

Příloha č. 6 k Zadávací dokumentaci

Obecné exporty

Zadání požadavků na řešení zkušebnímu příkladu

Anotace

Uživatelským cílem je vytvoření nástroje pro dynamickou tvorbu exportních úloh ve smyslu definice exportovaných entit, exportních struktur, omezujících podmínek, datového formátu apod. nad výběrovým (VS) schématem RSO.

1. Vytváření exportních úloh

V rámci vytváření exportu uživatel:

1. Definuje exportní úlohu, tzn.:
 - exportovanou entitu;
 - exportní strukturu;
2. Zadává parametry exportního běhu, tzn.:
 - omezující podmínku/filtr;
3. Zadává vlastnosti výstupního souboru, tzn.:
 - znakovou sadu;
 - datový formát.

Každou definici exportní úlohy bude možné uložit a znovu použít, případně uložit jako výchozí stav pro definici nových obsahově příbuzných úloh.

2. Exportovaná entita

Bude možné provádět export následujících „primárních“ entit:

- Stavební objekt (BUDOVA)
- Statistická budova (STABUD)
- Adresní místo (ADRESA)
- Bytová jednotka (BYTJED)
- Územní číselník (CISEL)

3. Exportní struktura/katalog exportních údajů

V závislosti na zvolené exportované entitě (viz tabulkové matice) bude exportní struktura definována množinami:

- numerických identifikátorů;
- přirozených identifikátorů;
- textových atributů;
- sledovaných atributů;
- obrazových bodů;

- územních identifikací;

Katalog exportních údajů

Aplikace nabídne pro každou entitu seznam přípustných a dostupných údajů organizovaných do věcných skupin (viz výše), ze kterého bude možné sestavit libovolnou kombinaci obsahu výsledného exportu.

K tomuto účelu bude na úrovni datového modelu udržován řídící **katalog exportních údajů**, kde bude, kromě informace o přípustné entitě, kategorii (věcné skupině) a vstupním zdroji, uložena i referenční (pro všechny exporty společná) definice výstupní datové struktury (název, datový typ a rozměr pole tabulky DBF, případně tag XML). Tato definice bude editovatelná pouze na úrovni katalogu jeho správcem a nebude možné ji dále měnit v průběhu tvorby exportu.

Numerické identifikátory

Množina údajů pro entity: Stavební objekt, Statistická budova, Adresní místo a Bytová jednotka.

Entita	Seznam přípustných údajů
Stavební objekt	CUZK_BUD_ID
Statistická budova	CUZK_BUD_ID, IDOB, CISBUD, ROZLIS, IDADR*
Adresní místo	CUZK_BUD_ID, IDOB, CISBUD, ROZLIS, IDADR
Bytová jednotka	CUZK_BUD_ID, IDOB, CISBUD, ROZLIS, IDADR, PORADI

* IDADR hlavní adresy (TYP_ADRESA=1)

Přirozené identifikátory

Množina údajů pro entity: Stavební objekt, Statistická budova, Adresní místo a Bytová jednotka.

Entita	Seznam přípustných údajů
Stavební objekt	PUKB, PUKA*
Statistická budova	PUKB, PUKA*
Adresní místo	PUKA
Bytová jednotka	PUKB, PUKA

* PUKA hlavní adresy (TYP_ADRESA=1)

PUKB: TYPCIS + CISDOM + CASTOBCE (+TEXT)

PUKA: PUKB + KODUVP (+TEXT) + CISOR

Textové atributy

Množina údajů pro entity: Stavební objekt, Statistická budova, Adresní místo a Bytová jednotka.

Entita	Seznam přípustných údajů (typ textu)
Stavební objekt	60*, 61*
Statistická budova	60*, 61*, 10, 20, 30, 40
Bytová jednotka	30, 40, 50

* první ze seznamu (PORADI=1)

Sledované atributy

Množina údajů pro entity: Statistická budova, Adresní místo, Bytová jednotka a Územní číselník.

Entita	Seznam přípustných údajů
---------------	---------------------------------

Statistická budova	Sledované atributy (SLEATR) k SB (viz Matice zdrojů)
Adresní místo	Sledované atributy (ADRATR) k AM (viz Matice zdrojů)
Bytová jednotka	Sledované atributy (ATRBYT) k bytu (viz Matice zdrojů)
Územní číselník	Atributy územních číselníků (KLAS + RSO); odvozené a agregované atributy ÚČ (VS)

Obrazové body

Množina údajů pro entity: Statistická budova, Adresní místo a Územní číselník.

