

## Využití Business Intelligence v kraji Vysočina

### The usage of Business Intelligence in the Vysocina region

Michal Šulc | sulc.m@kr-vysocina.cz

Krajský úřad kraje Vysočina, Odbor analýz

Miloš Vystrčil | vystrcil@vspj.cz

Vysoká škola polytechnická Jihlava, Katedra veřejné správy a regionálního rozvoje

#### ABSTRAKT:

Článek se věnuje využití metody zpracování a analýzy dat Business Intelligence v kraji Vysočina. Je popsána stručná historie a podstata fungování metody, současný stav řešení, hlavní přínosy a problémy spojené s dosavadním používáním metody v praxi. Načrtnut je budoucí potenciál metody, a to jak v rovině běžného operativního řízení chodu kraje, tak v rovině strategického regionálního managementu. Cílem článku je rovněž zdůraznit potenciál metody pro efektivní spolupráci veřejných institucí v oblasti zpracování a poskytování dat.

#### KLÍČOVÁ SLOVA:

Business Intelligence, zpracování dat, analýza dat, manažerský informační systém

#### KEYWORDS:

Business Intelligence, data processing, data analysing, managerial information system

### Úvod

Systémy Business Intelligence (BI) se řadí do skupiny nástrojů, které podporují manažerské rozhodování v organizaci (Managerial Information Systems, Executive Information Systems, Decision Support Systems). Jde o soubor komponent, které pomáhají data zpracovávat, analyzovat a doručovat uživatelům. Jejich základem je multidimenzionalita, tedy nazírání na organizační data a výsledky z mnoha úhlů pohledu. Nástroje BI slouží především pro strategické řízení organizace, své opodstatnění však najdou i v rovině takticko-operativního řízení. Lacko (2006) metodu definuje jako „proces transformace dat na informace a převod těchto informací na poznatky prostřednictvím objevování.“

„Business Intelligence“ v širším pojetí vyjadřuje určitý typ filozofie přístupu k datům a informacím potřebným pro realizaci různých činností. Ve spojení s výpočetní technikou byl pojem použit již v roce 1958, kdy člen vědeckého týmu společnosti IBM H. P. Luhn hovoří ve svém článku (Luhn, 1958) o automatickém systému zpracování informací v dokumentech pro průmyslové, vědecké a vládní organizace.

Následný vývoj metody je provázán s rozvojem kvantitativních metod sloužících pro podporu rozhodování v organizacích. Zejména se jedná o rozvoj modelové a datové

orientovaných systémů. Vývoj druhé skupiny systémů vyústil v tvorbu rozsáhlých datových skladů a manažerských informačních systémů, obsahujících široké spektrum dat pro podporu rozhodování (Powel, 2007).

Do běžné podnikové praxe vstoupily automatizované informační systémy v průběhu osmdesátých let. V této době také došlo k rozvoji systémů založených na multidimenzionálních datových modelech (OLAP – Online Analytical Processing), jejichž principy byly v roce 1993 definovány pomocí tzv. dvanácti hodnotících pravidel (Codd, Codd a Salley, 1993).

K zavedení pojmu „Business Intelligence“ do obchodní praxe došlo koncem osmdesátých let, kdy Howard Dresner, analytik Gartner Group, definoval BI jako sadu konceptů a metod, které by měly sloužit pro zkvalitnění rozhodovacích procesů firem. V průběhu devadesátých let vznikají robustní komerční řešení, která integrují podnikové provozní systémy v jejich plně šíři.

V současné době směřuje vývoj BI zejména k integraci strukturovaných a nestrukturovaných dat (viz například Inmon, 2007). Podle téhož autora (2005) je dokonce propojení strukturovaných a nestrukturovaných dat jedním z nejslibnějších směrů budoucího rozvoje BI. Dalším směrem vývoje je konstrukce mechanismů podporujících rozhodování v reálném čase.

