

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. 16

Specifikace bloku časové řady

Název souboru: RSIS_ZD001P16_CASOVE_RADY.pdf

Počet stran přílohy (bez tohoto krycího listu): 6

Administrace přílohy: Ing. Pavel Skotnica

Verze ke zveřejnění

Zadání subsystému SMS-CASOVE RADY

1. Základní charakteristika

- 1 | Subsystém SMS-CASOVE RADY umožňuje zadávat metadata potřebná pro vedení srovnatelných časových řad.
- 2 | Pojmy jsou přebrány a upraveny z „Úvodní studie k zajištění srovnatelných časových řad“, verze 1.2 z 03/06/2008, kterou vypracoval tým SMS-RADY.
- a) **Časová řada** (ČŘ) je množina pozorování určitého (zejména kvantitativního) ukazatele uspořádaná podle přirozené časové posloupnosti.
 - b) **Kritéria** srovnatelnosti. Pozorování v časové řadě jsou **srovnatelná**, jsou-li shodná všechna vymezení statistické proměnné:
 - b1) věcné (obsahové): vymezení statistické proměnné (STAPRO) a věcný obsah ukazatele, např. způsob, jakým je zjišťován obrat z podnikového účetnictví;
 - b2) prostorové: nejen vymezení území, k němuž se vztahuje, ale i druh statistické jednotky (podnik, místní jednotka, domácnost atd.);
 - b3) časové: vymezení ukazatele časově, pro určitý okamžik (např. 15. den měsíce) nebo za jaký časový interval (např. za rozpočtový rok), tj. zda je časová řada stavová či toková.
- 3 | Srovnatelnost časové řady podstatně ovlivňuje metodika zjišťování ukazatele (statistické zjišťování/administrativní data, plošné/výběrové šetření, vymezení základního souboru atd.).
- 4 | Pokud není srovnatelnost časové řady dle výše uvedených kritérií plně zajištěna, je třeba provést přepočty časové řady k zajištění srovnatelnosti, poskytnout dostatečné informace o charakteru a velikosti případných rozdílů, aby bylo možno zohlednit neúplnou srovnatelnost řad při analytické činnosti.
- c) **Typologie časových řad**. Z praktických důvodů je nutné rozlišovat, zda v celé časové řadě je shodné vymezení statistické proměnné (STAPRO)
 - c1) „pravé časové řady“ – shodné vymezení statistické proměnné (STAPRO) pro všechna období časové řady,
 - c2) „nepravé časové řady“ – časová řada je složena z několika na sebe navazujících časových řad s mírně odlišným vymezením statistické proměnné.
- 5 | Pozn.: Naprostá většina časových řad jsou „pravé časové řady“. Avšak někde je nutné vést „nepravé časové řady“, například počet registrovaných uchazečů na úřadě práce. Legislativa se často mění, nicméně má smysl sledovat tuto časovou řadu.

2. Popis současného stavu

- 6 | V současnosti se při sestavování časových řad postupuje více způsoby:
- a) Časové řady jsou vedeny ručně, nová resp. revidovaná pozorování časové řady jsou přidávána ručně. Jedná se zejména o časové řady přebírané z externích zdrojů. Mohou se provádět dodatečné přepočty k zajištění srovnatelnosti časové řady.
 - b) Časové řady se generují jako výstup ze zdrojových zpracovatelských databází:
 - b1) Celá časová řada pochází z jedné statistické úlohy, mohou se provádět přepočty k zajištění srovnatelnosti řady (zejména přepočet na shodné věcné a prostorové vymezení, zpětné přepočty při změně klasifikace);
 - b2) Časová řada se skládá z více datových zdrojů - jednotlivé úseky časové řady čerpají zdrojová data z různých úloh a VIP, mohou se provádět přepočty k zajištění srovnatelnosti časové řady (včetně zpětných přepočtů při změně klasifikace).
 - c) Časová řada je vedena ručně, část časové řady pochází z již ukončených statických úloh, nejnovější data se generují ze zpracování platné statistické úlohy. Přepočty se provádí většinou pomocí můstků.

7 | Metodické popisy o datových zdrojích, srovnatelnosti časových řad a přepočtech zajišťující jejich srovnatelnost se sestavují ad hoc. Nemají standardní podobu, jejich obsah, detail a úplnost se liší. Srovnatelnost časové řady se nezřídka posuzuje arbitrárně.

