

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. 19

Specifikace bloku VSTUP

Název souboru: RSIS_ZD001P19_VSTUP.pdf

Počet stran přílohy (bez tohoto krycího listu): 6

Administrace přílohy: RNDr. René Matýšek

Verze ke zveřejnění

Zadání subsystému VSTUP

1. Modul Dante

1.1. Obecná charakteristika modulu Dante

- 1 | Modul Dante je systém nahrazující a rozvíjející vlastnosti stávajícího systému DataMan. Jde o prostředí pro pořizování, sehrávání a kontrolu vstupních dat formou přímého pořízení zpracovateli ČSÚ, importu dat pořízených respondenty pomocí aktivních PDF formulářů, sehrání dat pořízených respondenty prostřednictvím webového rozhraní, sledování změn zpracovávaných hodnot, četnosti a důvodů těchto změn, prostředí umožňující sledování průběhu zpracování a prohlížení aktuálních pořízených dat i pro VPO, export dat pro centrální zpracování, tvorba kontrolních tabulek atd.
- 2 | Funkcionality modulu Dante vycházejí z požadavků specifikovaných v kapitole 5.9.1 [přílohy č. 78](#) („RSIS_ZD001P78_SP_navrh_architektury“).

1.2. Popis současného stavu

- 3 | Systém DataMan je využíván zpracovateli k pořizování a kontrole dat z výkazů, vytváření kontrolních tabulek a sestav, exportu dat do centrálního zpracování. Přebírá zdrojové soubory s popisem struktury výkazů a programy kontrol generované z PMAN, součástí je i systém EPV, což je programové vybavení určené primárně respondentům k pořizování a provádění kontrol na straně respondenta. Zpracovatelům také umožňuje data pořízená respondenty pomocí EPV importovat do zpracovatelské databáze jak jednotlivě tak v dávkách; data z EPV lze před importem prohlížet. Systém má svůj formalizovaný popis datové struktury, relací do dalších úloh a jazyk pro provádění kontrol nad daty, export dat a vytváření výstupních sestav, pracuje v lokálním (s vlastním datovým formátem) a síťovém (s datovým úložištěm v databázi Oracle) režimu. Má vlastní správu uživatelů. Nástrojem pro podporu šetření v domácnostech je systém Blaise (Corporate License, Usage 25-49 with BCP), předpokládá se jeho použití pro tuto oblast i do budoucna.

1.3. Cílový stav

- 4 | Modul Dante poskytne zpracovatelům prostředí pro pořízení, kontrolu a sehrávání vstupních dat. Dále umožní přímé pořízení výkazů respondenty do vstupní databáze úřadu (Dante - WEB). Spolupracuje s nástroji ostatních subsystémů SIS (např. manipulace a import dat ze souborů PDF). Umožňuje uložení dat jednak v klasické databázi (standardní zpracování), jednak na lokální stanici. Akceptuje a rozlišuje všechny typy kontrol podle SMS-ULOHY. Rozhraní pro editaci zpracovateli a rozhraní pro přímé pořízení výkazů respondenty budou vycházet ze stejného jádra (respondenti budou pracovat nad podmnožinou datové struktury i programů pro kontrolu dat) a zároveň budou, pokud to technologie dovolí, podobná i z hlediska uživatelů (jak vzhledem, tak logikou ovládání). Základem aplikace bude třívrstvá architektura s přístupem prostřednictvím webového rozhraní. Tento přístup by měl umožnit využití aplikace i u respondentů používající jiné operační systémy než OS na bázi Microsoftu (včetně možnosti případného využití například Linuxu v rámci ČSÚ), umožňuje operativní úpravy a upgrade nabízených programů bez nutnosti jejich reinstalace na stanicích uživatelů (jak respondentů, tak zpracovatelů ČSÚ) a klade na minimální nároky na uživatelskou stanici.