Článek je strukturován následovně: v první části jsou objasněny základní principy metody a jejich logických vrstev, další část se věnuje dosavadnímu využití BI v kraji Vysočina a jsou načrtnuty směry budoucího rozvoje řešení. Následně jsou uvedeny hlavní přínosy a problémy řešení. Konečně v poslední části práce je uveden praktický příklad využití metody.

### Principy Business Intelligence

Data každé větší organizace jsou standardně uložena v provozních systémech (Enterprise Resource Planning Systems) ve formě relačních tabulek a databází. Jejich struktura vyhovuje rychlému transakčnímu zpracování dat (Online Transaction Processing), nikoli však analytické formě uložení a zpracování dat, která je pro manažerský informační systém optimální.

Nástroje BI relačně uložená data integrují do nadstavbové datové vrstvy, která vyhovuje požadavkům analytické-

ho zpracování dat (OLAP). Strukturu systémů BI můžeme v obecné rovině rozdělit do tří vrstev:

#### Vrstva zpracování a importu dat

V této vrstvě probíhají načasované úlohy (tzv. datové pumpy, ETL procesy), které ve stanovených intervalech načítají data z primárních zdrojů do operačního úložiště datového skladu. Primárním zdrojem může být libovolný formát strukturovaných dat – databázový server (např. data účetního systému), soubor ve formátu xls (např. publikace či databáze Českého statistického úřadu) nebo textový soubor publikovaný na webu (např. měnové kurzy ČNB). Důležitou vlastností této vrstvy je její univerzálnost na vstupu, tedy připravenost na datová rozhraní poskytovatelů dat, a prakticky libovolná perioda aktualizace dat.

Jako praktický příklad fungování jednotlivých vrstev je uvedena publikace ČSÚ Regionální účty. Po publikování dat Českým statistickým úřadem na jeho webu je třeba publikaci uložit do pracovního adresáře. Datová pumpa poté data načte (v tomto případě data za rok poslední i data z minulých dvou let, tedy definitivní a semidefinitivní verze dat pro minulé roky) a importuje do operačního úložiště datového skladu.

#### Analytická vrstva (datový sklad)

Hlavním účelem analytické vrstvy je data transformovat do takzvaných multidimenzionálních objektů, které umožňují následnou efektivní a rychlou práci s daty. Multidimenzionální objekty si lze představit jako „datové kostky“, jejichž pomyslné hrany (tzv. dimenze) znamenají úhel pohledu na data. Tyto dimenze jsou často hierarchické (např. měsíc - čtvrtletí - rok, nebo obec - obec s rozšířenou působností - kraj - region soudržnosti - Česká republika) a historické, tj. zaznamenávají datové struktury v historickém pohledu. V hierarchii dimenzí lze nastavit různé způsoby agregace do vyšších úrovní, tedy nejen prostý součet, ale i další matematické operace s daty dle povahy vstupních dat. Rovněž způsob nastavení historizace jednotlivých dimenzí lze nastavit dle povahy a účelu následné analýzy dat.

V udávaném příkladu regionálních účtů je dimenzí čas (rok), region (NUTS 3 – NUTS 2 - ČR) a ukazatel (ukazatele regionálních účtů). Sledovanými fakty jsou %, Kč, mil. Kč, dle povahy ukazatele.

#### Prezentační vrstva

Data připravená v analytické vrstvě jsou již ve vhodném formátu pro prezentační vrstvu, která n-rozměrné objekty poté řeže, zobrazuje či analyzuje v dvourozměrném prostředí analytického software. Na analytickou vrstvu jsou nasazeny nadstavbové nástroje, které umožňují pokročilou práci s daty. Může se jednat například o:

1) Nástroj pro analýzu a reportování dat, který umožňuje vytvářet a ukládat různé pohledy na data ve formě in-

teraktivních tabulek, grafů a dalších grafických a analytických komponent. Po uložení pohledu jsou data vždy aktuální, tedy promítají se zde všechny aktualizace provedené v nižších vrstvách.