8 | V současnosti se časové řady vedou paralelně více způsoby: i) na webu ČSÚ jako *.xls soubory, ii) v ad hoc publikacích a zejména iii) ve Veřejné databázi (VDB), kde kromě časových řad jsou uložena rozsáhlá metadata.

3. Cílový stav

3.1. Popis

9 | Subsystém SMS-CASOVE RADY umožní založit, definovat, zadávat, editovat, schvalovat metadata potřebná pro popis výstupního objektu typu časových řad. Tato metadata lze připojit k příslušným statistickým údajům/hodnotám, které vznikají v průběhu centrálního zpracování a vytváření výstupních objektů a které jsou identifikovány – úlohou, VIP, obdobím úlohy, verzí a během úlohy, referenčním obdobím, statistickým objektem, statistickým doplňkovým objektem..., veškerými identifikacemi, které jsou potřebné pro centrální zpracování, vyhodnocení kvality a diseminaci.

10 | Subsystém by měl mít následující vlastnosti a zajišťovat tyto funkce:

- a) Je možné definovat požadavky na generování výstupního objektu „časová řada“:
 - a1) identifikace časové řady – po vymezení statistické proměnné (resp. více STAPRO, viz a5) se časová řada identifikuje v číselníku KLAS, tyto časové řady se dále sdružují do tematických celků (v souladu s Katalogem produktů – vazba na blok DISEMINACE), které jsou také obsaženy v číselníku KLAS;
 - a2) věcné vymezení (statistická proměnná - STAPRO);
 - a3) časové vymezení (periodicita – víceleté, roční, čtvrtletní, měsíční, týdenní ...);
 - a4) začátek a konec časové řady; (časová řada může být otevřená, tj. konec je vždy aktuální období);
 - a5) je nutné umožnit zadávat tzv. „nepravé časové řady“ (viz 1.c2), které se skládají z více souvislých úseků „pravých časových řad“, které se mírně liší v definici statistické proměnné.
- b) Jednotlivé úseky časové řady se mohou naplňovat z různých datových zdrojů, pro každý takový úsek stanoví:
 - b1) počátek a konec tohoto úseku;
 - b2) statistická úloha společně s dalšími identifikátory statistického údaje/hodnoty (verze/běh/období úlohy/ VIP (otevřené – může být více různých VIP anebo jen jedna pro pravou časovou řadu) a případně statistickým objektem, doplňkovým objektem, základním obdobím, z nichž se v daném úseku bude časová řada naplňovat.

11 | Pozn. Může nastat situace, kdy se pro dané období bude časová řada postupně naplňovat z různých zdrojů. Pak jsou vždy stanoveny priority, z kterých zdrojů údaj naplnit. Např. předběžné údaje čtvrtletní úlohy se naplňují z údajů měsíční úlohy a administrativních dat, údaje semidefinitivní z čtvrtletního výkaznictví a AZD a definitivní údaje z čtvrtletního výkaznictví, AZD a ročního výkaznictví.

- b3) metody, kterými se zabezpečí srovnatelnost celé časové řady, tj. jakým způsobem se na sebe jednotlivé úseky naváží (např. využití můstků) a přepočtou.
- c) Výstupní objekt časová řada bude verzovaný, bude se vést jeho historie a bude možno měnit platnost údajů.
- d) Subsystém je otevřený na doplňování atributů pro metainformací popis časové řady.
- e) Samotné přiřazení výstupu do časové řady je prováděno v bloku SMS-VYSTUPY pro údaje, které ještě nebyly vytvořeny a očekává se jejich vytvoření ve zpracování (viz f.), nebo je vytvářeno přes SMS-CASOVE RADY již nad existujícími statistickými údaji uložené

v centrální databázi (viz g.) s možností zachování, přiřazení následných (budoucích) údajů při zachování příslušné identifikace.