1.4. Popis funkcionalit modulu Dante

1.4.1. Funkcionality pro interní zpracovatele

- 5 | Pořízení a editace výkazů (optimální by bylo pořizovat do grafické podoby jakou má výkaz jako tomu je při EPV, úplná podobnost není podmínkou, podobnost ve schematickém rozvržení odpovídající výkazu je však nezbytná) – na pozadí ukládat záznamy změn jednotlivých údajů, včetně času, autora a důvodu změny údaje; možnost zobrazení historie změn údajů.
1. Opravy a doplnění informací o jednotce (vazba na EVID) – bude připraven export změn údajů sledovaných v EVID pro převzetí do systému EVID.
 2. Další typ kontrol – kontroly na celý soubor (resp. stratum) – například kontrola odlehlosti => relace definující příslušnost ke stratu (příslušnou tabulku aktualizuje odbor 23).
 3. Strukturovaný kontrolní chod – s možností filtrování a/nebo třídění podle respondenta, chyby, zpracovatele,...(možnost přidat komentář), filtrování podle kategorie chyb (závažné, informativní, signální,...); navigace k polím porušujícím vazby (otázka inteligentního

- vyhledávání nejvíce „podezřelých“ polí); označení chyb i pomocí notace OŘS.
4. Možnost volby formátu kontrolního chodu (.txt, .xml, .xls, .pdf, .odt,...).
 5. Příjem dat od jednotek (import dat z databáze navázané na „WEBové rozhraní“ + import editovatelných .pdf + import (po vytěžení) naskenovaných výkazů); importy, dávkové převody v kompetenci technologů; záznam historie akcí souvisejících s importem jak dávkovým tak importem jednotlivých vět.
 6. Minimalizace zátěže zpracovatele při akcích mimo vlastní pořizování.
 7. Kontroly s přiměřenou dobou odezvy (odezva kontroly jednotlivého formuláře v řádu sekund).
 8. Způsob práce s porušenými vazbami – tyto informace ukládat do samostatné tabulky se strukturou obsahující mj. IČO, číslo vazby, stav potlačení, komentář, - tabulku bude možné jakkoliv analyzovat.
 9. Přístup k administrativním zdrojům dat, jejich propojení s pořizovací úlohou a současné zobrazení příslušných údajů pro potřeby kontrol (opticky zpracovatelkou nebo i kontrolní vazbou), relace na další potřebné datové zdroje (minulá období apod.) – viz i bod 13.
 10. Možnost vyvolávání číselníků a doplnění zvoleného kódu do zpracování.
 11. Možnost uživatelského nastavení (např. rozlišení obrazovky, rozložení obrazovky, mód práce, hot keys, myš, menu, propojení (a způsob zobrazení) na administrativní data v centru); optimalizace oken dle potřeby uživatele; možnost maker.
 12. Parametrické spouštění úlohy např. s IČO (surfování po úlohách i jednotkách).
 13. Možnost práce off-line; na rozdíl od stejného požadavku pro respondenty ale s širšími možnostmi (kompletní číselníky, neomezená sada zpracovávaných výkazů, možnost následné synchronizace dat s centrálním úložištěm).
 14. Jednoduchá a přímá cesta k obnovení instalací (pro uživatele off-line zpracování) a dat po havárii techniky (repliky instalací a dat ukládané na serveru v okamžiku synchronizace).
 15. Možnost paralelně otevřených editačních oken (jedním uživatelem) pracujících nad stejným záznamem (souběžný aktivní pohled na více míst výkazu nebo relacemi svázaná data (AZD, předchozí období aj.) – využití i více monitorů).
 16. Zobrazení odpovídající mutace (respektive složení výkazu) pro konkrétní jednotku; spouštění odpovídajícího bloku kontrol.
 17. Přístup k provozní dokumentaci (odkazem nebo uložením statického dokumentu).

1.4.2. Funkcionality pro technology

1. Příjem VS a jejich aktualizace – relace EVID/RES, informace o připravené aktualizaci, možnost zvolit čas spuštění aktualizace VS (odložená aktualizace); propojení do EVIDu na adresní a kontaktní údaje.
2. Dávkové kontrolní chody (na úrovni serveru).
3. Exporty dat do centrálního zpracování – replikace dat uložených na krajských serverech (v počáteční fázi jen průřezových úloh s možností rozšíření na všechny úlohy (podle míry zatížení centrální výpočetní techniky při replikaci)).
4. Dávkové převody EPV + PDF + off-line tazatelé – importy do zpracování.
5. Možnost uživatelské definice ad hoc exportů (např. do KD).
6. Integrovaná možnost archivace (ruční x automatické v definovaných termínech) - konkrétní úlohy, skupiny úloh, úlohy a tabulek s nimiž je v relaci, kompletního zpracovatelského roku.
7. Integrovaná možnost vlastního systému záloh; stačí komplexní zálohy s volbou zpracovatelského roku do komprimovaného tvaru – není nutné používat inkrementální zálohy.
8. Vlastní generátor sestav s vazbou na výstupy definované v SMS-VYSTUPY (převod těchto požadavků prostřednictvím můstku VIP<->ÚOŘS).
9. Systém bude poskytovat předdefinovanou sadu matematických, statistických , logických

funkcí a funkce pro práci se základními typy proměnných (zejména řetězci, čísla, datem a časem, ...).

