

Veřejná databáze ČSÚ (VDB) a vizualizace demografických údajů

14.12.2011

diskusní večer

České demografické společnosti

Obsah

- Cíle, koncepce, principy
- VDB - součást SIS
- Schema VDB a metadata ve VDB
- Vstupy statistických údajů do VDB
- Výstupy z VDB – prezentační aplikace, animované grafy
- VDB 2.0

Cíle, koncepce

- Hlavní cíl – jednotný systém agregovaných statistických dat určených ke zveřejnění, prezentace na základě principu „jednoty čísla“
- interaktivní přístup uživatelů k informacím
- příprava standardních i nestandardních publikací
- zajištění výstupů pro další systémy tuzemské i zahraniční
- **system je založen na metadatové indettifikaci údajů – Statistický metainformační systém (SMS)**
 - (ukazatele – statistické proměnné, číselníky, statistické úlohy)
- **návaznost na proces zpracování statistických dat**

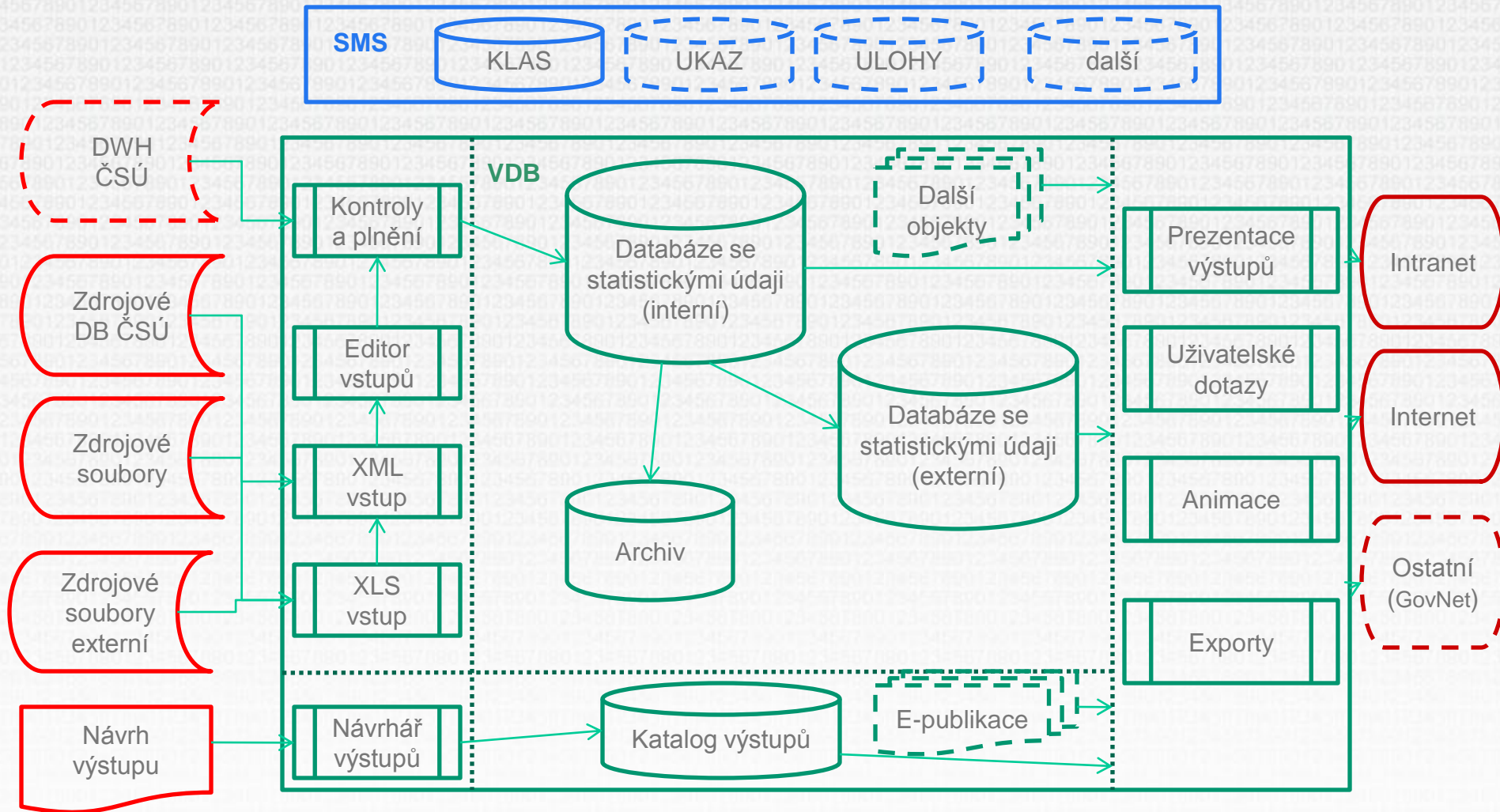
Některé principy

- statistické údaje ze **všech statistik** v jediné „maxi-tabulce“
- jeden zdroj dat všech výstupů pro veřejnost
- **jednota statistického údaje** – korekce přístupu ve VDB2
- možnost kombinací hodnot z různých vstupů do jednoho výstupu
- možnost uživatelských úprav výstupů a výběrů dat
- grafické formy výstupů – možnost přepínání T-G-M (tabulka – graf – mapa), datová základna pro další vizualizaci
- jasné porozumění údajům – metainformace ke každé hodnotě

VDB - součást SIS ČSÚ

- VDB - součást Statistického informačního systému (SIS)
 - obsahuje agregovaná statistická data jako výsledek zpracování – **VDB NIC NEPOČÍTÁ**, pouze prezentuje výsledek zpracování
