

## **TECHNICKÝ POPIS PŘEDMĚTU PLNĚNÍ**

### **Specifikace plnění veřejné zakázky**

Předmětem plnění veřejné zakázky je vytvořit, nasadit a pilotně naplnit funkční řešení Řídicího systému datových zdrojů a informačních výstupů rezortu životního prostředí (dále „**Řídicí systém**“). Jedná se o kombinaci dodávek metodik, softwarových nástrojů (SW), jejich parametrizace, naplnění vybranými vstupními/výstupními daty a navazujících servisních služeb k SW.

Vytvořený Řídicí systém bude základní součástí komplexního SW řešení a projektu MŽP STaR – Platforma pro statistiku, reporting a analýzy (zkr. „**STaR**“), jehož hlavním cílem je implementace informačního systému pro statistiku a reporting v rezortu MŽP.

STaR bude sloužit zejména ke čtyřem účelům:

- pro shromažďování dat o životním prostředí (dále jen „**ŽP**“) a vlivech na ŽP pro tvorbu informací určených veřejnosti;
- pro monitoring, hodnocení a podrobnou analýzu stavu a vývoje ŽP včetně modelování budoucího vývoje jednotlivých zájmových složek ŽP;
- k monitoringu naplňování závazků (vyplývajících z legislativy/strategických dokumentů) MŽP, zejména evidenci plnění reportingové povinnosti rezortu ŽP;
- k zajištění transparentnosti, věrohodnosti a veřejné ověřitelnosti datových výstupů, služeb, reportů a statistik resortu.

Motivací zadavatele v projektu STaR je:

- (1) nahradit stávající technicky a morálně zastaralé řešení statistického systému ISSaR<sup>1</sup> sloužící jako datová základna a publikační nástroj pro vybraná agregovaná data v oblasti statistiky ŽP;
- (2) zajistit dlouhodobě udržitelné technologické prostředí pro centrální zajištění statistických, reportingových a publikačních služeb nad operativními daty transakčních informačních systémů a
- (3) zajistit prostředí pro správu metadat informačních zdrojů rezortu ŽP včetně jejich zmapování.

Postavení předmětu a výstupů této veřejné zakázky v rámci projektu STAR znázorňuje schéma, jež je obsahem **Přílohy č. 5a** (Architektonické schéma projektu STaR) této Smlouvy. Vytvořeným Řídicím systémem, tj. nástroji a postupy pro katalogizaci metadat a ukládání vybraných dat v centrálním analytickém úložišti, bude vytvořen základ pro další využití a zpracování shromažďovaných dat a informací rezortu ŽP.

---

<sup>1</sup> Podrobnosti o ISSaR, viz **Příloha č. 6a** - Stručný popis stávajícího stavu ISSaR.

## Požadavky na Řídicí systém

Řídicí systém se bude skládat z níže uvedených dodávek a služeb:

### 1. Vytvoření a implementace metadatového katalogu

#### a. zpracování a dodání **Metodiky** katalogizace:

- výstupem bude návodný písemný dokument (metodika);
- srozumitelnost: pro odborníky mimo IT;
- požadavky na obsah - Metodika musí obsahovat minimálně následující:

#### i. definice předmětu katalogizace spočívající především v následujících aktivitách:

- identifikace potřeb, indikátorů, reportů aj. (dále „zájmové objekty“) a jejich rozčlenění do logických a hierarchických kategorií;
- identifikace a kategorizace datových vstupů;
- identifikace a kategorizace informačních výstupů, služeb (SW aplikací), potřeb;
- identifikace optimálního množství sledovaných atributů pro jednotlivé typové kategorie objektů.

Ukázka 7 kategorií zájmových objektů je specifikována v **Příloze č. 3a** (Orientační přehled zájmových objektů katalogizace) této Smlouvy. Výstupem bude tabulkový dokument podrobně a strukturovaně popisující všechny důležité atributy nutné pro katalogizaci zájmových objektů všech 7 úrovní a jejich typů. Rozsah sledovaných atributů se zpravidla bude dle úrovně (I-VII viz **Příloha č. 3a**) a konkrétního typu zájmových objektů lišit.