- 2) Nástroj pro statistickou analýzu dat, tedy statistický software, který podporuje databázové napojení na analytickou vrstvu.
- 3) Nástroj pro geografickou analýzu, který umožní geografickou vizualizaci pohledu na data.
- 4) Webové nástroje vizualizace a prezentace dat, například flashové animační a vizualizační komponenty.

V případě regionálních makroekonomických údajů z regionálních účtů je například možné v analytickém nástroji vytvořit pohled (například tabulku HDP na obyvatele podle krajů a roků), ve statistickém software zhodnotit vývoj ukazatele (např. variabilitu), v gisovém nástroji vizualizovat výsledky v mapě a konečně pomocí webu výsledky prezentovat na síti. Vytvořené pohledy či sestavy ve všech nástrojích jsou online, tedy po aktualizaci analytické vrstvy se nová data projeví i ve všech výstupech, např. novým sloupcem pro nové období či přepočítáním výsledných hodnot.

## Business Intelligence na Vysočině

Na krajském stupni veřejné správy se BI začíná prosazovat od roku 2004. Několik krajů buduje v současné době řešení založené na této metodě, například Plzeňský kraj, Zlínský kraj, Moravskoslezský kraj a Vysočina.

Kraj Vysočina vytváří datový sklad a s ním související analytické platformy pět let. Cílem výstavby datového skladu je vytvoření manažerského informačního systému, který z kvalitní zpracování a analýzy dat a který se stane základním informačním zdrojem pro organizační i regionální management. Jak je patrné z následujících částí článku, současný stav řešení již splňuje základní nároky kladené na manažerský informační systém. Základní směry budoucího rozvoje datového skladu kraje Vysočina je možno shrnout do následujících bodů:

#### Analýza nestruturovaných dat

V širším pojetí je možno do nástrojů BI zařadit i problematiku analýzy nestruturovaných dat. Do této oblasti patří především analýza textů, indexace a sledování vybraných webů, monitoring médií a v neposlední řadě i souborů lokálně uložených na počítači či v síti instituce.

#### Integrace s geografickými informačními systémy

Cílem je především jednotná datové základna a vzájemné sdílení dat obou platform. V budoucnu je možné uvažovat o elektronickém profilu kraje Vysočina či elektronickém atlasu Vysočiny, který bude obsahovat nejen popisné části, ale i určitou formu srovnání s ostatními kraji v ČR a v některých indikátorech i s evropskými regiony.

Spolupráce v oblasti výměny a využití dat mezi zainteresovanými subjekty v regionu

V současné době je již možno hovořit o množině regionálních dat na všech územních úrovních, která datový sklad jednak obsahuje a jednak jsou vytipována a zmapována pro budoucí zavedení do datového skladu. Analytická vrstva je tedy vhodným zázemím pro tvorbu standardizovaných setů regionálních dat, využitelných nejen pro vzájemné porovnávání krajů, ale například i pro další instituce, zabývající se regionální analýzou.

### Kvantitativní i kvalitativní posun řešení

Vzhledem k rozsahu krajem zpracovávaných agend nebude manažerský informační systém z pohledu obsahu nikdy dokončen. V další fázi budou zavedena již zmapovaná data a řešení upgradováno na modernější databázovou platformu.

### Přínosy řešení

Datový sklad kraje obsahuje v analytické vrstvě více než sto multidimenzionálních objektů, jejichž aktualizace je pravidelně zajišťována. Podařilo se navázat spolupráci s vybra-

nými institucemi veřejné správy a systematizovat předávání vybraných dat (regionální zastoupení ČSÚ v Jihlavě, krajské ředitelství Policie ČR v Jihlavě). K dispozici je tak poměrně široká datová základna, která může být využita jednak v běžném provozu kraje, ale i ve strategickém plánování kraje. Otevírají se možnosti využití dat nejen standardně publikovaných, ale i interně získávaných a pořizovaných, např. při hodnocení efektivity krajem zajišťovaných veřejných služeb, rozpočtu kraje a jejich vazbě na rozvojový potenciál kraje.