- f) V rámci zpracování SIS-CENTRAL dojde k faktickému přiřazení identifikace k časové řadě v případě, je-li v SMS-VYSTUPY definováno přiřazení k časové řadě.
- g) Subsystém umožní vybrat z již zpracovaných dat uložených v centrální databázi příslušnost do časové řady a poté subsystém SMS-CASOVE RADY zajistí identifikaci v centrální databázi.
- h) V případě postupného vytváření časové řady z různých úloh a v různém čase je možné stanovit priority úloh a nebude-li priorita zadána, subsystém DISEMINACE zajistí pro prezentaci prioritizaci dle platnosti dat v duchu „čím novější, tím lepší“.
- i) Subsystém umožní kontrolu souladu definic k již přiřazeným údajům k časovým řadám. Jelikož přiřazení k hodnotám je definovaná v subsystému SMS-VYSTUPY, ale i v subsystému CASOVE RADY a může dojít k redefinici, resp. k aktualizaci, je třeba v příslušné době provést kontrolu konzistence definovaných metadat s faktickou identifikací pro zvolené atributy, např. soulad počátku, STAPRO, zdrojové úlohy, ...

3.2. Technologie

- 12 | V rámci subsystému VYSTUPY s možností navázání na Katalog produktů (DISEMINACE, viz [příloha č. 24](#) („RSIS_ZD001P24_DISEMINACE“) se založí objekt časová řada pro zanesení metainformací - věcně, časově a prostorově se vymezí (STAPRO), určí se její počátek a případný konec. Tento výstupní objekt se uloží, založí se nová hodnota v číselníku časových řad. Dále se v rámci subsystému VYSTUPY doplní informace, z které statistické úlohy a z kterých VIP se naplní jednotlivé úseky časové řady, resp. z jakých již zpracovaných úloh lze provést přiřazení. Také se stanoví, z jakých datových zdrojů a s jakou prioritizací se řada tvoří (např. z jakých úloh/verzí zpracování - úroveň standardní, 1. roční revize, 2. roční revize). Tyto informace se využijí v subsystému DISEMINACE, resp. na kontrolu již uložených dat, resp. na přiřazení již uložených dat k příslušné časové řadě. DISEMINACE provede převzetí správných výstupních údajů. Objekt časové řady se bude dále doplňovat dalšími metadaty (např. o srovnatelnosti časové řady, o použitých metodách přepočtu časové řady).
- 13 | Z uvedeného vyplývá, že je několik možností vytváření hodnotové časové řady, jejíž metadatový popis je založen v subsystému SMS-CASOVE RADY:
 - A. Přiřazení již existujících dat do časové řady a zanechat identifikaci přiřazení pro statistické údaje/hodnoty uložené v centrální databázi.
 - B. V SMS-VYSTUPY přiřadit připravované výstupní objekty k časovým řadám a při vytvoření výstupu v SIS-CENTRAL dojde k přiřazení.
 - C. Dojde k vydefinování časové řady na základě poskládání z již definovaných časových řad.

4. Funkční specifikace předmětu

Obecné pojednání

- 14 | Tato kapitola specifikuje požadavky na funkcionalitu subsystému společně se zavedenými rolmi a s definovanou návazností a vazbami na další subsystémy SMS a subsystémy SIS. Funkcionalitu by měla zabezpečit navržená a později schválená aplikace, která by měla být postavena nad databázemi s datovým modelem, který je také předmětem zadání. Dále je předmětem zadání specifikovat náročnost na výpočetní techniku a SW požadavky.

4.1. Funkcionalita

4.1.1. Aplikace se chová podle role, která je k uživateli přiřazena

4.1.2. Subsystém je otevřený, lze doplňovat a odebírat další atributy objektu časové řady.

4.1.3. Způsoby zadávání vstupních dat:

- a) Výčtem hodnot z číselníků;
- b) Manuálně – text, číselné hodnoty, datum;
- c) Import dokumentů – popisy srovnatelnosti časových řad a metodiky přepočtu časových řad.