- ii. identifikace uživatelských skupin a rolí v Katalogu (*pozn.: jednou ze skupin jsou mimorezortní uživatelé, kteří vytváří a vkládají datové sady do katalogu – tyto uživatelé musí mít vzdálený přístup do Katalogu*)
- iii. návrh a vytvoření procesního modelu řízení a správy environmentálních dat a informací, jehož účelem je zajištění veřejně ověřitelné transparentnosti a věrohodnosti datových výstupů, služeb, reportů a statistik resortu.
  - vytvoření řídicích pravidel a nastavení přístupů k datovým zdrojům v Katalogu a chování jednotlivých rolí (odpovědnosti a pravomoci), a to zejména nikoliv výlučně pro role knihovníka, správců/dodavatelů datových sad, provozovatelů IS, správce reportů/zprávy aj. (viz bod 1. písm. a. ii této **Přílohy č. 1a**).

#### Dodavatel:

- navrhne vhodný způsob organizačního zajištění (pravidla, kompetence, odpovědnosti) tak, aby byly pokryty veškeré kanály příjmu dat od poskytovatelů;
- popíše, jak motivovat poskytovatele k poskytování dat preferovaným způsobem, resp. jaké má přijmout opatření k zajištění nepřerušovaného předávání dat od poskytovatelů (nyní na bázi předávání .xlsx souborů e-mailovou cestou);

- popíše procesy a vztahy mezi rolemi, např. způsob vkládání údajů jednotlivými poskytovateli dat (nyní cca 40 subjektů včetně zahraničních-EUROSTAT, EEA), schvalovací, validační, publikační, řídicí a jiné procesy. *Návrh řešení Řídicího systému musí respektovat strukturu a formáty dat poskytovatelů, jelikož MŽP nemá mandát k definování těchto parametrů. S ohledem na to, musí dodavatel vhodně navrhnout procesy tvorby, editace, změn, schvalování, předávání, publikace a vytěžování dat a metadat v Katalogu tak, aby byly pro uživatele Katalogu minimálně časově zatěžující, tj. návrh nesmí přinést v praxi hromadění chyb, chybových stavů a časových prodlev);*
  - zpracuje analýzu rizik nasazení katalogizace a návrh opatření na jejich snížení.
- iv. vytvoření řídicí hierarchické vrstvy (číselníků, pravidel, postupů, apod.) pro katalogizaci dat a informací rezortu ŽP (dále také „rezort“) - doporučené postupy - návod - pro katalogizaci relevantních informačních zdrojů pro jednotlivé typy zájmových objektů (viz bod 1. písm. a. i této **Přílohy č. 1a a Příloha č. 3a**) prostřednictvím Katalogu pro jednotlivé role uživatelů Katalogu, včetně vysvětlení a vhodných příkladů za použití zvoleného SW Katalogu:
- zdůvodnění volby a možnosti SW nástroje katalogizace (Katalogu);
  - Uživatelský „manuál“ Katalogu:
    - práce s formuláři - příklady katalogizace všech identifikovaných typů zájmových objektů, tj. jak katalogizovat identifikované typy vstupních/výstupních (popisovaných) informací a dat; *návod, jak vkládat/evidovat nové datové zdroje, jak vyplňovat hierarchickou strukturu metadatového záznamu (určení typu, úrovně v hierarchii, stanovení vhodného názvu objektu, vysvětlení, typy a upozornění k vyplňovaným položkám, možnosti číselníků, povinné nepovinné atributy, způsob vkládání příloh /datové, jiné/, vkládání linků);*
    - práce s přehledy - možnosti a postupy filtrování, vyhledávání objektů a atributů;
    - návrh a vysvětlení validačních a schvalovacích mechanismů *(Je požadován návrh vhodného postupu /organizačně i aplikačně/ řešení kontrol formální a věcné správnosti dat ve vztahu k navrženým uživatelským rolím a způsobu příjmu /automatizovaně, ručně/, a to od převzetí primárních dat od poskytovatele až po potvrzení správnosti garantem. Musí být popsán způsob a forma zajištění validací s příklady pro všechny typy zájmových objektů, a to, jak pro úroveň dat, tak metadat).*
  - Administrátorský „manuál“ Katalogu:
    - vysvětlení postupů pro obsahové správy Katalogu *(např. práce s metadatovým modelem, nastavení uživatelů a rolí, návod jak publikovat pro veřejný (opendata) a neveřejný režim, rady pro roli „knižovníka“ (správce Katalogu) - co znát, co mít, na co se ptát poskytovatelů dat/informací, co zjistit pro katalogizaci, možnosti automatizace přebírání /předávání/ harvestingu metadat, možnosti synchronizace aj.);*

- používání řízených slovníků a způsob jejich použití, rozvoje a změn;
- nastavení a úpravy číselníků, provázání úrovní metadatových objektů, spolupráce s EAP aj.);
- vysvětlení principu provázaného (linked) katalogu a možností jeho pokročilého použití.

