

Název projektu:

Redesign Statistického informačního systému v návaznosti na zavádění eGovernmentu v ČR

Příjemce: Česká republika – Český statistický úřad

Registrační číslo projektu: CZ.1.06/1.1.00/07.06396

## Příloha k zadávací dokumentaci veřejné zakázky „Integrační nástroje, vstupní a výstupní subsystém“

### **Příloha č. 20**

## **Specifikace bloku CENTRAL**

Název souboru: RSIS\_ZD001P20\_SMS\_CENTRAL.pdf

Počet stran přílohy (bez tohoto krycího listu): 6

Administrace přílohy: Ing. Pavel Skotnica

*Verze ke zveřejnění*

## Zadání subsystému SMS-CENTRAL

### 1. Základní charakteristika

- 1 | Subsystém SMS-CENTRAL zabezpečuje realizaci centrálního zpracování dle zadaných parametrů. Centrální zpracování zahrnuje následné automatizované procesy:
- a) nahrání dat do centrálních databází - ze VSTUPU, z administrativních zdrojů, z parametrizačních a datových tabulek potřebných pro zpracování;
  - b) zpracování dat - automatizovaná příprava dat, čištění, sehrání dat, nastavení meziproměnných, provedení odhadů, imputací, převážení, testování odlehlostí, identifikování primární důvěrnosti až po tvorbu dat vhodně uložených pro následnou tvorbu výstupů a definování datových zdrojů (vstupní VIP a zdrojové úlohy) pro výstupní objekty časové řady včetně zajištění zpětných přepočtů k zajištění srovnatelnosti časových řad;
  - c) tvorba výstupních objektů (exporty .dbf, exporty pro sezonní očišťování, pro stanovení důvěrnosti, SMS-KVALITA, generování časových řad, aktualizací dávky pro registry a jiné subsystémy SIS, exporty do UF a prostřednictvím operátorského rozhraní pro DWH a převod do DWH, Datová tržiště, VDB, apod.).
- 2 | Pro řízení zpracování je nutná parametrizace procesu a v rámci realizace i evidence samotného procesu (KDO, CO, KDY, ČÍM a JAK) se svými parametry, které jsou:
- a) získávané z jiných subsystémů;
  - b) nastavovány správcem procesu dle požadavku gestora úlohy;
  - c) stanovené před samotnou realizací gestory úlohy;
  - d) přebírány z parametrizačních nebo podpůrných tabulek, které spravují příslušní zaměstnanci zodpovídající za svou oblast a v určitých případech musí hlásit připravenost pro zpracování, resp. možnost uvolnění výstupů.
- 3 | Subsystém je silně napojen na ostatní subsystémy SMS (ULOHY, KLAS, UKAZ, KVALITA, VYSTUPY, CASOVE RADY) a systémy REGISTRY, transformace, VSTUP, EVID a zároveň připravuje podklady pro následné subsystémy (DISEMINACE, DWH, KVALITA, ...).

### 2. Popis současného stavu

- 4 | V současné době je samotné centrální zpracování řízeno útvarem, který spravuje chod centrálního zpracování pro vybrané podnikové statistiky. Zdejší popis se netýká zpracování sociálních, zemědělských, cenových a demografických statistik, které jsou spravovány přímo gestory úloh na svých pracovištích.
- 5 | Samotné řízení centrálního zpracování se odvíjí od technologie popsané v technických projektech v produktu PMAN, které v sobě mimo jiné zahrnují specifikace pro centrální zpracování, jako jsou harmonogramy, primární nastavení parametrů, souvislosti s jinými úlohami apod. Realizují se zde jak úlohy přímo vycházející z účelu statistiky, tak i podpůrné úlohy, jako jsou nahrání administrativních dat, aktualizace KDB, aktualizace RES apod.
- 6 | Některé procesy centrálního zpracování jsou prováděny mimo výše uvedený útvar a někdy dochází k neplánovaným kolizím. Také ne vždy, vzhledem k sofistikovanosti některých úloh, jsou spouštěny ve správném sledu nebo úplně, někdy jsou naopak spuštěny programové moduly zbytečně. Ani není vždy jasné kdy a kým a s jakým nastavením parametru byla příslušná úloha realizována. Je vytvořeno už několik nezávislých tabulek v databázi, z kterých lze vyčíst vybrané metainformace, ale metodika je zatím nejednotná a týká se vybraných procesů.
- 7 | Pro samotné spouštění je vytvořena aplikace, přes kterou se spouští zásadní vybrané procesy, jako je:
- nahrání dat z OSZ;
  - zpracování dopočtů;
  - tvorba výstupů do .dbf, .xlsx, .xls, tvorba výstupních sestav;
  - aktualizace VS, ZS;

- nahrání do KDB z transformace;
- nahrání do KDB dle návrhu OSZ.