10. Přístup k provozní dokumentaci (odkazem nebo uložením statického dokumentu).

1.4.3. Funkcionality pro respondenty (Dante – WEB)

1. Pořízení a editace *dat* z výkazů v datovém úložišti ČSÚ s možností sledování aktuálního stavu pořízení *dat* respondenty.
2. Možnost opakovaného přístupu k již uloženým výkazům (problém autentizace k úloze).
3. Možnost práce off-line.
4. Jednoduché odesílání *dat*.
5. Přístup a pořízení přes obecné webové rozhraní (nezávislost na operačním systému).
6. Respondent se při přihlášení dostane k poslední potvrzené verzi hodnot (nikoliv například k hodnotám upraveným podle požadavků VPO) – data pořízená respondenty prostřednictvím on-line webového rozhraní se budou držet v samostatné databázi.
7. Možnost přístupu i k ostatním přiřazeným úlohám (včetně předchozích období), možnost respondentů definovat další uživatele s právy přístupu ke konkrétním úlohám (delegování práce na jiné zaměstnance subjektu nebo na pracovníky externí firmy).
8. Bude přístupná uživatelská dokumentace k aplikaci a úloze.
9. Možnost exportu a následného importu celého nebo části výkazu do/z editovatelného formátu (.xls, .xml) pro možnost nezávislého zpracování více pracovníky na straně respondenta nebo pro možnost naplnění *dat* prostřednictvím informačního systému používaného respondentem.
10. V rámci pořizování – možnost zobrazit vysvětlivky k polím při najetí myši na konkrétní pole.
11. Označení chyb ve výpisu pomocí OŘS notace.
12. Zobrazení odpovídající mutace (respektive složení výkazu) pro konkrétní jednotku; spouštění odpovídajícího bloku kontrol.

1.4.4. Funkcionality pro pracovníky VPO

1. On-line sledování průběhu zpracování sleduje se v systému EVID.
2. Sledování kvality pořízených *dat* na centrální replice (suma krajských db – viz požadavky z hlediska technologií – položka exporty do centrálního zpracování – replikace se zpožděním max. 24 hodin + na pozadí dle možnosti a kapacity linek).
3. Možnost přímého pohledu na kontrolní sestavy (dnes zajištěno php skripty nad pořizovací databází) – pohled prostřednictvím „můstku“ (možnost zvážit jako součást obecného GVO (příloha 22)).
4. Označení kontrol výčet a typ pro PDF, WEB, CAPI, ZPRAC (zavést a používat číselník typů a kategorií kontrol).
5. Možnost analyzovat data získaná na základě historie jednotlivých údajů (bod 1.4.1/1 – důvody změn, četnost => zpětná vazba pro špatně pochopené popisky a vysvětlivky atd).
6. Systém bude poskytovat předdefinovanou sadu matematických, statistických, logických funkcí a funkce pro práci se základními typy proměnných (zejména řetězci, čísla, datem a časem, ...).
7. Přístup k provozní dokumentaci (odkazem nebo uložením statického dokumentu).

1.4.5. Funkcionality pro programátory

1. Nová úloha se objeví pouze „zařazením“ naprogramované úlohy do seznamu pro pořizovatele (resp. technology)
2. Programátoři budou připravovat aplikační vrstvu => ověřování bude probíhat nad jedinou verzí, nenastanou komplikace s distribucí jednotlivých verzí programového vybavení.
3. Vlastní editační prostředí není třeba – lze použít buď obecný editor (typu PSPad nebo v případě použití prostředí typu Delphi for PHP editor tohoto prostředí).

4. Generování strukturované provozní dokumentace k úloze (seznamy položek, typ a rozsah položek, kontrolních vazeb).
5. Vývoj a ladění programů bude probíhat nad testovací/vývojovou databází, po odsouhlasení uživateli bude hotová sada programů exportována do produkční databáze.
6. Systém bude poskytovat předdefinovanou sadu matematických, statistických, logických funkcí a funkce pro práci se základními typy proměnných (zejména řetězci, čísla, datum a časem, ...).

1.4.6. Vazby na cílový stav existujících funkčních bloků

- 6 | Datová vazba se systémem EVID na úrovni volání dat z EVID pro aktuální větu a sledování stavu formuláře a zpracování.