b. dodání SW nástroje/řešení pro katalogizaci (dále jen „**Katalog**“) zájmových objektů:

i. zvolený SW musí:

- být postaven na bázi standardu CKAN;
- být plně webovou aplikací založenou na standardu HTML 5;
- umožňovat:
  - výběr svodného metadatového standardu (podporovat min. Dublin core, aj.), knihovnického standardu pro popis produktů, standardu pro popis procesů, standardu popisu dat včetně prostorových a časových řad, vše ve vzájemné interakci;
  - užití různých metadatových standardů pro různé záznamy;
  - disponovat webovým GUI pro ukládání<sup>2</sup> nejen metadatového popisu, ale i zdrojových dat (např. datových agregátů) přímo v Katalogu (souborový/databázový systém) ať lokálně nebo v cloud infrastruktuře (provozované v rámci infrastruktury datového centra zadavatele);
  - validaci vkládaných údajů, resp. datových sad, webových služeb;
  - zpřístupnění datových agregátů/sad přes (1) webové rozhraní a (2) přes API;
  - přebírání metadatových záznamů z jiných rezortních/mimorezortních katalogů formou standardizovaných dotazů;
  - harvesting obsahu z/do jiných katalogů (včetně možnosti prolinkování instancí vlastního Katalogu a v nich uložených záznamů) a propojení s metadatovým katalogem geoportálu (geoportal.cenia.cz, resp. <http://ms2.cenia.cz>) a pravidelný odběr metadat, připravenost na zapojení dalších metadatových katalogů podřízených organizací rezortu (CENIA, ČGS, VÚV TGM, AOPK ČR, ČHMÚ, ŠUNAP, KRMAP, Správa národního parku České Švýcarsko, Správa národního parku Podyjí);
  - prostřednictvím linků přímý přístup k datům a jiným objektům (obrázky aj.), resp. jejich stažení;
  - nastavení přístupu k jednotlivým záznamům, agregátům a datovým zdrojům na základě jednotlivých rolí - víceúrovňové možnosti nastavení/administrace přístupových práv v návaznosti na organizační strukturu zadavatele;
  - nastavení schvalovacích workflow pro jednotlivé role uživatelů pro editaci, publikaci a selektivní zobrazení záznamů dle přístupových práv;

---

<sup>2</sup> Zadavatel předpokládá jednoduché webové rozhraní, v rámci něhož uživatelé v roli editorů budou moci nahrávat ve vazbě na katalogový záznam zdrojové datové soubory. Katalog, resp. jeho vhodná nastavba či EAP musí umožňovat validaci (na úrovni škodlivého kódu, kontrolu struktury datové sady vůči šabloně, kontrolu úplnosti dat, kontrolu atypických hodnot) a víceúrovňové schvalování (autorizaci garantem). V rámci metodiky dodavatel vytvoří optimální workflow pro řešení celého procesu. Součástí dodávky Katalogu je i implementace (či vytvoření a dodání) validačních a schvalovacích funkcionalit.

- přístup i mimorezortním uživatelům.
  - disponovat vícejazyčným rozhraním - minimálně čeština a angličtina včetně vícejazyčnosti metadat;
  - podporovat výměnné standardy RSS, Atom, RDF;
  - obsahovat:
    - API: pro přístup k záznamům, buď podle standardu RPC nebo REST a umožnit poskytnutí záznamů minimálně ve formátu JSON a optimálně i XML/CSV;
    - webové (GUI) rozhraní: umožní správu obsahu i uživatelů; zpřístupnění obsahu, je snadno změnitelné formou nezávislých šablon:  
➤ Katalog umožňuje napojení na LDAP a IDM Objednatele.
  - splňovat níže uvedené nefunkční technické požadavky na:
    - Výkon - SW je prokazatelně (formou oficiální reference nebo demonstrační aplikace) schopen zajistit odezvu devíti z deseti současně pracujících uživatelů do čtyř vteřin na virtuálním stroji s jedním CPU a 8GB RAM nad tisíce záznamy v úložišti;
    - Škálovatelnost - výkonnost SW je možné zvýšit převedením do cloud infrastruktury beze změny pro uživatele, finančních nároků na SW část a změn pro správce systému;
    - Spolehlivost - SW je aktivně udržován a rozvíjen, existuje na trhu již více než 5 let a má alespoň padesát instalací;
    - Rozšiřitelnost - pro SW existují zásuvné moduly (pluginy), kterými je možné rozšířit funkcionalitu;
    - Udržitelnost a spravovatelnost - SW má integrovaný monitoring/logování a nástroje pro snadnou správu přes webové rozhraní a API;
    - Bezpečnost - vzhledem k určení SW pro centrální management rezortních informačních zdrojů je nezbytně nutné zajistit:
      - Autentizaci s možností využití jednotného rezortního vstupu (single sign on);
      - Autorizaci - správa rolí a odpovědností;
      - Šifrování - https pro přístup k webovému a API rozhraní, možnost šifrování dat;
      - Integrita - garance, že obsah zprávy není změněn při přenosu, uložení, veškeré změny jsou archivovány, existuje konzistentní log veškerých provedených akcí;
      - Neodvolatelnost - jednou uložená data a záznamy nelze smazat, pouze aktualizovat či označit jejich platnost za ukončenou.
- ii. zvolený SW musí obsahovat správcovskou příručku technické správce SW s návodem na zprovoznění a doporučeními pro konfiguraci, zálohování, škálování, rozšiřitelnost apod. pro všechny moduly.
- c. vytvoření metadatového modelu založeného na zpracované metodice a seznamu zájmových objektů a identifikovaných atributů:
- i. metadatový model:
- musí být v souladu s normami ČSN ISO 15836, ISO19139;
  - bude obsahovat:
    - dokumentaci použitých metadatových prvků;
    - příklady použití metadatových prvků v konkrétních katalogových listech.