8 | Nicméně informace z tohoto procesu nejsou vždy zaznamenány v patřičné šíři, nebo se nesnadno dohledávají. Dále jsou zde aplikace na správu KDB společně s nastavením úrovně odhadů (aplikace ODSOUHLAS a KDB), na nastavení verze a běhu zpracování a evidence revidovaných období (aplikace REVIZE), na evidenci změn/mazání/přidání dat (aplikace extico2). Existují nástroje pro nahrání dat z administrativních zdrojů. K dispozici je aplikace pro nahrání dat do DWH (aplikace OR). Na druhou stranu neexistuje obecný nástroj pro převod externích dat z vybraných formátů a ani obecný nástroj pro export s možnou specifikací do vybraných formátů s výjimkou aplikace pro import,/export z/do .dbf.

### 3. Cílový stav

- 9 | Mít k dispozici souhrnnou aplikaci, která v sobě integruje:
- a) specifikaci procesu automatizovaného zpracování na základě volby programových chodů, resp. programových modulů;
  - b) spouštění definovaného procesu, který eviduje procesní atributy, vstupní parametry a samotné verze programových modulů;
  - c) úložiště programových modulů, storovaných databázových programových objektů a pohledů, creatovacích skriptů;
  - d) evidenci požadovaných datových změn (opravy dat), nastavení extrému;
  - e) nastavení úrovně odhadu, verzí a běhu zpracování, uvolnění a zveřejnění dat, nastavení datumu pohledu do podpůrných databází a naplnění dalších požadovaných metainformací ke zpracování;
  - f) systém KDB a vkládání oprav pro KEČ;
  - g) volání/prolinkování se na aplikaci operátorského rozhraní na správu a nahrávání pro DWH, DM a VDB.
- 10 | Další záležitosti týkající se centrálního zpracování jsou uvedené níže, ale budou předmětem zadání jiné aplikace:
- h) aplikaci pro zadávání a samotnou realizaci nahrání externích a administrativních dat do Centrální databáze z různých dostupných formátů (.dbf, .xls, .xlsx, .xml,...);
  - i) aplikaci pro zadání a samotnou realizaci pro buďto převod uložených externích a administrativních dat v centrálním úložišti do standardních formátů uložení pro centrální zpracování nebo pro opačný převod ze standardního uložení do zvolené struktury a zvoleného formátu;
  - j) aplikaci GVO;
  - k) napojení na SMS-KLAS, UKAZ,KVALITA, VYSTUPY a CASOVE RADY a plnění subsystémů KVALITA;
  - l) napojení a plnění SIS-DWH, CENTRAL, DISEMINACE, VDB.
- 11 | Hlavní aplikace subsystému umožní spouštění a evidenci programových chodů, které se mohou skládat z několika programových modulů. Také **umožní přístup do zadávání**, evidence a uložení vstupních parametrů pro řízení procesů, zásahů do datových souborů, zadávání metainformací, tvorbu metainformací procesů související s Centrálním zpracováním, Kvalitou, DWH a tvorbou výstupů. Všechny změny jsou evidovány a identifikovány.
- 12 | Cílový systém poskytne pracovníkům OCZ řízení, spouštění, parametrizování, plnou kontrolu a sledování nad průběhem zpracování statistické úlohy v centrální databázi.
- 13 | Pracovníkům VPO umožní parametrizovat zpracování a sledovat jeho průběh, případně analyzovat proces zpracování po jeho ukončení.

14 | Systém směřuje přes datové uložení v Centrální databázi k propojení na ostatní subsystémy SMS-ULOHY, KLAS, UKAZ, VYSTUPY a připravuje vstupy pro GVO, SMS-KVALITA, SIS-DWH, DISEMINACE, VDB.

15 | Navrhovaný systém by měl být otevřený pro případné další doplňování funkčnosti a měl by poskytovat takové rozhraní vůči svému okolí, které bude jeho schopnosti dále rozvíjet a zvyšovat komfort pro uživatele zabývající se správou zpracování a řízení úloh.

