

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. 22

Specifikace bloku Generátor výstupních objektů

Název souboru: RSIS_ZD001P22_GVO.pdf

Počet stran přílohy (bez tohoto krycího listu): 4

Administrace přílohy: Ing. Zdeňka Burešová

Verze ke zveřejnění

1. Generátor výstupních objektů (GVO)

1.1. Základní charakteristika

- 1 | Aplikace GVO slouží k předpřípravě/úpravě částí programů pro generování výstupních objektů a k vlastnímu generování výstupních objektů, tj. výstupních pracovních/kontrolních sestav pro VPO a agregaci údajů popsaných všemi metadaty pro uložení do DWH a VDB.
- 2 | Vstupem jsou metadata vytvořena a uložena v subsystému SMS-VYSTUPY, viz [příloha č. 15](#) („RSIS_ZD001P15_SMS_VYSTUPY“).
- 3 | Výstupem jsou buď vytvořené výstupní sestavy v požadovaném formátu (pracovní/kontrolní tabulky pro VPO), nebo oraclovská tabulka (nebo sada tabulek) s napočtenými agregačními údaji, které jsou popsány všemi potřebnými a očekávanými metadaty včetně případného příznaku primární důvěrnosti, hodnot atributů kvality a příznaku, že údaj je součástí časové řady. Z takto uložených dat a metadat bude tvořen textový soubor ve formátu UF pro nahrání do DWH a jeho prostřednictvím do VDB k prezentaci.
- 4 | Aplikace umožní programátorům případné úpravy předgenerovaných kódů, které vznikly převodem ze SMS-VYSTUPY, a umožní jim testovací generování tabulek nebo agregací.
- 5 | Aplikace umožní automatizované spouštění (přes menu s výběrem) generování tabulek pro VPO a generování agregovaných dat s vytvořením souboru v UF vč. výpočtu (identifikace) a nastavení primární důvěrnosti a případně i atributů kvality dat.
- 6 | Aplikace navazuje na SMS-VYSTUPY, SMS-ULOHY, SMS-UKAZ, SMS-KLAS. Výstupem aplikace jsou vytvořené tabulky v požadovaném formátu nebo soubor ve formátu UF s agregovanými daty, daty a metadaty naplněné tabulky v centrální databázi CSU1.

1.2. Popis současného stavu

- 7 | Pro tvorbu pracovních, kontrolních a příp. publikačních výstupů v .xls (popř. .txt, .dbf nebo HTML) formátu je v současné době používána aplikace GVS, která má 2 části – programátorskou (úprava a ladění programů) a uživatelskou (spouštění generování jednotlivých výstupů přes uživatelské menu). Aplikace je vytvořena ve Forms 6i.
- 8 | Uživatelská část – je vytvořen formulář pro zadání úlohy, období, typu dat (předložená data, dopočtená data, výstupní data), popř. i výběru dle kódů hlavičky, legendy, apod. Po označení/výběru výstupů, které se mají generovat je spuštěno vlastní generování, jehož výsledkem je uložení vygenerovaných výstupů do unixových adresářů nebo do tabulky pro HTML. Jako podklad pro tvorbu seznamu výstupů slouží uživatelský rozdělovník, který vznikne rozbalením programátorského rozdělovníku (období generování je v programátorském rozdělovníku nahrazeno kódem, v uživatelském rozdělovníku je kód období nahrazen konkrétními hodnotami dle periodicity úlohy, obdobně je zacházeno i s volbou za jaké území mají být výstupy generovány).
- 9 | Dále je vytvořena aplikace (Statistické tabulky), která takto uložené výstupy v HTML formátu prezentuje přes webové rozhraní (opět je možnost volby úlohy, období, typu výstupu).
- 10 | Programátorská část – vstupem je XML soubor, vytvořený v aplikaci PMAN, obsahující parametry pro hlavičky (včetně hlavičky v .xls souboru), legendy, elementy legend, výběry, patičky, podmínky pro typ dat a rozdělovník, viz [příloha č. 15](#) („RSIS_ZD001P15_SMS_VYSTUPY“). Jednotlivé části jsou naparsovány do oraclovských tabulek a současně transformovány z formálního popisu PMANu do kódu SQL, popř. PL/SQL. Takto připravené části jsou nabídnuty programátorovi k případným úpravám. Následná část vlastního generování výstupu (po individuálním zadání kombinace hlavičky-legendy-výběru atd. nebo výběrem z formuláře dle připraveného rozdělovníku) vygeneruje z předpřipravených částí PL/SQL program, který po spuštění vytvoří výsledný výstup v požadovaném formátu.
- 11 | Současná aplikace umožňuje jen tvorbu tabulek jako souborů, netvoří údaje popsané metadaty. Nastavení primární důvěrnosti i výpočet hodnot atributů kvality se provádějí v samostatných programech.
- 12 | Aplikace používá při zpracování/ladění kromě elementů legend uložených v tzv. databázi elementů legend a lokálních (pro každou úlohu individuálně definovaných) elementů legend číselníky uložené v tzv. Zrcadle METISu (METIS_MIRROR, jedná se o číselníky SMS-KLAS převedené do