Důležitým přínosem řešení je i zefektivnění vlastního procesu zpracování a analýzy dat uvnitř krajského úřadu kraje. Za nejdůležitější přínosy v této oblasti považujeme:

- centrální uložení dat v pevné a logické struktuře,
- snadnou interní i externí online dostupnost dat,
- úspora času při zpracování dat,
- pravidelná aktualizace dat ve stanovených intervalech,
- zvýšení vypovídající hodnoty a kvality dat,
- možnost spojování dat provozních systémů a realizaci křížových dotazů,
- eliminace duplicitního zpracování dat několika pracovníky,
- jednotnost datových výstupů kraje a zvýšení otevřenosti směrem k veřejnosti.

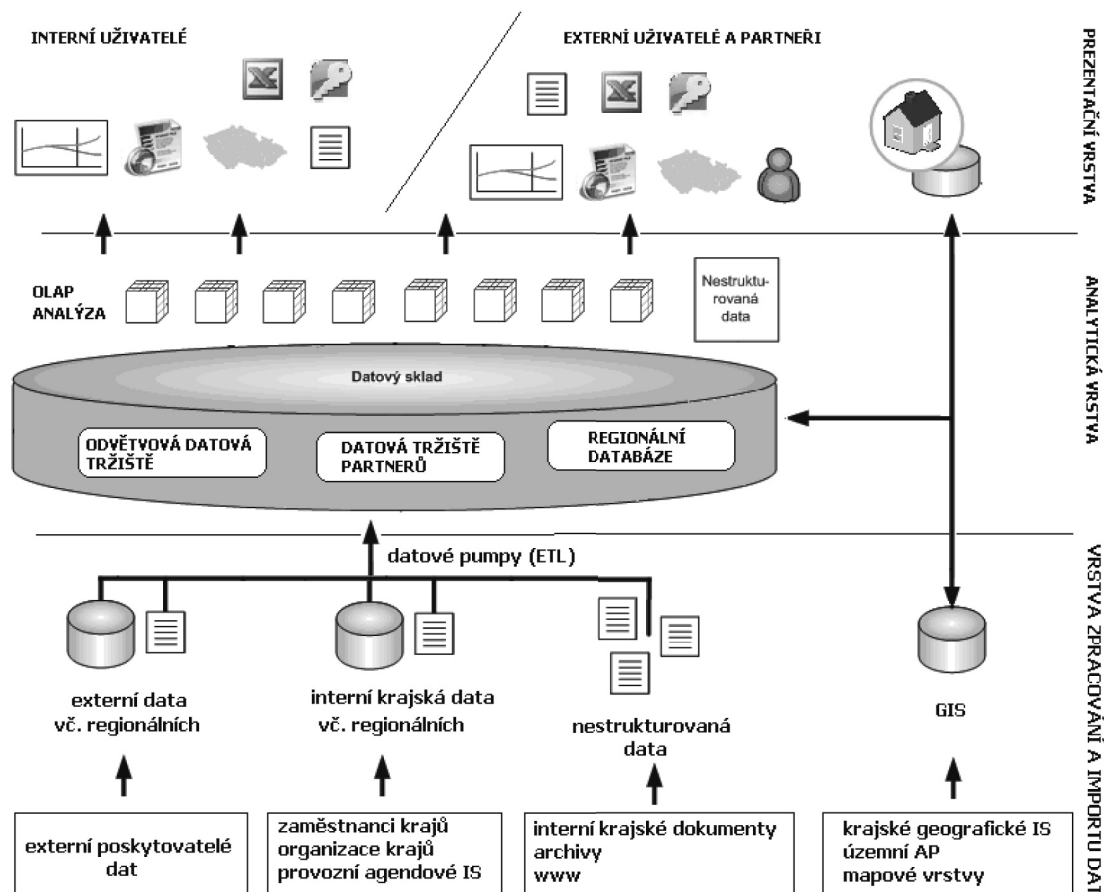


Schéma č. 1: Schéma datového skladu kraje – cílový stav

Zdroj: autoři

Datová oblast (datové tržiště)	Zdroje dat
1. Statistika	ČSÚ, MPSV ČR, ČNB, Eurostat, MF ČR
2. Fond Vysočiny	Krajský úřad
3. Integrovaný záchranný systém	Krajský úřad, Policie ČR, ZZS a HZS kraje Vysočina
4. Příspěvkové organizace	Krajský úřad, příspěvkové organizace, ÚIV
5. Ekonomika a rozpočet kraje	Krajský úřad
6. Provoz	Krajský úřad
7. Obce a dotační tituly	ČSÚ, Krajský úřad
8. Neziskové organizace	Krajský úřad, neziskové organizace, ČSÚ
9. Metadata	Krajský úřad

Tabulka č. 1: Rámcový přehled obsahu a datových zdrojů Datového skladu kraje Vysočina

Zdroj: autoři

## Hlavní problémy řešení

Dosavadní výstavba manažerského informačního systému založeného na BI ukázala i praktické problémy implementace metody. Jedná se jednak o problémy na straně technologie, ale zejména o problémy na straně lidského faktoru.

Hlavním technologickým problémem je v současné době stará databázová platforma a s ní spojená bezpečnostní rizika. Zastarávání technologie znamená obecně nutnost migrace na novější verze zvolené platformy a tedy i finanční náklady instituce, zejména při přechodu na vzájemně nekompatibilní platformy či nekompatibilní verze jedné platformy. Další problémy spojené s technologií, např. nefunkčnosti či havárie systému, jsou řešeny za běžného provozu a jsou srovnatelné s problémy běžného kancelářského software.