#### **Obecné pojednání**

16 | Výstupem projektu je navržení a později schválená aplikace, která by měla splnit níže vyjmenované funkcionality společně se zavedenými rolemi a s definovanou návazností a vazbami na další subsystémy SMS-a subsystémy SIS. Předmětem zadání je také datový model databází, nad kterými je aplikace postavena a specifikování náročnosti na výpočetní techniku a SW požadavky.

17 | Obecně je centrální zpracování jednou z nedílných součástí procesu zpracování statistické úlohy. Existují úlohy zpracovávající (odhady a imputace, editace dat, řešení nonresponse apod.) mimo systém Centrální databáze, přesto i pro tyto úlohy musí aspoň dojít:

- k nahrání výstupních dat do Centrální databáze;
- nahrání příslušných metainformací (v omezené míře nadefinovat během importu);
- převedení dat s metainformacemi do DWH, VDB, a metainformací do SMS-KVALITA.

## **4. Funkční specifikace předmětu**

### **4.1. Funkcionalita**

18 | Funkcionalita aplikace CENTRAL lze rozdělit do následujících věcných bloků.

1. Blok INICIAL - Natažení předdefinovaného zadání ze subsystému SMS--ULOHY týkající se samotného procesu zpracování a přebírání metadat z jiných subsystémů SMS a SIS.
2. Editace zadání procesu zpracování.
3. Parametrizace zpracování ze strany gestor(ů) úlohy:
  - úroveň zpracování, revize, oprava/insert dat, nastavení KEČ.
4. Realizace zpracování.
5. Naplnění rozhraní pro další subsystémy, resp. spouštění z Centrální databáze exporty, které jsou importována do svých systémů (DWH, DISEMINACE,VDB, VSTUP, SMS-KVALITA a jiných externích zdrojů).
6. Správa aplikace CENTRAL.

#### *Ad 1. INICIAL*

19 | V subsystému SMS (zejména v subsystému ULOHY) se provádí zadání a příprava statistických úloh. Zde jsou předdefinovány konkrétní procesy zpracování společně s harmonogramy a příslušnou parametrizací spouštění běhů a verzí zpracování. Tím je vymezené, které konkrétní programové chody a moduly mají být spouštěny a to nejen v rámci Centrálního zpracování. Aplikace SMS-CENTRAL inicializačně pro zvolenou úlohu a referenční rok(y) převezme příslušné informace, které chápe jako inicializační nastavení pro zpracování, které jsou citované v následném funkčním bloku. Většina informací je metadatového charakteru, jejichž popis je svázán ze SMS-KLAS, UKAZ, VYSTUPY, CASOVE RADY a vybrané informace mohou být editované je s návazností na SMS-KLAS, resp. SMS-UKAZ

#### *Ad 2. Editace zadání procesu zpracování*

20 | Realizace centrálního zpracování konkrétní úlohy/verze/běhu zpracování znamená automatizované zpracování nad centrální databází při zvolené parametrizaci. Je provedeno spouštěním určitých programových modulů a chodů.

21 | Aplikace CENTRAL umožní před samotným spuštěním změnu specifikace konkrétních programů. Změna může být podle volby jednorázová nebo trvalejší (do další změny). Z evidence programů lze vybrat, které konkrétní programy a v jakém sledu jsou příslušným procesům přiřazeny a při samotné realizaci spuštěny. V rámci správy systému jsou programy evidovány i se svým metainformačním popisem (verze programu, popisem činnosti, autorem, datумы tvorby/změn a evidence apod.). Dále

je zde možnost specifikace, které procesy mají být zahrnuty pro realizaci a to opět s možností jednorázového využití, resp. trvalého nebo časového, resp. typového omezení. Typové omezení je svázáno s typem verzí, resp. běhu zpracování. Časové omezení je dané konkrétním datem platnosti.

#### *Ad 3. Parametrizace gestory úloh*

22 | Již byla zmíněna nevyhnutelnost centrálního zpracování pro jakoukoliv statistickou úlohu, jejíž výstupy mají být uloženy do DWH a případně prezentovány VDB. Pro evidenci jakéhokoliv údaje musí být vedena informace o:

- způsobu získání údaje (pořízen, odhadován,...) – povinné;
- důvěrnosti – nepovinné;
- kvalitě – nepovinné;
- VIP (zahrnuje v sobě napojení na SMS-UKAZ, SMS-KLAS, ale i případný statistický popis – MJ, ocenění, funkci, časové rozlišení...) – povinné;
- referenčním obdobím – povinné, případně i základním (bazickém) obdobím – nepovinné;
- úloze, verzi zpracování a běhu zpracování – vše povinné;
- příslušnosti k časové řadě – nepovinné, pokud není součástí časové řady.