1.4.7. Datová architektura

- 7 | Data budou uložena ve vstupních databázích (krajské databáze, předřazené pro web přístup respondentů a souhrnné před centrálem) v tabulkách odpovídajících struktuře výkazu - podle oddílu (tabulka s identifikátorem a dalšími vybranými informacemi o jednotce, tabulky pro jednotlivé oddíly – buď 1:1 nebo 1:N – volba relace bude podle specifického atributu v SMS-ULOHY).
- 8 | Data u respondentů bude možné ukládat a zpracovávat i na lokálních stanicích.
- 9 | V těchto db bude tabulka definující jednoznačnou relaci mezi VIP a položkou (úloha-oddíl-řádek-sloupec), tato tabulka bude sloužit jednak jako podklad pro export do UF, jednak jako můstek pro přístup k datům prostřednictvím tabulek vytvořených v SMS-VYSTUPY.
- 10 | Významová identifikace může být složena z více položek formuláře.
- 11 | Optimalizace rozložení transakcí, operací na centrálním serveru, lokálním serveru a lokální stanici z hlediska četnosti, počtu uživatelů, objemu dat.
- 12 | Systém nemá využívat specifické vlastnosti konkrétního typu databáze (pouze v případě zásadního zlepšení užitečných vlastností a po dohodě se zadavatelem).
- 13 | Pro webové rozhraní bude podporován protokol HTML5 nebo vyšší (zejména v případě potřeby lokálního ukládání dat).

1.4.8. Vnitřní rozhraní subsystému

- 14 | Funkční blok DANTE musí mít otevřené relace s funkčním blokem EVID. Popis bloku EVID viz [příloha č. 35](#) („RSIS_ZD001P35_TP_EVID“).

1.5. Vnější rozhraní subsystému

- 15 | Vazba na modul ISAAC subsystému PROGRAM – Dante přebírá SQL skripty pro vytvoření sady tabulek pro danou úlohu, včetně skriptu pro přidělování přístupových práv jednotlivým rolím, SQL skripty pro kontroly na úrovni serveru, dvě verze skriptu v JS pro interaktivní kontroly na úrovni stanice a pro PDF formulář, seznam číselníků pro danou úlohu a datum platnosti, popřípadě další informace. Takto získané skripty bude možné doladovat a doplňovat.
- 16 | Výstupem ze systému jsou data pro centrální zpracování ve formátu UF (a to jak statistické údaje, tak i jejich atributy, informace o kvalitě a historii údajů). Podkladem bude tabulka pro přiřazení popisu dat „úloha-oddíl-list-řádek-sloupec“ předpokládaného pro uložení dat v systému DANTE a VIP (respektive identifikátoru UF). V případě nutnosti lze tuto tabulku použít i jako interface pro použití návrhů sestav vytvořených v generátoru výstupních sestav v části CENTRAL (zde nad formátem uložení dat založeným nad VIP). Podrobný popis UF viz [příloha č. 32](#) („RSIS_ZD001P32_UF_popis“) a [příloha č. 36](#) („RSIS_ZD001P36_DWH_datovy_model“) a související dokumenty.
- 17 | Funkční blok EVID subsystému VSTUP, který není součástí tohoto předmětu plnění, musí mít otevřené relace k subsystému, který obsahuje popis objektů, z nichž se skládají TP (SMS-ULOHY – viz [příloha č. 26](#) („RSIS_ZD001P26_prostredi ICT“) a [příloha č. 27](#) („RSIS_ZD001P27_parametry_uziti SIS“)). Z něj získává seznam číselníků pro danou úlohu a datum platnosti, podklady pro vytvoření harmonogramu přípravy a zpracování, podklady pro popis zpravodajské povinnosti.

18 | Vstupy jsou také přijímány z generátoru grafické podoby výkazu, tj. subsystém PŘÍPRAVA, modul ENRICO SMS.VYSTUPY pro tvorbu maket Kontrolních zpracovatelských výstupů.

1.6. Dopad na vnější rozhraní SIS

19 | Podle současného předpokladu bude vnější rozhraní využíváno obecně např. pro správu a autentizaci uživatelů, správu aplikací a databází.

20 | Funkční blok bude zapojen do režimu přihlašování k aplikacím SSO s použitím interního trezoru identit.

21 | Funkční blok musí spolupracovat s poštovním úřadem Novell Groupwise, elektronickou spisovou službou ABC SUITE, portálem Intranet a Internet.

1.7. Popis zdrojů ICT

1.7.1. Obecné

22 | Viz [příloha č. 26](#) („RSIS_ZD001P26_prostredi ICT“).