Největším problémem na straně lidského faktoru je ignorace systému některými zaměstnanci instituce. Objektívni důvody nezájmu, např. zdánlivá složitost systému či nevhodnost struktury výstupu, jsou řešeny osvětou a maximálním přizpůsobením výstupů potřebám uživatelů. To však vyžaduje pracovní kapacitu na straně instituce. Vhodným řešením je i možnost jednoduchého vyhledávání konkrétních dat a výstupů v metadatovém systému. Subjektívni důvody jsou řešeny běžnými manažerskými prostředky.

Kritickým místem BI je kvalita a aktuálnost dat vstupující do vrstvy zpracování a importu. Tento problém má několik aspektů:

- pokud se formát a struktura předávaných vstupních dat mění, což je bohužel poměrně obvyklé, je nutné přeprogramování datové pumpy,
- pokud se mění obsah vstupních dat, např. vypuštění některých ukazatelů či přidání jiných, jsou nutné změny až do úrovně prezentační vrstvy,
- mění-li se objektívni metodika zjišťování dat a tedy vy-

povídací schopnost a časová řada přebíraných dat, jsou nutné podstatné změny až do úrovně prezentační vrstvy,

- neaktuální data znamenají konec využívání každého informačního systému ze strany uživatelů.

Nikoli nepodstatným problémem je přehlednost uživatelského rozhraní systému, tedy zejména způsob logického uspořádání výstupů a vlastní vizualizace dat. Velké množství zahrnutých agend a dat klade vysoké nároky zejména na přiřazení jednotlivých datových i funkčních oblastí oprávněným uživatelským skupinám. Všechny datové zdroje, objekty a výstupy jsou proto opatřeny metadatovým rámcem, který uživateli umožňuje potřebná data dohledat. Dodržován je rovněž princip „one page managementu“, kdy ze základního jednostránkového výstupu je možné se intuitivně dostat na detailní přehled.

