celý svět	index	ročně	csv	ano	https://govdata360.worldbank.org/indicators/ha03bec65?country=BR&indicator=41472&viz=line_chart&years=2017,2019	ručně
Hrubá přidaná hodnota podle sektorů	ČSÚ, Databáze regionálních účtů	ČR	mld.Kč	ročně	xlsx	ano	https://apl.czso.cz/pli/rocenka/	ručně
Produktivita práce podle sektorů (HPH/zaměstnanec)	ČSÚ, Databáze regionálních účtů	ČR	mld.Kč	ročně	xlsx	ano	https://apl.czso.cz/pli/rocenka/	ručně
Vývoz a dovoz High-tech zboží a služeb	ČSÚ, Databáze regionálních účtů	ČR	mil. Kč	ročně	xlsx	ano	https://www.czso.cz/csu/czso/zahranicni_obchod_s_high_tech_zbozim_vav	ručně
The Global Innovation Index (GII)	World Intellectual Property Organization	celý svět	index	ročně	pdf, csv	ano	https://www.wipo.int/publications/en/details.jsp?id=4514&plang=EN	ručně
The Digital Economy and Society Index (DESI)	Evropská komise	EU	index	ročně	csv	ano	https://digital-agenda-data.eu/datasets/desi	API, ručně

The European innovation scoreboard (EIS)	Evropská komise	EU	index	ročně	xlsx	ano	https://ec.europa.eu/docsroom/documents/41864	API, ručně
Výdaje na výzkum a vývoj v ČR (GERD)	ČSÚ	ČR, kraje	mil. Kč	ročně	xlsx	ano	https://www.czso.cz/csu/czso/statistika_vyzkumu_a_vyvoje	ručně
Celkové výdaje na VaV (GERD) - podíl na HDP (intenzita výzkumu a vývoje)	ČSÚ	ČR, EU	podíl na HDP	ročně	xlsx	ano	https://www.czso.cz/csu/czso/statistika_vyzkumu_a_vyvoje	ručně