 - obsahuje i statistická data z externích zdrojů (např. některá data regionální statistiky z UIV, MPSV, ÚZIS apod.)
- VDB je **jedním** z nástrojů zajišťujícím **diseminaci** statistických dat
 - datové portfolio - veškerá agregovaná data určená k prezentaci ukládána ve VDB
 - vedle VDB ovšem mohou existovat další diseminační nástroje, zejména nástroje umožňující ad-hoc výběry a zpracování dotazů nad databázemi mikrodat (např. databáze zahraničního obchodu, data-mart SLDB, ...)
- Výstupy VDB
 - **předdefinované výstupní statistické objekty**, tedy makety pro konkrétní tabulky, grafy, mapy
 - uživatelem definovaný výběr dat

Schéma VDB



VDB a metadata

- Datový model VDB – založen na identifikaci statistických údajů pomocí metadat (SMS)
- Metadatový popis – **podmínka pro uložení dat i navrhování výstupních objektů**
- Zajišťuje metodik VDB – doplňuje za všechny úsekové statistiky - vlastní analýza a návrhy
- Konzultace s VPO a SMS

Identifikační metadata

123

CO ?

Počet přistěhovalých

KodUk

podle

Pohlaví:	Muži
Věk:	15-64 let

Druh_Id

KDE ?

Z:	Francie
Do:	Česká republika

Pro_Id

KDY ?

Od:	Rok 2005
Do:	Rok 2006

Cas_Id

Další?

osoba

MJ_Id

běžná hodnota

CharH

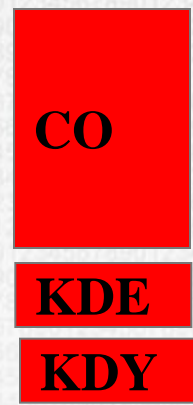
předběžný

TypZvu

Identifikace dat

7 dimenzí

- Ukazatel
- Měřicí jednotka
- Druhová specifikace
- Prostorová identifikace
- Časová identifikace
- Charakter hodnoty – absolutní údaj, index apod.
- Typ údaje – předběžný, revize, definitivní apod.