Entita	Seznam přípustných údajů
Statistická budova	Obrazové body SB
Adresní místo	Obrazové body SB* + Obrazové body AM
Územní číselník	Obrazové body ÚČ

* v případě, že neexistuje obrazový bod AM

Územní identifikace

Množina údajů pro entity: Stavební objekt, Statistická budova, Adresní místo, Bytová jednotka a Územní číselník.

Entita	Seznam přípustných údajů
Stavební objekt	Nadřazené územní vazby číselníku 55* (CHODNOTA+ZKRTEXT)
Statistická budova	Nadřazené územní vazby číselníku 55 (CHODNOTA+ZKRTEXT)
Adresní místo	Nadřazené územní vazby číselníku 55 (CHODNOTA+ZKRTEXT)
Bytová jednotka	Nadřazené územní vazby číselníku 55 (CHODNOTA+ZKRTEXT)
Územní číselník	Nadřazené územní vazby daného číselníku (CHODNOTA+ZKRTEXT)

*Sledovaný atribut statistické budovy (SLEATR).

4. Omezující podmínka

Exportovaná data bude možné v závislosti na zvolené entitě omezit/filtrovat:

- územně (dle územního číselníku);
- věcně (dle sledovaných atributů).

Územní filtrace

Pro entity: Stavební objekt, Statistická budova, Adresní místo, Bytová jednotka a Územní číselník.

Entita	Filtrace
Stavební objekt	Nadřazené územní vazby číselníku 55
Statistická budova	Nadřazené územní vazby číselníku 55
Adresní místo	Nadřazené územní vazby číselníku 55
Bytová jednotka	Nadřazené územní vazby číselníku 55
Územní číselník	Nadřazené územní vazby daného číselníku

Věcná filtrace

Pro entity: Statistická budova, Adresní místo a Bytová jednotka

Entita	Filtrace
Statistická budova	Dle sledovaných atributů SB (SLEATR)
Adresní místo	Dle sledovaných atributů AM (ADRATR)
Bytová jednotka	Dle sledovaných atributů bytu (ATRBYT)

5. Datový formát

Primární datovým formátem bude dBase IV (prioritně), alternativním pak XML.

Metapopis exportu

Ke každému exportu bude automatizovaně vytvořen soubor (log) metapopisu generování a provedeného exportu dat RSO, jehož obsah bude definován zadavatelem.

6. Popis výběrového schématu RSO

6.1. Výběrové schéma RSO

Výběrové schéma RSO (VS) obsahuje konsolidovaná a stabilizovaná data RSO. Slouží jako zdroj pro poskytování dat z RSO pro vnitřní potřeby ČSÚ i pro veřejnost. Výběrové schéma je vytvářeno na základě rozhodnutí správce RSO jako opakovaný proces a datový obsah VS RSO není dále aktualizován. Všem výběrům z něj vytvářeným proto je možné zaručit stejné podmínky, a tedy i konzistentní datové výstupy.

6.1.1. časový rozměr výběrového schématu

Výběrové schéma (VS) obsahuje pouze data platná k danému časovému okamžiku definovanému v rámci vytváření VS.

VS neobsahuje historii jednotlivých entit, jde o časový snímek editační báze.

Záznamy ukončené, resp. budovy zaniklé před určeným datem pohledu, do VS RSO přenášeny nejsou.

6.1.2. Věcný obsah výběrového schématu

Obsahem výběrového schématu jsou všechny základní entity RSO:

- **Evidenční část RSO**
(Budovy, Statistické budovy, Bytové jednotky a Adresy se všemi atributy a texty)
- **Atributy územních celků**
(Sledované atributy a Obrazové body územních celků včetně textů)

Součástí VS jsou všechny platné (ZRUS=N) záznamy uvedených entit.

Datový obsah VS RSO je po celou dobu existence aktuální instance VS neměnný. Není proto nutné řešit zpětnou aktualizaci RSO daty VS.

Replika číselníků KLAS není součástí VS, je využito společné repliky umístěné v databázi soustavy registrů.

6.2. Odvozování atributů územních celků

Atributy územních celků v RSO obsahují většinu statistických a geografických atributů pouze na nejnižší nutné územní úrovni. VS je obohaceno o hodnoty části těchto atributů agregované za nadřazené územní úrovně.