23 | Gestor úlohy v rámci procesu centrálního zpracování nastavuje:

- verzi a běh zpracování;
- míru revizí (která referenční období se mají revidovat ve výstupech);
- úroveň schválení dat, která příslušné výstupy uvolní pro další úlohy;
- výstupy a datum a úroveň zveřejnění;

navíc pro podnikové statistiky:

- úroveň využití databází KDB a související s klasifikací ekonomických činností (podle úrovně odhadu, resp. data pohledu);
- využití RES a jiných podpůrných databází, které většinou vycházejí z administrativních zdrojů dat.

24 | Dalším aspektem ovlivňovanými gestory úloh jsou nastavení

- odlehklých dat (ELIMů) a to na úrovni úlohy, období, výkazu, část výkazu, oddílu, položky;
- změn resp. insertů dat/meziproměnných/atributů a to z možností trvalé ne dočasné úrovně.

25 | V rámci toho je umožněno specifikovat, zda-li se změna má promítnout před zpracováním, resp. na konci zpracování. Také je umožněn hromadný import změn ELIMů i změnu dat z .xls, resp. .xlsx.

26 | Vedle výše uvedených funkcionalit systém umožní rozšíření o vyvolání spuštění dalších aplikací, které souvisí s parametrizací zpracování ze strany gestora úlohy.

#### *Ad 4. Realizace zpracování*

27 | Aplikace zajišťuje samotné spuštění procesu zpracování, při kterém lze formou menu vymezit vybrané programové moduly a chody, resp. jejich skupiny. Moduly je možné spojovat v různé skupiny a různou hierarchií. Procesy lze pouštět v posloupnosti, ale obecně lze provádět zpracování i souběžně (paralelně). Při samotné realizaci je evidován průběh programových modulů – čas zahájení a čas ukončení společně s chybovostí procesu, kým byl proces zahájen a s jakým nastavením. Hlavní výběr začíná volbou úlohy a referenčního roku zpracování. Pro definování skupin programovacích modulů a chodů lze předdefinovat šablony, které jsou použitelné pro více úloh.

28 | Spuštění lze parametrizovat daty zahájení s možností vymezení pro celé zpracování, nebo pro jednotlivé skupiny, moduly a chody. Také je možnost spuštění zpracování pomocí Batchu, tedy dávkově s parametrizací posloupnosti, resp. sériovosti zpracování.

29 | Vedle výše uvedených funkcionalit systém umožní rozšíření o vyvolání dalších aplikací, které souvisí s automatizovaným centrálním zpracováním, např. po ukončení necht' dojde k vytvoření exportu do

UF a následného převodu do DWH (aplikace operátorské rozhraní OR pro nahrání do DWH existuje a musí být užíván i v nové technologii) nebo umožní rozšířit menu o další volby.

*Ad 5. Rozhraní, resp. exporty z Centrální DB, resp. importy do jiných subsystémů*

30 | Tak jak aplikace CENTRAL využívá rozhraní pro své inicializační nastavení, resp. přebírá informaci z již výše uvedených zdrojů, tak na druhou stranu zpřístupní své informace pro navazující subsystémy, ať už ve formě naplnění datového rozhraní (tabulek, pohledů), nebo export dat do UF, resp. tak pomocí jiných aplikací provedení importu do navazujících subsystému, jako je SMS-KVALITA (předmětem zadání), DWH (již zavedeno přes operátorské rozhraní, jde jen o zanesení využití do systému), DISEMINACE (předmětem zadání) podle potřeb, resp. optimalizací procesu celého systému GA SMS a GA SIS.

*Ad 6. Správa systému*

31 | Každá akce uskutečněná pomocí aplikace je evidována (kdy, kde, kým a co). Umožní verzování definovaných „objektů“, např. skupin programových chodů. Dále umožní reportování o zadávání změn parametrů, nebo samotného spouštění zpracování, resp. aktuálním průběhu zpracování. Také bude možné sledovat naplněnost jednotlivých databázových uživatelů, resp. odhad volného databázového prostoru. Samozřejmě je možnost správy uživatelů, přidávání rolí jednotlivým uživatelům. Je nutná evidence, resp. systém na správu programových modulů, chodů a jejich skupin.