- 4.1.4. Export popisů časových řad v *.pdf, *.rtf, *.xls, *.xlsx souhrnné reporty metadat o časových řadách.
- 4.1.5. Lze ukládat a používat předvolené (dědit) popisy časových řad pro další časové řady. Rovněž by mělo jít vytvářet popisy časové řady pro více časových řad současně.
- 4.1.6. Provádět kontrolu, resp. porovnání vydefinovaných atributů v SMS-CASOVE RADY s metadatovými informacemi údajů přiřazených v SIS-CENTRAL (v centrální databázi) a výsledky kontroly reportovat.
- 4.1.7. Dodatečně přiřadit údaje k časové řadě v SIS-CENTRAL (v centrální databázi)
- 4.1.8. Existence rozhraní pro subsystém DISEMINACE
- 4.1.9. Odlišovat prostředí návrhu/editace a prostředí schválených se ctěním rolí uživatelů, ke schválené verze může dojít jen po odsouhlasení uživatelů s rolí STATISTIK, tak METODIK.
- 4.1.10. Evidence přihlášení a provedených činností, umožnit dělat reporty z této evidence.
- 4.1.11. Každý objekt a atribut má evidenci, kdy a kdo prováděl změnu a také má své datumy platnosti.
- 4.1.12. Subsystém umožňuje verzování objektu časových řad, specifikaci platnosti objektu v duchu GA SMS.
- 4.1.13. Provést export z centrální databáze pro zvolené časové řady s ohledem na volbu určení dle datumu pohledu, resp. s možností zvolené maximální parametrizace identifikace údajů (úloha/verze/běh/..)

4.2. Uživatelské role

15 | Pro každou z výše uvedených funkcionalit lze nadefinovat role s následujícími vlastnostmi:

- 4.2.1. **Role prohlížení – bez možnosti provádění změn, právo prohlížet si vybrané atributy časových řad, editovat a exportovat výstupy a souhrnné reporty o metadatech časových řad.**

4.2.2. Role správcovská (administrátor)

- a) obvykle pracovník útvaru metodiky, spravuje číselníky vztahující se k časovým řadám;
- b) označuje, které atributy časové řady jsou veřejné, zakládá atributy nové;
- c) spravuje přístupová práva pro další skupiny uživatelů, určuje, které položky mohou naplňovat;
- d) přiděluje role uživatelům;
- e) hlavní úkolem je zajištění správné konfigurace subsystému;
- f) zahrnuje rovněž roli prohlížení.

4.2.3. Role statistik (role editační se schválením)

- a) obvykle pracovník věcně příslušného odboru (VPO), je to gestor příslušné časové řady. Definuje časovou řadu, naplňuje atributy časové řady, zejména definice STAPRO/VIP, datové zdroje. Má právo zakládání, aktualizace, editování, rušení objektů, změna platností, schvalování a verzování objektů, ke kterým má právo.
- b) schvaluje veškeré změny konfigurace objektů;
- c) určuje, které atributy časové řady jsou veřejné;
- d) uvolňuje časovou řadu pro interní/externí využití;
- e) může vznášet požadavky na doplnění metainformací o časových řadách;
- f) zahrnuje rovněž roli prohlížení.

4.2.4. Role Metodik (role editační se schválením)

- a) výkonná role, obvykle pracovník útvaru metodiky, ale také diseminace. Má právo zakládání, editování, rušení objektů, změna platností, schvalování a verzování objektů, ke kterým má právo. Zejména naplňuje atributy časové řady, vztahující se zejména k metodice zajištění srovnatelnosti časových řad.
- b) Může vznášet požadavky na vedení časových řad a doplňování atributů časových řad.
- c) Zahrnuje roli prohlížečí. Navíc může uvolňovat tyto metainformace pro externí uživatele, které však budou uloženy jako zrcadlo mimo samotný subsystém.

16 | Pozn.: Pracovník útvaru metodiky může být i gestor časové řady, zejména u časových řad historických, přebíraných z jiných zdrojů (průměrná teplota v ČR). Nicméně v tomto případě má roli Uživatel-Statistik.

4.3. Vazby

4.3.1. Vazby subsystému SMS-CASOVE RADY a SMS-KLAS

17 | Subsystém SMS-CASOVE RADY využívá standardní klasifikace a číselníky z SMS-KLAS při definování a identifikaci časové řady.

4.3.2. Vazby subsystému SMS-CASOVE RADY a SMS-UKAZ

18 | Subsystém SMS-CASOVE RADY využívá definice statistické proměnné (STAPRO) při definování časové řady.

4.3.3. Vazby subsystému SMS-CASOVE RADY a SMS-VYSTUPY

- a) V subsystému SMS-VYSTUPY se provádí přiřazení výstupu (údajů) k časové řadě.
- b) Dále se v subsystému SMS-VYSTUPY určuje, z kterých datových zdrojů se naplňuje časová řada: i) počátek a konec úseku, ii) STAPRO a datový zdroj – statistická úloha.