Tyto agregované hodnoty atributů jsou dopočítávány v rámci vytváření VS. Příloha č. 1 obsahuje přehled všech sledovaných atributů územních celků sledovaných v RSO. V tabulce je vyznačena ta část aplikace RSO, kde odvozování a dopočet probíhá. Při vytváření VS jsou počítány hodnoty všech atributů, které nejsou součástí editačního schématu RSO, ale jsou obsaženy ve VS RSO.

Definice agregovaných atributů RSO je součástí **Katalogu atributů územních celků RSO**. Katalog obsahuje i definici jejich agregačních algoritmů. Katalog atributů tvoří řídicí strukturu celého procesu odvozování, a je proto tvořen jako pevná struktura bez nástroje uživatelské správy jeho obsahu.

6.2.1. Metody odvozování atributů

V modelu RSO v 4.1 jsou vytvořeny nástroje pro odvozování hodnot některých atributů. Následující odstavec uvádí jejich stručný popis.

6.2.1.1. Odvození v RSO

Dávkové odvozování hodnot atributů probíhá nad evidenční částí RSO formou vytvoření datové dávky pro hromadnou aktualizaci RSO a jejím následným použitím pro hromadnou aktualizaci obsahu evidenční části RSO. Vytvoření datového obsahu dávku probíhá výpočtem z tabulek evidenční části RSO, a to buď nad budovami, byty nebo adresami za definovanou úroveň území.

Takto jsou počítány hodnoty atributů platné ke zvolenému časovému okamžiku. Do evidenční části RSO se tyto hodnoty promítají jako hodnota atributů platná od zvoleného časového okamžiku do konce platnosti položky číselníku.

Každý atribut má v katalogu atributů definovanou funkci, která je pro výpočet hodnot atributu za zvolené území použita. V katalogu je jako zdroj hodnot těchto atributů označeno Výpočet nad (adresami, budovami, byty) – dávka.

6.2.1.2. Odvození při vzniku VS

Metoda odvození při vzniku VS je využita pro atributy, které se nevyskytují v evidenční části RSO, ale až ve VS a případně iRSO.

Vytvoření datového obsahu atributů probíhá výpočtem z tabulek evidenční části RSO, a to buď nad budovami, byty nebo adresami za definovanou úroveň území, v době vytváření VS. Data tedy nejsou ukládána do RSO, ale pouze do nově vznikajícího VS.

Každý atribut má v katalogu atributů definovanou funkci, která je pro výpočet hodnot atributu za zvolené území použita. V katalogu je jako zdroj hodnot těchto atributů označeno Výpočet nad (adresami, budovami, byty).

6.2.1.3. Agregace

Metoda agregace je použita pro výpočet hodnoty atributu za vyšší územní celek jako sumy hodnot atributu nižšího celku. Agregace je použita pouze na atributy, které jsou definovány pouze ve VS RSO nebo iRSO. Probíhá tedy v průběhu vytváření VS.

V katalogu atributů jsou tyto atributy označeny ve sloupcích Agregace pomocí kódu atributu, jehož hodnota má být sčítána a informace o vazbě v KLAS, která má být pro agregaci použita. V dokumentu v příloze je uveden jako identifikace atributu jeho akronym, který ale z definice KLAS nemusí být unikátní. V katalogu atributů RSO proto bude nutné nahradit identifikaci agregovaného atributu pomocí unikátního kódu atributu.

V případě, kdy je možné zaručit, že mezi číselníky (cílovým a agregovaným) existuje jednoznačně vyhledatelná přímá vazba, není nutné identifikovat v agregacním vymezení vazbu, ale stačí pouze kód atributu.

Agregace v průběhu vytváření VS probíhá vždy až po skončení odvozování atributů při vzniku VS. Proto je zaručeno, že agregací mechanismus použije správné hodnoty zdrojových atributů, pokud jsou tyto zdrojové atributy odvozované při vytváření VS.

Standardně jsou všechny agregace na stejné úrovni, jejich pořadí proto není zaručeno. Pokud by agregací mechanismus definoval agregaci z jiného také agregovaného atributu, je nutné v katalogu specifikovat pořadí agregací.

Návrh katalogu atributů RSO je obsažen v příloze **Katalog atributů územních celků**.