zjednodušené struktury, jak byly vedeny v aplikaci pro číselníky předcházející SMS-KLAS). Tyto číselníky jsou ve zpracovatelské/vývojové databázi zpřístupněny jako public synonyma přes DB link.

- 13 | Vlastníkem každé úlohy (projektu) je právě jeden uživatel-programátor. Vlastnictvím projektu jsou zajištěna práva editovat objekty tohoto projektu.

1.3. Cílový stav

- 14 | Aplikace bude mít dvě samostatné části – uživatelskou a programátorskou. Měla by fungovat na podobném principu jako současná aplikace (kterou tato nová bude nahrazovat a současně bude rozšiřovat její funkcionalitu).
- 15 | Součástí aplikace bude parser metadatového popisu jednotlivých částí zadání, viz [příloha č. 15](#) („RSIS_ZD001P15_SMS_VYSTUPY“), uloženého v SMS-VYSTUPY a překladač metadatových popisů do kódu SQL (PL/SQL), popř. jiného (z důvodu udržitelnosti a použitelnosti systému musí být zvolen jazyk, který jsou schopni zpracovávat současní zaměstnanci ČSÚ). Jednotlivé části by měly zůstat jako samostatné objekty i v této aplikaci.
- 16 | Dále aplikace umožní zásahy programátora do předgenerovaných kódů z důvodu nutných oprav případných nedokonalostí překladu. Současně umožní programátorovi ladící a testovací generování vlastních výstupních objektů do všech požadovaných výstupních formátů (vč. tvorby souboru v UF formátu).
- 17 | Součástí generátoru agregovaných dat musí být i část identifikující a nastavující příznaky primární důvěrnosti podle předem zadaného algoritmu (minimálně pro několik základních typů bude identifikace součástí této aplikace, bude také umožněno k identifikaci a nastavení důvěrnosti volat "externí" program, resp. načítání příznaku z jiné tabulky), nastavení příznaku, že údaj je součástí časové řady, vč. její identifikace, výpočet hodnot požadovaných atributů kvality opět na základě předem daných algoritmů/postupů.
- 18 | Samostatnou částí bude uživatelské menu/formulář pro dávkové spouštění generování výstupních objektů v rámci centrálního zpracování (samostatné spuštění "souborových" výstupů a samostatné spuštění generování agregovaných dat pro DWH a VDB).
- 19 | Použití generování výstupů bude možno nejen jejich volbou přes toto uživatelské menu, ale bude připraveno i řešení umožňující spouštění generování výstupů i voláním procedury (programu apod.) z jiné aplikace přímo zadáním požadovaných parametrů.
- 20 | Výstupem aplikace budou jednak soubory s tabulkami (všechny požadované formáty, zejména .xls) – výstupy pro VPO, jednak soubory agregovaných dat s metadatovým popisem ve formátu UF (publikační výstupy) určené pro uložení do VDB prostřednictvím DWH. Soubor s agregovanými daty bude obsahovat za jednu úlohu/dávku každý statistický údaj (hodnotu) s daným metadatovým popisem jen jednou, tzn. v jednom souboru se nesmí opakovat hodnoty, které by měly stejný popis (metadatový popis je unikátní).
- 21 | Aplikace bude používat číselníky uložené v SMS-KLAS. Při řešení bude rozhodnuto, zda bude obsah SMS-KLAS replikován do zpracovatelské DB nebo bude použito DB linků do databáze SMS. Dále bude používat i metadata (zejména podmínky) pro "složitě kombinované číselníky – elementy legend" zadané v SMS-VYSTUPY, viz [příloha č. 15](#) („RSIS_ZD001P15_SMS_VYSTUPY“).