123	KodUk	Druh_Id	Pro_Id	Cas_Id	MJ_Id	CharH	TypZvu
------------	-------	---------	--------	--------	-------	-------	--------

Toto je „DNA“ pro hodnoty ve VDB

Každá hodnota v celé VDB má unikátní DNA

Doplňková metada

- rozšiřují informace o statistickém údaji
 - důvěrnost
 - zdroj údaje
 - interval zveřejnění
 - poznámka k údaji
- rozvoj SMS umožní připojovat další informace – např. o kvalitě
- Vlastnost **AKTUAL** – nejčerstvější údaj

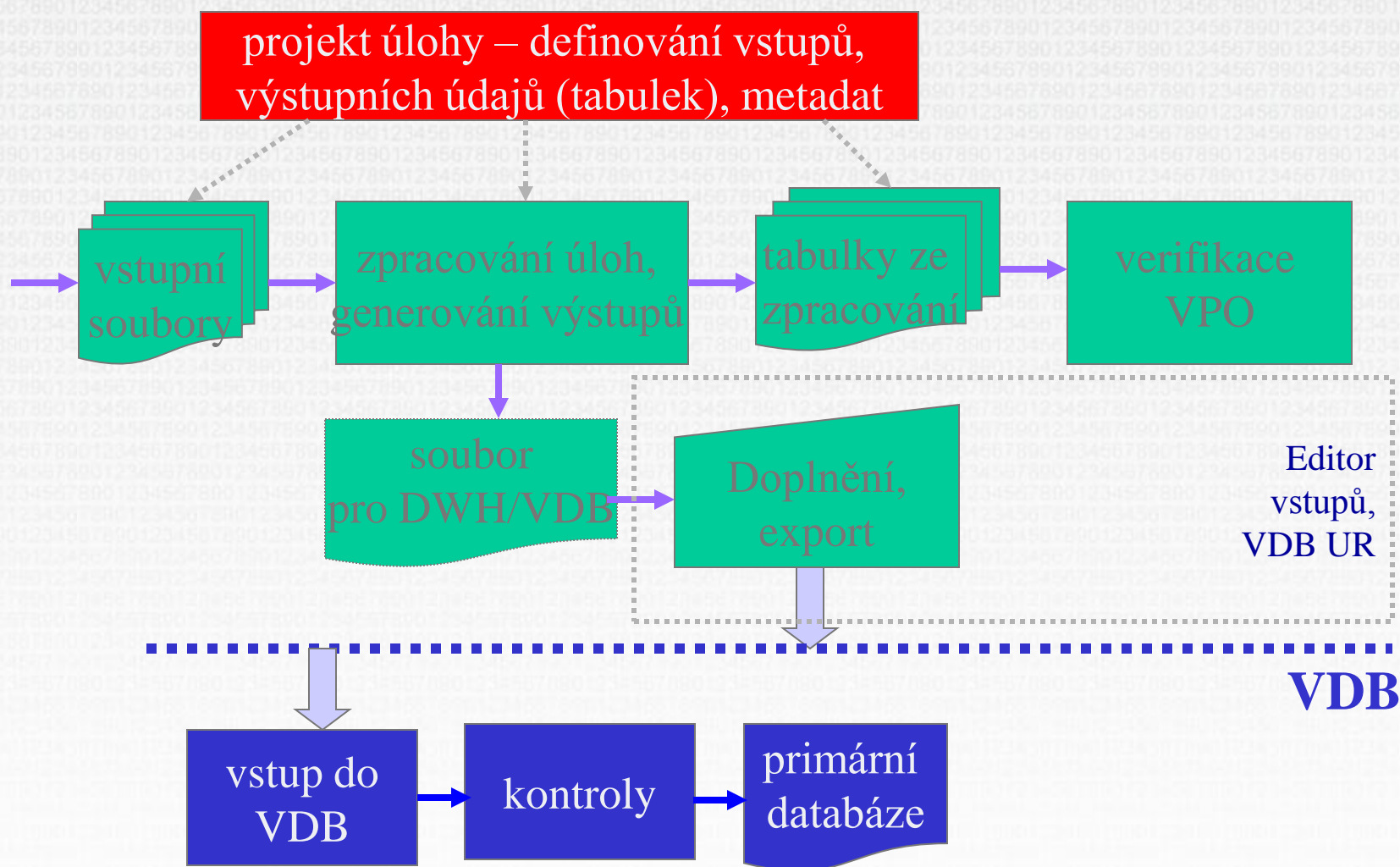
Vstupy statistických údajů do VDB

- Správce datového zdroje (VPO) – zajistí dodání vstupních dat správě VDB
- Správa VDB
 - Plnění VDB pro interní uživatele – kontroly, test kolizí a duplicit
 - Synchronizace údajů do VDB pro externí uživatele – bez důvěrných dat – automatický proces 1 denně, možnost i okamžitého ručního přenosu dat