- d. dodání/postoupení licencí zvoleného katalogového SW/Katalogu (včetně všech integrovaných pluginů/modulů)
- e. zprovoznění katalogu (SW)
  - i. zprovoznění instance/í Katalogu;
  - ii. konfigurace, parametrizace, customizace (včetně grafiky, CSS);
  - iii. instalace pluginů (*zejména pluginů pro správu rolí katalogu, práci s geoinformacemi včetně harvestingu*) nebo autonomních SW komponent (*např. validačního a příjmového modulu pro data, schvalovacího modulu apod. nebude(ou)-li k dispozici jako nativní plugin/y*);
  - iv. integrace vytvořeného metadatového modelu (viz bod 1. písmeno c.);
  - v. publikace Katalogu v rámci vnitřní sítě MŽP, resp. rezortní sítě a na webu;
  - vi. propojení se SW MICKA (<http://ms2.cenia.cz>) a harvesting vybraného množství (min. 40) metadatových záznamů typu geoDCAT do Katalogu.
- f. naplnění modelovými (pilotními) metadaty
  - i. úvodní naplnění základními statistickými a geostatistickými daty (základní územní číselníky ČSÚ včetně hranic - možnost využití [opendat](#), tj. administrativní územní členění a NUTS/LAU) pro umožnění rychlého přechodu uživatelů na novou platformu;
  - ii. vytvoření formulářů pro katalogizaci všech kategorií zájmových objektů na základě Metodiky (viz bod 1. písm. a. této **Přílohy č. 1a**);
  - iii. vytvoření nejvyšší obsahové úrovně hierarchie (viz **Příloha č. 3a**, úroveň I-IV) a jejich číselníků (minimálně pro úroveň: tematická oblast/doména ŽP, zákon, agenda/úkol/projekt/strategický cíl, informační služba).
  - iv. vytvoření metadatových záznamů ke:
    - [Zprávě o životním prostředí za rok 2016](#) (dále „Zpráva“) - *katalogizace všech informačních výstupů Zprávy - cca 220 grafů, tabulek, map - a jejich provázání na hierarchicky navazující metazáznamy*;
    - všech Objednatelům identifikovaných reportingovým povinností MŽP (243), (viz <https://owl.mzp.cz/index.php/s/XEZjUnpg2KTtK7W>).
  - v. za vybrané agregované datové sady dodané zadavatelem:
    - provedení podrobné katalogizace 7 dodaných datových zdrojů (2 RDMS, 3 textový soubor (XLSX, CSV, JSON), REST služba, SOAP služba) tak, aby byly typově zastoupeny všechny možnosti.
      - zadavatel poskytne součinnost a podklady;
      - bude existovat provázání na nejvyšší úroveň hierarchie včetně přiřazení role pro oprávněné uživatele;
      - obsahem bude strojově čitelný datový slovník datového zdroje (soubor CSV nebo JSON).
    - provedení katalogizace 3 typových indikátorů Zprávy včetně úplné dekompozice na zdrojové datasety (*např. Emise primárních částic a prekursorů sekundárních částic*):
      - zadavatel poskytne součinnost a podklady;
      - bude existovat provázání na nejvyšší úroveň hierarchie.
    - provedení katalogizace 2 typových reportů/reportingových povinností:
      - zadavatel poskytne součinnost a podklady;
      - bude existovat provázání na nejvyšší úroveň hierarchie.