1.4. Funkční specifikace předmětu plnění – obecné pojednání

- 22 | Výstupem projektu bude návrh a později schválení aplikace, která musí splnit všechny vyjmenované funkcionality společně se zavedenými rolemi a s definovanou návazností a vazbami na další subsystémy SMS a subsystémy SIS. Aplikace bude postavena nad datovým modelem pro evidenci a správu, který je předmětem zadání a který bude schválen zadavatelem. Dále je předmětem zadání specifikování náročnosti na výpočetní techniku a SW požadavky.

1.4.1. Funkční specifikace

- 23 | Aplikace vytvořená na základě této ZD musí plně odpovídat cílovému stavu.
- 24 | Uživatelská část aplikace bude tvořena menu/formulářem umožňujícím v rámci centrálního zpracování kdykoli pustit generování výstupních objektů z programátory předem připravených částí. Volit se bude minimálně úloha, období úlohy (za jednotlivá období nebo libovolné kumulanty), základní území, za která se mají výstupy generovat (ve většině případů je územím celá Česká republika, ale je nutno umožnit generovat i výstupy minimálně za kraje a okresy), typ dat, ze kterých

se mají výstupy generovat (u agregovaných dat vždy za revidovaná-výstupní data, u výstupů pro VPO i za pořízená nebo dopočtená data). Dle takto zadaného předvýběru nabídnout seznam tabulek, které výběru odpovídají. Z tohoto výběru ještě dále umožnit vybírat dle kódu hlavičky, legendy, výběru, resp. priority.

- 25 | Soubory s výstupy (pro VPO i pro DWH) budou ukládány do adresáře na unixu, odkud budou ukládány na standardní místa pro zpřístupnění VPO, resp. pro načtení do DWH. Ve zpracovatelské databázi budou zůstat (volitelně) zachovány tabulky s napočtenými hodnotami, popř. i s popsány metadaty. Stanovení názvů souborů s výstupy pro VPO budou součástí navrženého řešení. Základní návrh datového modelu a názvů tabulek pro ukládání agregovaných dat je uveden v [příloze č. 30](#) ("RSIS_ZD001P30_CSU1_datovy_model").
- 26 | Vstupem do programátorské části aplikace jsou metadata uložená v SMS-VYSTUPY, viz [příloha č. 15](#) ("RSIS_ZD001P15_SMS_VYSTUPY"). Budou se přebírat metadata k výstupům, které jsou označeny v rozdělovníku jako výstupy pro VPO (pracovní tabulky apod.) a VDB. Aplikace provede jejich načtení (naparsování) a převedení do programovacího jazyka této aplikace (SQL nebo jiný – kap. 1.3), vč. platností vstupních objektů. Jednotlivé objekty navržené v SMS-VYSTUPY zůstanou jako samostatné objekty/části i nadále. Při opakovaném načtení objektů ze SMS-VYSTUPY bude provedeno automaticky uložení pod další verzi daného objektu (nebude docházet k přepsání již existujících). Pro více stejných objektů bude jeden označen jako platný, tudíž používaný při vlastním generování výstupů.
- 27 | Při parsování bude aplikace doplňovat i názvy zdrojových tabulek s daty, propojení mezi nimi (joiny) a podmínky výběru dat pro období úlohy a území (zvážit možnost i použití předdefinované tabulky, která by byla plněna zadavatelem).
- 28 | Aplikace umožní takto převedené jednotlivé části programů oprávněnému programátorovi upravovat, ostatní budou moci všechny objekty prohlížet. K editování bude umožněno použití libovolného vybraného (uloženo např. v parametrech uživatele) editoru.
- 29 | Bude umožněno parametrické zadání typu dat do jednoho objektu (centrální zpracování probíhá nad třemi sadami tabulek, samostatně pro pořízená data, pro dopočtená data od začátku roku a pro revidovaná-výstupní data), aby pro různé typy dat nemusely být různé objekty. Dále musí být umožněno načítání dat do výsledného výstupu pomocí více spojených selectů – union all.
- 30 | Další funkcionalitou této části bude možnost spuštění ladících výstupů individuálním zadáním kombinace hlavičky-legendy-výběru atd. Výstupy pro testování bude možno použít i hromadně obdobným způsobem jako v uživatelské části aplikace.
- 31 | Podrobnější upřesnění k některým částem:

a) Hlavičky

- i. bude zajištěno automatické doplnění funkce NVL při parsování;
- ii. agregační funkce SUM bude doplňována automaticky při generování programu (nutno dát pozor na vkládání funkce SUM u podílových proměnných);
- iii. bude možno použít i jiné agregační funkce, které dostanou přednost před SUM.