32 | Tzn.

- zajistit evidenci přihlášení a provedených činností, umožnit dělat reporty z této evidence;
- každý objekt a atribut má evidenci kdy a kdo prováděl změny společně se svými daty platnosti;
- verzování objektu a specifikace platnosti objektu je v souladu s pravidly systému GA SMS.

## 4.2. Uživatelské role

33 | Každá z výše uvedených funkcionalit má možnost nadefinování role s následujícími vlastnostmi:

- role prohlížení – bez možností provádění změn, resp. spouštění;
- role preeditační – provést změny, aktualizace, nové nastavení, ale zatím neschválené, tudíž neplatné, fakticky se jedná o návrh;
- role editační se schválením – provést změny, aktualizace, nové nastavení, které jsou platné a zároveň může editovat jiné návrhy a provést případné schválení návrhu;
- role operátorská – umožní provádět spouštění zpracování s možností vymezení spouštěné úlohy;
- role správcovská – přidělení uživatelů, rolí, ale i doplnění aplikací, atributů (parametrů) pro sledování, resp. nastavování metainformací související s centrálním zpracováním.

## 4.3. Vazby

***Převzetí z procesu přípravy***

- SMS-KLAS, SMS-UKAZ – zdroj číselníků a popisu pro naplnění /čtení metainformací;
- SMS-ULOHY – zdroj pro popis úloh, harmonogramy společně s verzemi SMS-CASOVE RADY – zdroj metadat pro definování výstupního objektu časová řada.
- a běhy zpracování, případných programových chodů a modulů;

***Prostřednictvím SW pro CENTRAL, kdy převážně dojde ke spouštění příslušných programů v rámci realizace zpracování nad centrální databází a rozhraní s ním spojené v SIS CENTRAL:***

- C - I/O - spouštění pro provedení importu/exportu mezi centrální databází a externími datovými zdroji;
- EVID – nahrání VS, ZS a jejich aktualizací pro VSTUP;
- VSTUP – nahrání dat do CENTRAL;
- SMS-VYSTUPY – přebírání informací pro zpracování pomocí GVO;

- GVO – spouštění aplikace pro tvorbu výstupních objektů v SIS CENTRAL;
- DWH – spouštění importu do DWH prostřednictvím tvorby UF z Centrální databáze;
- DM – aplikace pro tvorbu pohledu a dat pro DM.

***Vytvoření rozhraní, resp. pohledů pro plnění uvedených aplikací společně s informacemi přebíraných z procesu přípravy.***

- VDB ;
- DISEMINACE;
- SMS-KVALITA;
- REGISTRY.

#### **4.4. Očekávaná datová architektura**

- 34 | Aplikace bude mít veškerá data uložena v centrální oraclovské databázi v uživateli podle referenčního období úlohu (CSUrrrr) a/nebo centrální databázi SMS pod jedním uživatelem. Veškerá data budou ukládána se všemi potřebnými metadaty a bude zajištěna kontrola jejich konzistence.
- 35 | Metadata budou zadávat interní zaměstnanci ČSÚ z centra i z krajských pracovišť. Každý uživatel bude mít vlastní konto a bude mít právo editovat jen objekty, ke kterým má právo. Všichni uživatelé budou mít právo vidět (jen pro čtení) všechny objekty.

#### **4.5. Rozhraní subsystému**

- 36 | Vstupním rozhraním je rozhraní vytvořené v subsystému SMS-ULOHY a SMS-KLAS, popř. i SMS-UKAZ (view v Oracle DB SMS), SMS-VYSTUPY, SMS-KVALITA, SMS-CASOVE RADY; a rovněž bude přístup nad Centrální databázi (SIS-CENTRAL).
- 37 | Z aplikace SMS-CENTRAL bude vytvořeno výstupní rozhraní (view) pro subsystém SMS-KVALITA, SIS-VSTUP, SIS-CENTRAL, VDB, DISEMINACE, REGISTRY
- 38 | Struktura a obsah rozhraní budou definovány dodatečně při schvalování datového modelu.
- 39 | Aplikace bude využívána pouze interními zaměstnanci, nebude mít definováno výstupní rozhraní pro externí uživatele.