## Příklad využití BI

Jako příklad platformy využití BI v regionální analýze uvádíme dynamický model komparace krajů (Vysočina, 2008). V modelu je srovnáváno celkem 12 indikátorů ve třech oblastech – ekonomické výkonnosti, inovačním potenciálu a kvalitě života.

Pro každý indikátor je sestaven datový algoritmus pro určení pořadí krajů ve statické složce (výkonnost v posledním známém roce) a dynamické složce (změna ukazatele mezi rokem 2000 a posledním známým rokem) která jsou následně vizualizována v bodovém grafu včetně agregací (metodou průměru přidělených bodů podle pořadí) do tří vyšších úrovní a celkového umístění krajů (čím více bodů, tím lepší výsledek).

Po aktualizaci příslušných dat v analytické vrstvě se v souladu se zvoleným způsobem výpočtu zaktualizují i všechny

výstupy, v tomto případě grafy, animace a tabulky na příslušném webu. To je v případě jednoho modelu vcelku zanedbatelná úspora času, avšak v případě několika podobných výstupů tento princip „jedné aktualizace dat“ a „mnoha aktuálních výstupů“ znamená velmi velkou přidanou hodnotu. Již fungující „Statistický portál bezpečnosti“ a „Monitoring nezaměstnanosti“ na analytické stránce kraje (<http://analytika.kr-vysocina.cz>) jsou toho dokladem.

## Závěr

Výstavba funkčního řešení Business Intelligence v podmínkách krajské veřejné správy je dlouhodobou záležitostí. To je dáno především velkou rozmanitostí řešených agend a z ní vyplývajícím potřebným objemu relevantních dat. Řešení není bez běžných provozních problémů, avšak vzhledem k potenciálu, které řešení znamená pro řízení dat a informací v krajích, je nastoupená cesta správná.

Budoucí rozvoj a využití metody má kromě běžného praktického využití na krajských úřadech i jeden důležitý aspekt, kterým je potenciální partnerství institucí v regionu. Jednotná datová základna pro praktickou regionální politiku i výzkumné aktivity realizované např. vysokými školami může podpořit a zvýšit efektivitu jejich spolupráce a cílení jejich činností.

## English summary (abstract)

The article deals with the usage of the data processing and analysing method Business Intelligence (BI) in the Vysocina Region. The BI method is a proper basis for regional data integration and the creation of the managerial informational system of the region.

A short history and the technical framework of the method is described, including main advantages and problems of the current solution. The future development of the method is drafted and the aspect of the future partnership of respective institutions is emphasized.

## Literatura

- CODD, E.F., CODD, S.B., SALLEY, C.T. (1993): Providing OLAP to User-Analysts: An IT Mandate. E.F. Codd and Associates
- INMON, W. (2005): Looking Ahead: Unstructured Data. DM Review, December 2005. Citace 19.3.2010. Dostupné na WWW: <http://www.information-management.com/>
- INMON, W. (2007): Tapping into Unstructured Data: Integrating Unstructured Data and Textual Analytics into Business Intelligence. Prentice Hall, 264 s., ISBN 978-0132360296
- LACKO, L. (2006): Business Intelligence v SQL Serveru 2005: Reportovací, analytické a další datové služby. Brno, Computer Press, 392 s., ISBN 80-251-1110-5
- LUHN, H.P. (1958): A Business Intelligence System. IBM Journal. October 1958
- POWER, D.J. (2007): A Brief History of Decision Support Systems. Version 4.1. DSSResources.COM. Citace 19.3.2010. Dostupné na WWW: <http://DSSResources.COM/history/dshistory.html>
- Vysocina (2008): Dynamická komparace krajů [analytická webová aplikace]. Citace 15.12.2009. Dostupné na WWW: <http://analytika.kr-vysocina.cz/komparace>