b) Legendy

- i. protože jsou data popsána jen položkami ZČ a AČ, je nutno pro kombinované číselníky pro propojení s daty používat v KLASu definované vazby na ZČ a AČ;
- ii. nutno doplnit odkazy na rozhraní SMS-KLAS vč. podmínek propojení a podmínek where pro návrat záznamu z číselníku;
- iii. vyřešit odkazy, propojení a podmínky pro specifikace elementů legend, viz [příloha č. 15](#) ("RSIS_ZD001P15_SMS_VYSTUPY"),
- iv. pro VDB – každá hodnota TEP a příslušného EP (položka z číselníku) bude přenesena do metadatového popisu vypočteného údaje.

c) Výběry

- i. pro "omezení" – jen doplnění zadaných podmínek do podmínky where;

- ii. "identifikační" – pro výstup VDB musí být podmínka doplněna do popisu dat jako odkazy na EP pro příslušné TEPy.

1.4.2. Uživatelské role

32 | Aplikace bude mít minimálně tyto uživatelské role:

- role pro programátory – právo zakládání, editování, rušení a verzování objektů, ke kterým má právo;
- roli pro náhledy/prohlížení objektů – právo prohlížet všechny objekty;
- role pro centrální zpracování – uživatel s právem spouštět generování výstupních objektů;
- role pro správce - přidělování přístupových práv, nastavování parametrů aplikace.

33 | V rámci uživatelské role pro programátory aplikace zajistí povolení editovat jednotlivé objekty projektu/úlohy jen uživatelům, kteří jsou jednotlivým projektům přiřazeni, tj. mají právo tento projekt/úlohu editovat.

1.4.3. Vazby na cílový stav existujících funkčních bloků

34 | GVO má velmi těsnou vazbu na SMS-VYSTUPY. Z tohoto subsystému čerpá metadata o požadavcích na výstupy. Dále využívá zprostředkovaně informace ze SMS-ULOH, SMS-UKAZ. Pro metadatové popisy je samozřejmě používán i SMS-KLAS.

35 | GVO je součástí systému CENTRAL (uživatelská část je používána v průběhu centrálního zpracování). Výstupy z GVO jsou prostřednictvím DWH přenášeny do VDB (publikační výstupy).

1.4.4. Očekávaná datová architektura

36 | Aplikace bude mít veškerá data uložena v centrální oraclovské databázi CSU1 pod uživateli jednotlivých zpracovatelských roků. Struktura databáze CSU1 je uvedena v [příloze č. 30](#) („RSIS_ZD001P30_CSU1_datovy_model“)

37 | Programátorskou část GVO budou používat 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.

38 | Uživatelská část bude používána na centrální úrovni.

1.5. Vnitřní rozhraní subsystému

39 | V rámci aplikace bude vytvořeno rozhraní mezi programátorskou a uživatelskou částí (mezi vývojovým a testovacím/reálným prostředím), která zajistí, že jen programátory upravené a schválené objekty úlohy/projektu budou zpřístupněny pro zpracování v uživatelské části. Ladění (programování) výstupů bude probíhat na vývojové databázi. Po jeho ukončení bude umožněno jejich předání/uvolnění pro testování resp. reálné zpracování.

1.6. Vnější rozhraní subsystému

40 | Vstupním rozhraním je rozhraní vytvořené v subsystému SMS-VYSTUPY a SMS-KLAS, popř. i SMS-UKAZ (view v Oracle DB SMS).

41 | Pro přenos dat do DWH je použit univerzální formát, viz [příloha č. 32](#) („RSIS_ZD001P32_UF_popis“) – výstup z generování agregačních dat s metadatovým popisem.

42 | Očekává se zapojení funkčního bloku do režimu přihlašování k aplikacím SSO s použitím interního trezoru identit.

1.7. Dopad na vnější rozhraní SIS

43 | Aplikace bude využívána pouze interními zaměstnanci, nebude mít definováno výstupní rozhraní pro externí uživatele.

1.8. Popis zdrojů ICT

44 | Viz [příloha č. 27](#) („RSIS_ZD001P27_parametry_uziti“).

1.9. Provozní parametry

45 | Viz [příloha č. 27](#) („RSIS_ZD001P27_parametry_uziti“).