7. Funkční model

7.1. Vytváření výběrového schématu

Proces opakovaného vytváření VS RSO je spouštěný ručně z formuláře aplikace RSO. Vlastní proces pro vytváření VS RSO probíhá na pozadí tak, aby jeho činností nebyl blokován uživatelský program, a aby chod vytváření nebyl závislý na připojení klientského PC k aplikačnímu serveru a databázi. V rámci formuláře správy vytváření VS je možné monitorovat stav procesu vytváření VS.

7.1.1. Mechanismus vytváření VS RSO

Výběrové schéma je vytvářeno pokaždé zcela od počátku. Po zahájení procesu vytváření VS je napřed zcela vymazáno do té doby platné VS, přičemž objekty datového modelu VS jsou zachovány. V době vytváření VS jsou z důvodu výkonnosti potlačena veškerá integritní omezení objektů VS. Tato omezení jsou aktivována až po naplnění obsahu VS.

Vlastní vytváření VS probíhá formou uložené PL/SQL procedury spuštěné jako úloha na pozadí (job). Úloha vytváření VS je složena z ucelených kroků, které na sebe navazují. Dokončení kompletního kroku je vždy indikováno do tabulky běhů tak, aby v případě přerušení běhu z důvodů technické závady bylo možné pokračovat v započatém běhu od posledního dokončeného kroku.

7.1.2. Dostupnost VS RSO

Vzhledem k tomu, že VS RSO je vytvářeno zcela od počátku, není VS po dobu vytváření VS dostupné pro žádnou aplikaci RSO.

8. Datový model

Datový model odděleného VS RSO pro potřeby zkušebního příkladu tvoří dvě části:

- Model výběrového schématu evidenční části RSO (budovy, byty, adresy)
- Model atributů územních číselníků výběrového schématu RSO

Vztahy mezi jednotlivými tabulkami zobrazují následující schémata. Detailní popis všech tabulek uvedených ve schématech je obsažen v přílohách.

8.1. Základní vlastnosti datového modelu

Datový model VS RSO je uložen v odděleném datovém schématu RSOVS. Data modelu jsou uložena v tabulkových prostorech RSOVS_DATA a indexy v prostoru RSOVS_INDXX.

Datový model RSO VS obsahuje vždy nejvýše jeden platný časový snímek dat evidenční části RSO a části atributů územních číselníků RSO. Tento časový snímek je datově podmnožinou primárních dat RSO. Aktuální velikost diskového prostoru evidenční části RSO ve verzi 3.5 je zhruba 100 GB. Lze předpokládat, že výsledná požadovaná velikost diskového prostoru VS RSO nepřesáhne 200 GB.

8.2. Model výběrového schématu evidenční části RSO

Evidenční část VS RSO je určena pro uložení časového snímku informací o budovách, bytech a adresách včetně všech atributů a textů.

Protože rychlost aplikace RSO při výběrech z evidenční části RSO není kritická, nebylo nutné optimalizovat uložení dat s ohledem na rychlost výběru. Optimalizace uložení dat VS RSO byla vytvářena zejména s ohledem na minimalizaci datového objemu uložených dat a na minimalizaci nutných změn v těch částech aplikace, které přecházejí ze současné evidenční části RSO do VS RSO.

Datový model VS RSO proto vychází z datového modelu evidenční části RSO s tím, že obsahuje některá společná omezení:

- byly zrušeny všechny servisní atributy (datum a autor poslední aktualizace záznamu, příznak logického zrušení záznamu, odkaz na historii změn záznamů)
- byla zrušeny všechny informace o platnosti jednotlivých položek podřízených tabulek s výjimkou data založení a data zániku budovy v RSO.