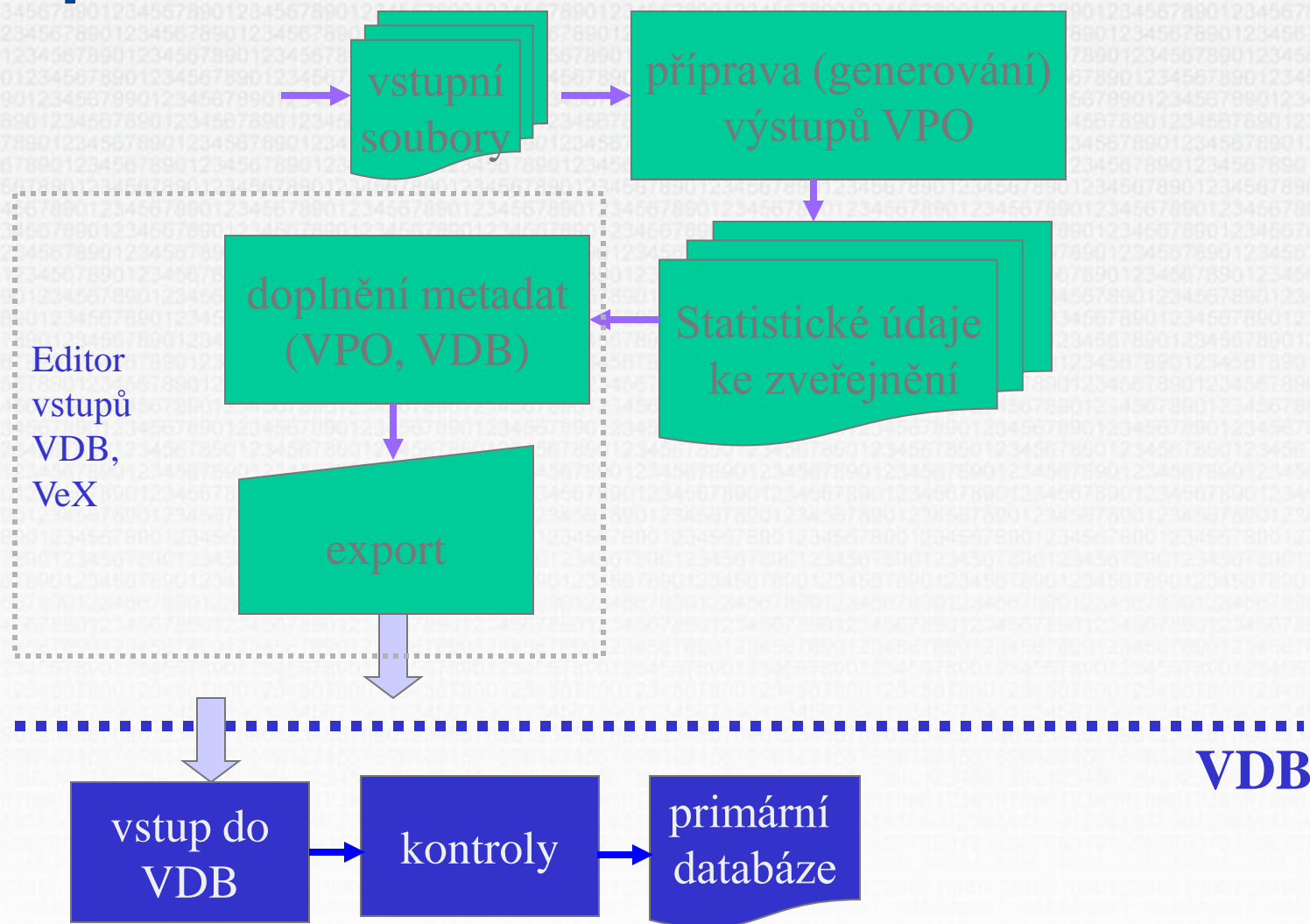
Připojování metadat

- Definované univerzální vstupní rozhraní VDB - struktura souborů (údaje, metadata)
- Standardní cesta – bez nutnosti ex-post doplňování metadat
 - přímá vazba na výstupy ze zpracování – volební statistika
 - generování souborů (XML) případně využívání editoru vstupních dat – národní účty, VŠPS, zahraniční obchod, konjunkturální průzkumy...
- Nestandardní cesta – nutno dodatečně připojovat metadata, zajišťuje oddělení VDB
 - automatické připojení (speciální aplikace) - demografie, cestovní ruch, RES, ...
 - ruční připojení – k tabulkám XLS (od VPO) se ad-hoc provádí analýza a připojování metadat (zajišťuje oddělení VDB)

Vstup souborů do VDB – standardní cesta



Vstup souborů do VDB – nestandardní cesta



Výstupy z VDB

- Předdefinované výstupní objekty – „prázdné“ makety tabulek s možností prezentace grafů a map, navrhuji autoři (sekce 7, VPO)
- Výstupní objekty (VO)
 - Návrh VO
 - Autorské makety - vidí pouze autor a správa VDB
 - Zařazení do navigačního stromu, klasifikace VO (čas, typ území)
 - Typologie výstupu (základní, doplňkový apod.)
 - Schvalování VO – zpřístupnění pro interní uživatele – metodik VDB na základě žádosti autora
 - Veřejné – vidí všichni interní uživatelé
 - Zveřejnění VO – ruční převedení do databáze pro externí uživatele na základě žádosti autora
 - Veřejné makety se stanou dostupné externím uživatelům
 - **Externí uživatelé nevidí důvěrná data!**