1.7.2. Specifické

23 | Specifické zdroje nejsou vyžadovány.

24 | Modul Dante bude používán ve vývojovém, testovacím a produkčním prostředí.

1.8. Provozní parametry

25 | Jsou uvedeny v příloze parametry užití.

26 | **Požadavky na výkonnost:** Modul musí být schopen pracovat na stávající výpočetní technice úřadu. Systém musí být postaven tak, aby s rezervou zvládl špičky zpracování

27 | Předpokládaný počet konkurenčních uživatelů pro systém na úrovni OSZ je cca 100 (v rámci jednotlivých krajů); požaduje se přechod mezi editačními poli formuláře v době řádu jednotek ms (včetně provedení interaktivních kontrol souvisejících s opouštěným polem), odezva na kontrolu aktuálně editovaného výkazu v řádu jednotek sekund (při cca 150 kontrolách na výkazu), doba kontroly jednoho výkazu na straně serveru řádově desítky sekund (pro současně nejkomplikovanější úlohy typu P5-01 na stávajícím HW do cca 120 minut při kompletní kontrole úlohy s 10000 jednotkami ve VS); modul Dante bude pracovat nad databázemi lokalizovanými v místě jednotlivých pracovišť.

28 | Předpokládaný maximální počet konkurenčních uživatelů webového rozhraní pro respondenty je řádově 10000 přístupů během pracovní doby jednoho dne (v případě kolize termínů odevzdání více měsíčních a čtvrtletních výkazů). Je požadována certifikace ISVS + šifrování přenosů zejména pro přístup respondentů k externímu webovému rozhraní modulu Dante.

1.8.1. Počty uživatelů dle rolí

29 | Viz [příloha č. 27](#) („RSIS_ZD001P27_parametry_uziti_SIS“).

1.8.2. Objem dat, počty položek dle datové architektury

30 | Viz [příloha č. 27](#) („RSIS_ZD001P27_parametry_uziti_SIS“).

1.8.3. Dostupnost a odezva podpory

31 | Viz [příloha č. 27](#) („RSIS_ZD001P27_parametry_uziti_SIS“).

2. Technické požadavky

32 | Pro web rozhraní bude použit protokol HTML 5 a vyšší.

3. Bezpečnost

33 | Všechny zadávané moduly musí, není-li výslovně uvedeno jinak, splňovat následující:

1. použité programovací prostředky musí být dostatečně obecné a rozšířené (podpora, jednodušší řešení problémů, snazší případná výměna nebo doplnění zaměstnanců, jednodušší zadání vývoje části systému mimo úřad, ...);

2. moduly budou „open-source“ pro maximální zjednodušení případných vlastních doplnění nebo zadání části vývojových prací na tomto systému externímu dodavateli;
3. jednoznačně musí být zachovány „okraje“ komponent, respektive jednotlivých modulů – rozhraní musí být striktně definována a popsána, není přípustná jakákoliv nedokumentovaná vazba a funkce;
4. správa přístupu uživatelů bude na úrovni úloh; při použití SSO&TI je nezbytné zajistit lokální fungování této autentizace pro případ výpadku linek do centra;
5. vzhledem k dalšímu provozu a rozvoji dodaného SW a nutnosti převzetí zdrojových kódů zaměstnanci ČSÚ je nutné individuálně vyvinutý SW připravit v prostředí základních programovacích jazyků, které je pro zaměstnance standardem, tj. SQL, PL/SQL, C++, C#, Delphi, Java, JavaScript, PHP, PERL, Python a další dle seznamu standardních SW produktů;
6. moduly systému budou založeny na třívrstvé architektuře s přístupem prostřednictvím webového rozhraní;
7. navržené řešení, nástroje, aplikace nesmí vyžadovat konkrétní HW řešení a musí být přenositelné i na jinou platformu HW a OS;
8. podpora prohlížečů IE, FF, Chrom, Opera ...;
9. ve fázi návrhu řešení budou identifikovány sdílené komponenty řešení v rámci celé dodávky;
10. návrh řešení musí minimalizovat vícenásobné uložení dat (netýká se záloh a archivace);
11. identifikace objektů (metainformací, položek, VIP) musí být jednoznačná v rámci celého SIS;
12. tiskové výstupy, pokud není specifikováno jinak, se očekávají pro textové dokumenty (např. TP) ve formátu RTF a PDF, výstupy strukturované do tabulek ve formátu XML, HTML, PDF.