- vytvoření vzorových CSV a JSON šablon pro 12 vybraných dekomponovaných datasetů získaných v předchozím úkolu:
    - šablony budou připojeny k relevantním katalogizačním záznamům.
  - vytvoření datových slovníků (datových struktur) pro na základě zpracovaných šablon k výše uvedeným 12 dekomponovaným datasetům:
    - tyto datové slovníky budou sloužit pro skladování a analýzu dat získaných z uvedených dekomponovaných datasetů.
  - provedení úplné katalogizace aktuálně publikovaných OpenDat rezortu (nyní 36 sad z 12 organizací, blíže viz [opendata.mzp.cz](http://opendata.mzp.cz)).
- vi. zpracování doporučení ve smyslu identifikace nezbytných kroků, jež musí zajistit Objednatel, pro zajištění kompletního naplnění Katalogu daty a metadaty za agregované datové sady, jež jsou nyní vedeny v ISSaR a pro zajištění přechodu mezi stávajícím způsobem přebírání a vedení agregovaných dat novým způsobem navrženým v Metodice (viz bod 1. písm. a. této **Přílohy č. 1a**)
- g. Prezentace a školení
- vyškolení budoucího „knihovníka“ a vybraných 5 editorů (správců datových sad) v použití navržených postupů a nástrojů v rozsahu 8 hodin.
  - praktický workshop pro knihovníka a editory v rozsahu 6 hodin navazující na školení (nejpozději do 14 dnů od vyškolení budoucího knihovníka dle předcházejícího odstavce). Součástí bude prezentace výstupů a poznatků Objednateli.

## **2. Implementace centrálního objektového úložiště pro analytické a publikační účely**

- a. dodávka SW nástrojů pro ukládání datových sad rezortu určených pro analytické a publikační účely (dále jen jako **environmentální analytická platforma**, zkr. „**EAP**“) - Smyslem EAP je zabezpečit objektové úložiště, prohledávací, indexovací a ETL nástroj/prostředí, které by umožňovaly přebírat vybraná data z produkčních databází či souborů, zabezpečit transformaci dat do standardizované objektové struktury, uložení a případné obohacení (enrichment) tak, aby bylo možné vytvářet konsolidované datové sady určené pro statistické vytěžování, předávání či publikaci „surového“ datového obsahu přes jednotný katalog (viz bod 1 písm. b. této **Přílohy 1a**).
- i. zvolené SW řešení EAP musí:
- být kompatibilní se zadavatelem používanou technologií Elasticsearch, kterou již zadavatel pro vybrané datové zdroje používá;
  - disponovat nativním REST API pro automatizovanou výměnu dat;
  - obsahovat výkonný prohledávací engine a objektové datové úložiště optimalizované pro rychlé dotazování nad uloženými daty obsahujícími desítky až stovky atributově heterogenních datových objektů;

- umožňovat sledovatelnost<sup>3</sup> (tracebilitu) dat - tj. pokročilé vyhledávací funkce, a to včetně webového GUI pro snadné vyhledávání a filtrování uživateli;
- umožňovat škálování výkonu;
- být provozovatelný na virtualizovaném HW prostředí;
- disponovat nástroji ETL pro automatizované předávání a transformaci dat z různých datových zdrojů (minimálně podpora souborových vstupů typu .xlsx, .csv, .xml, .json, .dbf, aj., relačních databází - JDBC, ODBC /mysql, mssql, postgresql, oracle/, webové služby /SOAP, REST/; nerelační databáze-mongodb; ideálně také RDF formáty);
- umožňovat replikaci dat - data lze uživatelsky snadno importovat z formátu CSV nebo z Excelu, nebo pro informační systémy pomocí API; uložená data lze následně externími nástroji vizualizovat jako tabulky, grafy či mapy;
- umožňovat mapování a fulltextovou indexaci strukturovaných i nestrukturovaných dat (PDF, DOCX, apod.);
- umožňovat použití a definování textových slovníků pro pokročilou indexaci obsahu dat;
- Pozn.:
  - zadavatel preferuje, aby SW úložiště EAP bylo, s ohledem na úsporu nákladů na provoz a správu, integrované s úložištěm Katalogu;
  - zadavatel umožňuje použít pro řešení opensource produktů (konkrétně použít technologii Elasticsearch nebo alternativní technologie poskytující minimálně stejný rozsah služeb jako technologie Elasticsearch a je zároveň kompatibilní na úrovni REST API s technologií Elasticsearch).