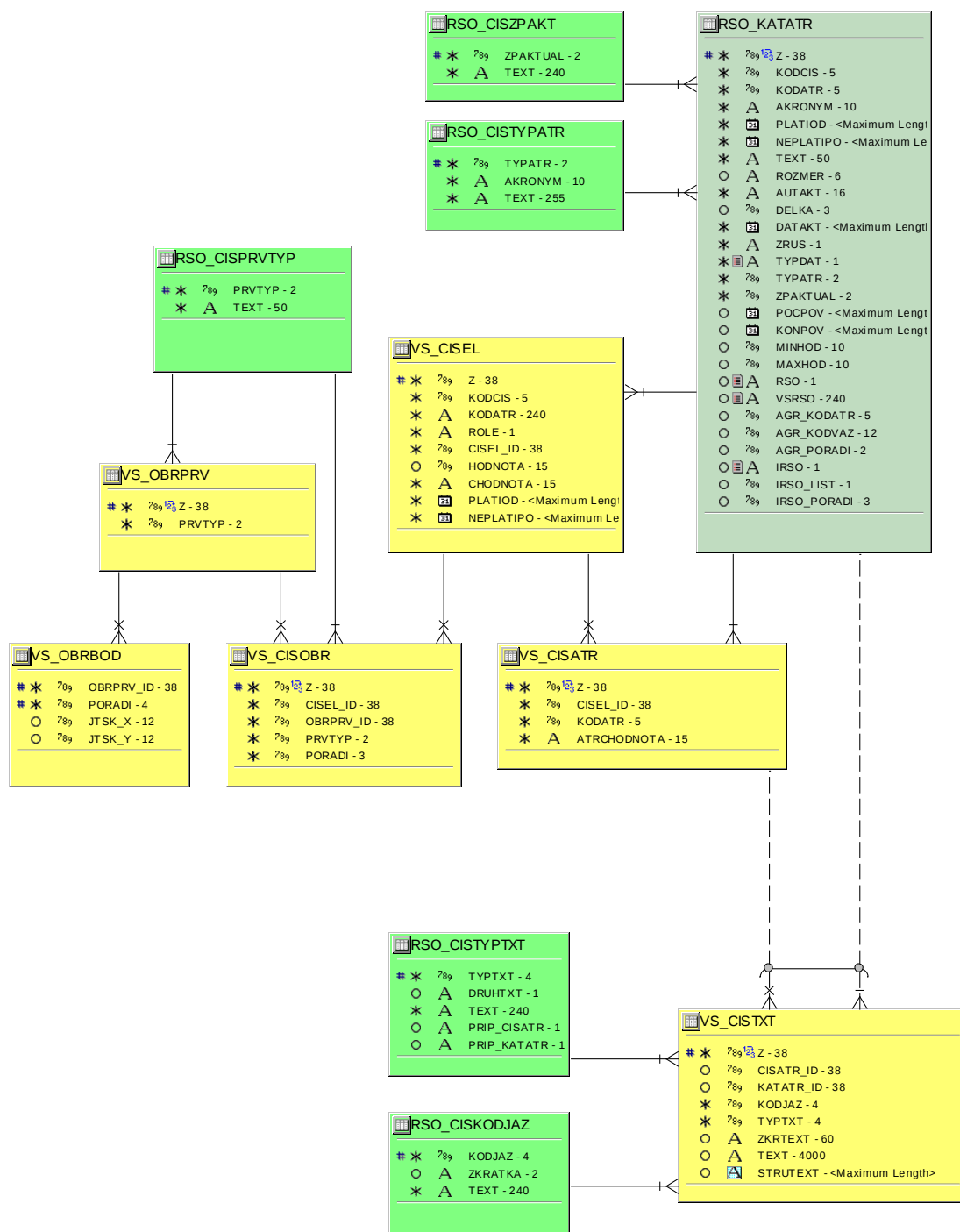
Tabulky RSO_KONTYPZDR a RSO_IDEATR jsou řídicí tabulky evidenční části RSO a ve schématu jsou zahrnuty kvůli úplnosti a kvůli vyznačení nutných vazeb a integritních omezení.



8.3. Model atributů územních číselníků VS RSO

Atributy územních číselníků ve VS RSO jsou uloženy v modelu postaveném na modelu atributů územních číselníků v primárních datech RSO. Ve VS RSO je tabulka položek číselníků VS_CISEL řešena jako samostatná fyzická tabulka, nikoliv jako materializovaný pohled do dat KLAS. Proto bylo možné postavit mezi touto tabulkou a jejími podřízenými tabulkami integritní omezení primární – cizí klíč za využití indexů. To umožňuje efektivnější práci s daty atributů a textů.

Tabulka RSO_KATATR je řídicí tabulka aplikace RSO (katalog atributů) a je na schématu zobrazena pro vyznačení vzájemných vazeb a integritních omezení.



Obrázek 2: Schéma datového modelu atributů územních číselníků VS RSO

9. Přílohy

1. Katalog atributů územních celků



SSREG_3_6_oddělení
schémat_katalog_v1..

2. Datový model evidenční části VS RSO



VRSO_Evidenční_čá
st.pdf



VRSO_Evidenční_čá
st_Tabulky.pdf

,

3. Datový model atributů územních číselníků RSO.



VRSO_Územní_čísel
níky.pdf



VRSO_Územní_čísel
níky_Tabulky.pdf

,