4.3.4. Vazby subsystému SMS-CASOVE RADY a SMS-POZADAVKY

19 | V subsystému SMS-POZADAVKY se evidují požadavky na vedení časových řad plynoucí především z legislativy (periodicita, počátek časové řady). Tyto informace se dále rozpracovávají v subsystému SMS-ULOHY. Subsystémy spolu přímo nekomunikují, veškeré informace se přebírají zprostředkovaně přes subsystém SMS-ULOHY.

4.3.5. Vazby subsystému SMS-CASOVE RADY a SMS-KVALITA

- a) Ze subsystému SMS-KVALITA se přebírají atributy kvality vztahující se k jednotlivým obdobím časové řady i celé časové řadě.
- b) Do subsystému SMS-KVALITA vstupují požadavky na atributy kvality specifické pro časové řady, jako např. variační koeficient meziročních temp růstu. Přebírají se informace o srovnatelnosti časových řad (zda se ČR skládá z více datových zdrojů, informace o metodice přepočtu časových řad, srovnatelnost časových řad).

4.3.6. Vazby subsystému SMS-CASOVE RADY a DISEMINACE

20 | Subsystém DISEMINACE přebírá potřebná metadata vztahující se k časové řadě, zejména vymezení časové řady, počátek a konec časové řady, informace o datových zdrojích, srovnatelnosti atd.

21 | Některé informace vcházejí do subsystému DISEMINACE zprostředkovaně přes subsystém SMS-KVALITA, kde se informace o časové řadě sumarizují v rámci dimenze kvality „srovnatelnost“.

4.3.7. Vazby subsystému SMS-CASOVE RADY a SIS-VSTUP

22 | Mezi subsystémy žádná přímá vazba není.

4.3.8. Vazby subsystému SMS-CASOVE RADY a SIS-CENTRAL a SMS-CENTRAL

- a) V subsystému SIS-CENTRAL jsou uloženy hodnotové údaje společně s identifikací příslušnosti k dané/daným časovým řadám, ostatní metainformace k řadě lze získat v subsystému SMS-CENTRAL a SMS-CASOVE RADY.

- b) Dále se v subsystému SMS-CENTRAL dále konkretizuje (viz 4.3.3.b), z kterých datových zdrojů se naplňuje časová řada: i) počátek a konec úseku, ii) konkrétní VIP (STAPRO) a datový zdroj – statistická úloha, iii) operace k zajištění srovnatelnosti časové řady, iv) datové zdroje k naplnění předběžných (1.verze), semidefinitivních (1. roční revize) a definitivních dat (2. Roční revize) a priority k jejich naplňování.

4.3.9. Vazby subsystému SMS-CASOVE RADY a SMS-ULOHY

- 23 | Mezi subsystémy žádná přímá vazba není. Požadavky na vedení časových se do subsystému SMS-CASOVE RADY dostávají zprostředkovaně přes SMS-VYSTUPY.
- 24 | Ostatní atributy charakterizující úlohu se získává přímo z datových zdrojů nebo ze SMS- CENTRAL.

4.4. Očekávaná datová architektura

- 25 | Aplikace bude mít veškerá data uložena v centrální oraclovské databázi SMS-pod jedním uživatelem. Veškerá data budou ukládána se všemi potřebnými metadaty a bude zajištěna kontrola jejich konzistence.
- 26 | 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 - roli prohlížení) všechny objekty.

4.5. Rozhraní subsystému

- 27 | Vstupním rozhraním je rozhraní vytvořené v subsystému SMS-ULOHY a SMS-KLAS, popř. i SMS-UKAZ (view v Oracle DB SMS), SMS-CENTRAL a bude přístup nad centrální databází (SIS-CENTRAL).
- 28 | Z aplikace SMS-CASOVE RADY bude vytvořeno výstupní rozhraní (view) pro subsystémy SMS-KVALITA, SIS-CENTRAL, SMS-CENTRAL, VDB, DISEMINACE, SMS-VYSTUPY.
- 29 | Struktura a obsah rozhraní budou definovány dodatečně při schvalování datového modelu.
- 30 | Aplikace bude využívána pouze interními zaměstnanci, nebude mít definováno výstupní rozhraní pro externí uživatele.