b. dodání/postoupení licencí SW řešení EAP zadavateli

c. zprovoznění a konfigurace EAP včetně oddělených instancí pro data ze zdrojových systémů (nebo datových sad) a obohacená data určená pro konkrétní analýzy nástroji<sup>4</sup> BI

d. dodání metodiky pro práci s EAP

- i. popis (výstižný návod) postupů ETL pro všechny identifikované druhy datových zdrojů (*struktura ukládání dat, konfigurace služeb, spouštění služeb, logování, datové zdroje, konektory, mapování dat*), obohacování dat, filtrace s daty aj. s komentovanými příklady
- doporučení pro oblast dekompozice komplexních datových sad
  - doporučení pro oblast vytváření datových slovníků k datovým sadám

---

<sup>3</sup> Ukládané data v EAP jsou určena pro další analytické využití. Tato data mají charakter vysoké dynamičnosti (jak na straně vstupů pro EAP - změna struktury, hodnot, a tak následných výstupů z EAP). Je proto nezbytné, aby v každé chvíli, o každém uloženém údaji (hodnotě) v EAP bylo možné zjistit, kdy a v jaké kvalitě byl přijat, zda je stále platný, pokud již platný není, od kdy a jakým údajem byl nahrazen. Kromě toho musí EAP být schopen s těmito informacemi pracovat (tj. např. vypublikovat současně data s různou kvalitou, dobou platnosti, automaticky nahrazovat starší data aktuálnějšími atd.).

<sup>4</sup> Dodávka těchto nástrojů je mimo Předmět plnění této veřejné zakázky.



- ii. metodické postupy pro vkládání dat na příkladech analyzovatelných datasetů<sup>5</sup> i primárních datových zdrojů do datového skladu EAP (identifikace potřeb konsolidace dat, doporučení pro pseudonymizaci zdrojových dat, importní konektory, konfigurace získávání a extrakce dat pomocí ETL) včetně možností automatizace zejména ve vazbě ke katalogu a rolím uživatelů;
  - iii. metodické postupy pro využití EAP pro analýzu dat - způsoby napojení na nástroje BI s příklady;
  - iv. metodické postupy pro katalogizaci a ukládání odvozených/obohacených dat (platí princip, že všechny konverze musí být dokumentovány; *nutné respektování metodiky dle bodu 1 písm. a.*);
  - v. doporučení pro publikaci formou opendat (musí být konzistentní s Metodikou dle bodu 1. písm. a. této **Přílohy č. 1a**);
  - vi. návod a pravidla pro vytváření synchronizačních /“push“/ skriptů.
- e. vytvoření a otestování modelových „pipelines“ (ETL procedur) pro automatizované anebo poloautomatizované předávání dat EAP:
- i. vybraná agregovaná data (v podobě CSV dokumentů):
    - s vybranými 12 dekomponovanými datasety (viz bod 1. písm. f. v. této **Přílohy č. 1a**) vytvořit příslušné ETL procedury a uložit je do platformy EAP;
    - uložit kódový zápis (algoritmy) ETL procedur pro opakovatelnost použití;
    - pro procedury vytvořit relevantní katalogizační záznamy v katalogu včetně odkazu na kódový zápis procedury;
    - vytvoření strukturovaného schématu (CSV, JSON) pro každou datovou sadu a jeho umístění ve zvoleném úložišti dostupném přes jedinečné URL.
  - ii. kopie dat 2 produkčních SQL databází vybraného IS rezortu:
    - na příkladu MA ISOH (Firebird, velikost DB cca 40 tabulek - v hlavní tabulce cca 1,2 mil. záznamů) a IRZ (MS SQL, 8 tabulek, v hlavní tabulce cca 50 tis. záznamů);
    - provedení analýzy, vytvoření ETL procedury a kompletní katalogizaci MA-ISOH, IRZ.
  - iii. data vybrané existující komplexní WSDL služby rezortu ([IPPC](#) nebo [ISPOP](#))
    - včetně:
      - vytvoření konverzního konektoru umožňujícího parametrizovatelné dotazování a obsahujícího synchronizačního agenta na WSDL (*v případě, že by toto nebylo možné zajistit ETL nástrojem*);
      - dodání licence k tomuto konektoru s možností SW nezávisle upravovat;
      - provedení úplné katalogizace zdroje WSDL pocházejících z IPPC (ISPOP), včetně ETL dat (*eventuálně také konverzního konektoru*).
    - provedení vzorového výpočtu alespoň jednoho indikátoru ŽP na základě katalogizačních záznamů v katalogu a dat uložených v EAP.