Další výstupy


- Uživatelský výběr dat
 - Uživatel definuje požadavek na data a finální podobu výstupu – *zatím jen pro interní uživatele*
 - Různé scénáře (*moje území, moje téma*)
- Animované grafy
 - Speciální VO pro dynamické grafy (časové řady) - využití Google Motion-Chart
- *Připravuje se návrhář PDF publikací*
 - *Výstupní objekty VDB (tabulky, grafy, mapy)*
 - *Statické externí objekty (metodika, komentáře, obrázky, kartodiagramy, ...)*

Prezentace – hlavní zásady

- přístupnost
 - česká a anglická jazyková verze
 - plná funkčnost pro nejrozšířenější prohlížeče
 - zohlednění pravidel dle W3C a Blind Friendly Web
- komunikace s uživateli
 - základní info o VDB, stručný návod
 - zpětná vazba – email, FAQ, připomínky
 - sledování návštěvnosti, nejžádanější tabulky (TOP10)

Prezentace VDB na webu

<http://vdb.czso.cz/>

CSU Český statistický úřad  Kont




Lidé a společnost | Ekonomika | Trh práce a mzdy | Věda, IT | Zemědělství, ekologie | Vícetématické informace

NEJŽÁDANĚJŠÍ

- [ČR v číslech](#)
- [Inflace](#)
- [Intrastat](#)
- [Makroekonomické údaje](#)
- [Obyvatelstvo](#)
- [Regiony, města, obce](#)
- [Ročenky](#)

Nejnovější čísla | Analýzy | Časové řady

NEJNOVĚJŠÍ ČÍSLA - RYCHLÉ INFORMACE

-  **16.05. Indexy cen výrobců (duben 2011)**
Ceny zemědělských a průmyslových výrobců dále rostly
-  **16.05. Indexy cen vývozu a dovozu (březen 2011)**
Vývozní a dovozní ceny rostly
-  **13.05. Předběžný odhad čtvrtletního HDP (1. čtvrtletí 2011)**