---

<sup>5</sup> Data, která MŽP pro vytváření svých výstupů pro tvorbu Zprávy potřebuje, jsou data agregovaná, většinou je proto nejde získávat prostým napojením na informační systémy, shromažďující primární data - na straně dodavatele dat je většinou nutná úprava těchto dat.

- f. zprovoznění vybraných ETL procedur a naplnění úložiště daty z výše uvedených zdrojů (bod 2 písm. e.) - *v případě, že bude v rámci zprovoznění ETL procedur vytvořeno autorské dílo, je požadováno, aby byla zároveň Dodavatelem Objednateli dodána k užívání tohoto díla nevýhradní licence umožňující nezávislý rozvoj díla s možností udělování podlicence Objednatelem třetím stranám bez písemného souhlasu Dodavatele.*
- g. vytvoření typového synchronizačního nástroje (agenta) pro průběžnou aktualizaci dat poskytovaných z níže uvedených zdrojů
- i. provedení a otestování na příkladu WSDL IPPC (viz bod 2. písm. e. iii. této **Přílohy 1a**);
- ii. pro 3 vybrané sady agregovaných dat<sup>6</sup> (viz bod 2. písm. e. i. této **Přílohy 1a**);
- iii. jedné relační databáze (viz bod 2. písm. e. ii. této **Přílohy č. 1a**).
- h. zpracování obohacené datové sady pro vybraný modelový indikátor Zprávy
- i. včetně:
- uložení kódového zápisu procedury pro opakovatelnost použití;
  - vytvoření relevantního katalogizačního záznamu v katalogu včetně odkazu na kódový zápis procedury;
  - zpracování návodu/postupu, jak v rámci EAP provádět obohacení datových sad, jak tyto sady předávat externímu BI nástroji a to včetně doporučení požadavků na nezbytné funkcionality BI nástroje, které jsou nezbytné pro to, aby bylo možné obohacená či primárně uložená data EAP prostřednictvím REST API použít jako datový zdroj pro BI, a to včetně požadavků na zabezpečení synchronizace - nebude-li zajišťovat EAP.
- i. provedení demonstrace modelové statistické vizualizace pomocí Dodavatelem vybraného externího nástroje BI (*ten není součástí řešení*); vybraný statistický/BI SW dokáže přistoupit nativně (bez manuálního exportu dat) přímo k EAP přes nativní konektor nebo REST API a získat z něj data, nad kterými provede příkladový výpočet (např. sumarizace hodnot) a grafovou+mapovou vizualizaci.
- j. školení technického správce EAP (5 hodin)
- k. prezentace postupu řešení a výstupů bodu 2. této **Přílohy č. 1a** formou workshopu (5 hodin)

---

<sup>6</sup> V situaci, kdy jsou v systému publikované stovky datových výstupů (tabulky, grafy, kartogramy) vzniklé kombinací a přepočty řádově stovek vstupních datových zdrojů, není fyzicky možné manuálně zajišťovat aktualizaci výstupů při každé změně vstupních dat. Tento proces musí být nastaven automaticky.

### **3. Maintenance SW dodaných a implementovaných v rámci bodů 1. a 2. na 5 let (60 měsíců)**

Standardní údržba a servis SW<sup>7</sup> - maintenance znamená provádění všech prací spojených s podporou a aktualizací SW částí Řídicího systému včetně jejich SW platformy, odstranění závad či optimalizace výkonnostních limitů jednotlivých prvků SW architektury Řídicího systému a udržování Řídicího systému v řádném a plynulém provozním stavu tak, aby byly dodržovány parametry 98 %<sup>8</sup> dostupnosti v pracovní dny od 8:00 do 16:00 hodin („provozní doba“) a bylo minimalizováno riziko ohrožení dodávky služeb (funkcionality) Řídicího systému uživatelům a minimalizováno riziko zneužití údajů v něm obsažených:

Standardní údržba SW zahrnuje především, nikoliv však výlučně, následující činnosti:

- správa, dohled (sledování zejména parametru dostupnosti) a podpora SW (včetně webových a databázových serverů);
- eliminace technických a bezpečnostních incidentů po nahlášení Objednatel dle následující parametrů - v případě havárie do 12 hodin, výpadku do 48 hodin, závady do 5 pracovních dnů od nahlášení:
  - Havárie = činnost SW je zcela nebo podstatně omezena, všechny nebo důležité části selhaly nebo jsou nedostupné, poskytují výrazně zhoršenou odezvu, selhaly a/nebo jsou zcela nefunkční nebo je jejich funkčnost omezena tak, že je kritickým nebo zásadním způsobem ovlivněna činnost SW. Stav některé ze SW částí Řídicího systému neumožňuje plnění základních funkcí.
  - Výpadek = SW je funkční pouze částečně, je ovlivněn selháním nebo omezením některé ze systémových funkcí podporujících důležité činnosti SW, např. některá z webových služeb/skriptů vykazuje funkční vady, pouze některé funkce nejsou plně funkční. Stav některé ze SW částí Řídicího systému umožňuje plnění základních funkcí, avšak s omezením rychlosti zpracování nebo za mimořádných provozních opatření (např. provizorní provoz s vynaložením většího úsilí či se zvýšenými náklady)
  - Závada = SW je funkční, vyskytují se však nedostatky nepodstatné povahy, které způsobují například nekomfort obsluhy nebo zvyšující se pracnost činností nad rámec pracovní obvyklé v běžném provozu nebo některé funkce selhaly, ale nejsou v daný moment využívány nebo nemají žádný vliv na řádný chod SW anebo je mírně zvýšena odezva SW. Stav některé ze SW částí Řídicího systému umožňuje plnění základních funkcí, avšak s vyskytujícími se drobnými chybami, které nebrání nebo mají zcela minimální vliv na řádné užívání a funkcionality.
  - Kategorizaci incidentu provádí Objednatel a povinností Dodavatele je reagovat bez zbytečného odkladu, nejpozději však do 12 hodin v pracovní dny od nahlášení incidentu a započít řešení nastalé situace. Objednatel je oprávněn v odůvodněných případech prodloužit lhůtu na obnovení činností SW do řádného provozního stavu.
- Sledování znalostní báze výrobců jednotlivých SW komponent Řídicího systému, vyhledávání a implementace vhodných oprav, konfigurace SW, údržba, podpora a aktualizace SW (minoritní i majoritní aktualizace);
- Aktivní vyhledání a identifikace oprav, bezpečnostních záplat, patchů, hotfixů, nebo servicepacků včetně jejich vývoje/stažení, uložení a implementace;

<sup>7</sup> SW komponenty Řídicího systému (dále „SW“) budou technicky provozovány Objednatel na jeho HW a síťových prostředcích včetně virtualizačního SW a serverového operačního systému.

<sup>8</sup> akceptovatelná nedostupnost SW (mimo předem dohodnuté servisní odstávky SW) bez sankce je kumulativně max. 3 hodiny a 12 minut v rámci Garantované provozní doby. Trvajících nedostupnost více jak 120 min. v kuse je považována za incident typu havárie.

- Implementace aktualizace či rekonfigurace SW i na vyžádání Objednatelem;
- Garance dodávky a implementace nových verzí SW (po vydání jejich autorem/komunitou) s vyhodnocení rizik případné nekompatibility;
- Příprava a dodání exportů dat a metadat vedených v SW, které nebudou dostupné Objednateli (nebo administrovatelné Objednatelem) přes aplikační prostředí Řídicího systému na jeho vyžádání;
- Reporting – zabezpečení a vedení on-line Provozního deníku s dálkovým přístupem Objednateli – sledovaným parametrem je řádné zapisování úkonů specifikovaných body výše s datem provedení příslušného úkonu a krátkým výstižným popisem a jejich kategorizací (tj. instalace updatů, patchů, řešení incidentů, záznamy nedostupnosti SW-aplikačních služeb, provedené rekonfigurace, poskytnutí exportů apod.). Vždy ke konci kalendářního měsíce bude Dodavatelem zaznamenán do Provozního deníku 1 souhrnný záznam, který konstatuje minimálně (1) počet provedených úkonů maintenance za dané měsíční období zaznamenaných do deníku (tj. hodnotu 0 až n odpovídající počtu záznamů deníku za dané období) a (2) konstatováním, zda v daném měsíčním období došlo či nedošlo k porušení parametru dostupnosti SW. Součástí Provozního deníku musí být zároveň sekce pro zápis porušení provozních parametrů s možností editace zástupcem Objednatele. Nebude-li záznam o provedení daného úkonu zaznamenán do 5 pracovních dnů od jeho provedení, je tento fakt důvodem pro vyměření smluvní sankce. Provozní deník musí být zřízen nejpozději v den vydání Finálního akceptačního protokolu. Provozní deník musí být vytvořen v technologii umožňující min. časové logování provádění záznamů.