PUBLIKACE, DATA

[Tématické skupiny](#)

VDB Veřejná databáze ČSÚ  Napište nám | Web ČSÚ | English 

Úvodní stránka | Navigační strom | O Veřejné databázi | Časté dotazy | Definice ukazatelů | Nápověda

- 1** Souborné informace
- 2** Životní prostředí, zemědělství
- 3** Práce, sociální statistiky
- 4** Obyvatelstvo, volby
- 5** Makroekonomika
- 6** Zahraniční obchod
- 7** Ceny
- 8** Průmysl, staveb., energ.
- 9** Služby

Hledej

v kódech v textech

 Tabulky  Grafy  Mapy [Nastavit filtry](#) 

- Aktuality
- Animované grafy
- Návštěvnost
- Odkazy
- Služby pro Vás

VDB Web 1.3.14 © Český statistický úřad Vygenerováno 17.05.2011 17:44

Prezentační aplikace

- Stromová navigační struktura – temata
- Seznam předdefinovaných objektů
 - Základní výstupy
 - Doplnkové výstupy
- Filtrování – grafické výstupy, typy území, referenční období
- Vyhledávač
 - texty v hlavičkách, legendách, názvech
 - Kódy výstupních objektů
- Přepínání formy zobrazení – tabulka, graf, mapa
- Hypertexty – metainformace (ukazatel, položka číselníků, metainformační popis každé hodnoty)
- Výběr parametrů pro vybraný výstupní objekt
- Exporty a tisky – XLS, XML (včetně metadat)
- Animované grafy – [dynamické](#) a [statické](#)
- Návštěvnost

Návštěvnost a připomínky

- Návštěvnost
 - využití statistik Google Analytics
 - měření přístupů ke stránkám VDB
 - vlastní načítání zobrazení, tisku, exportu
 - výsledky návštěvnosti
 - v prezentační aplikaci – Top10 a celkové přístupy
 - Pro správu VDB – XLS analýza (Intranet)
- Připomínky uživatelů
 - e-mail – možnost odeslání připomínky přes GW
 - evidence připomínek (XLS)
 - interní workflow vyřizování – ve spolupráci s informačními servis

VDB 2.0

- Založený na SMS
 - Poprvé využívání statistických proměnných (SMS-UKAZ)
 - upřesňování pomocí elementárních proměnných (SMS-KLAS)
 - definování výstupů (SMS-ULOHY a SMS-DISEMINACE)
- Předpoklad výhradního čerpání z datového skladu
- První použití na datech [SLDB 2011](#) – v roce 2011-2012
 - přebírání dat z DWH – standardní cesta – nutno upravit aplikace vstupu a synchronizace dat
 - nový návrhář výstupních objektů
 - návrhář publikací PDF

VDB a nová role v diseminaci

- Výstupní objekty VDB - on-line dostupnost, nahradí tabulky v publikacích, možnost vzájemných vazeb (odkazování)
- Publikace – kombinace výstupních objektů VDB (tabulky) a dalších objektů (textové části, speciální grafické výstupy)
- Interní výstupy – pracovní výstupy VPO, verifikace dat před zveřejňováním
- Jednotná data – data uložena jednou, ale použita v různých výstupech
- Integrovaná součást webu – datová základna moderních vizualizačních technik

Díky moc za pozornost...

Kontakt:
Eduard Durník, ČSÚ

tel: 2 7405 3143, 604 871 271

mail: eduard.durnik@czso.cz

Odkazy

- VDB4 <http://vdb.czso.cz/vdbvo/uvod.jsp>
- AG na VDB4 <http://vdb.czso.cz/vdb/charts.jsp>
- VDBE – SLDB <http://vdb.czso.cz/slodbvo/>
- Statické AG

<https://docs.google.com/spreadsheet/cc?key=0AsAMX8fmHJPMdFA5NVNyUGRVZW5yRzJrandmTnY3bVE#gid=2>

- Ukázka vložených grafů – <http://mcerny.blogspot.com>