



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Název projektu	Rozvoj vzdělávání na Slezské univerzitě v Opavě
Registrační číslo projektu	CZ.02.2.69/0.0./0.0/16_015/0002400

Prezentace předmětu: **Business Intelligence**

Vyučující:
doc. Mgr. Petr Suchánek, Ph.D.



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

Business Intelligence

Přednáška 2



**SLEZSKÁ
UNIVERZITA**

**OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ**

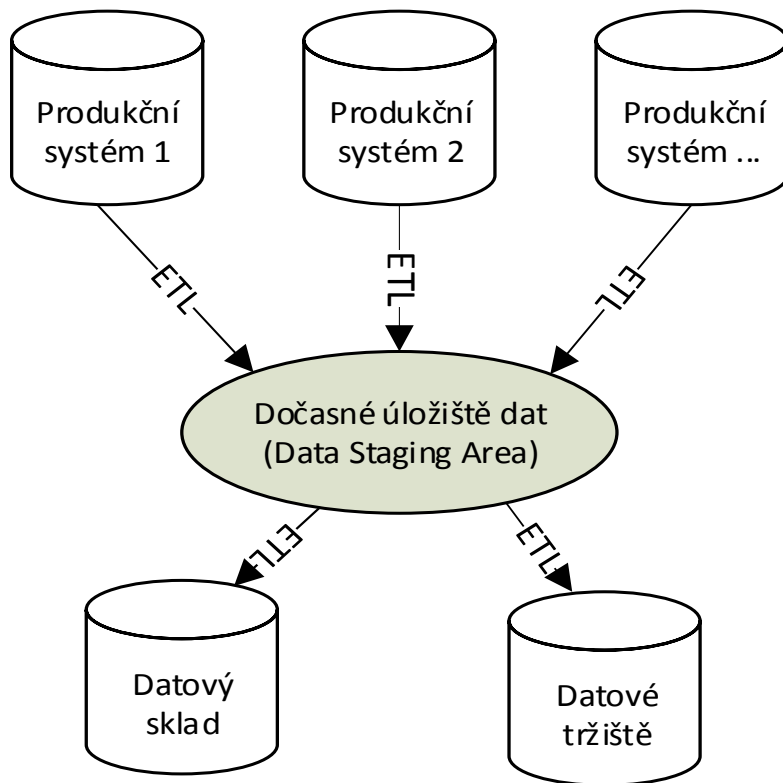
doc. Mgr. Petr Suchánek, Ph.D.

Komponenty BI – produkční a zdrojové systémy

- Produkční zdrojové systémy = produkční zdrojové databáze
 - základní vrstva vstupu pro BI;
 - různorodé zdroje jak z hlediska struktury dat, tak i technologií;
 - mají obvykle transakční charakter a nepatří do skupiny BI aplikací.
 - Vnitřní
 - interní zdroje dat (ERP (Enterprise Resource Planning), CRM (Customer Relationship Management), SCM (Supply Chain Management), soubory v tabulkových kalkulátorech (například MS Excel) apod.).
 - Vnější
 - například statistické údaje statistického úřadu, výstupy vládních institucí, seznamy podnikatelských subjektů, telefonní seznamy, on-line přístupné internetové databáze a zdroje apod.
-

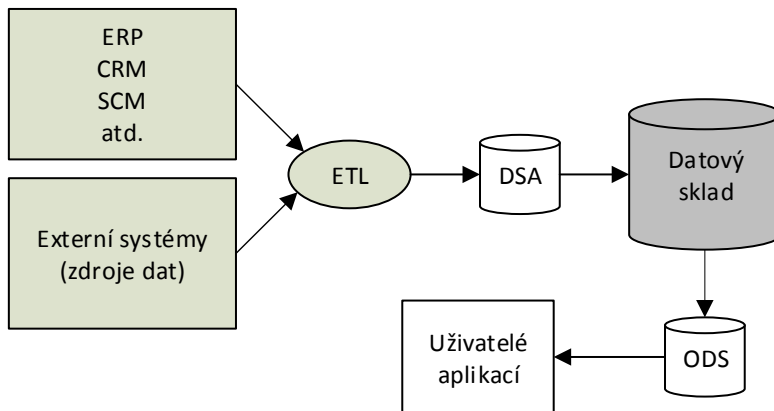
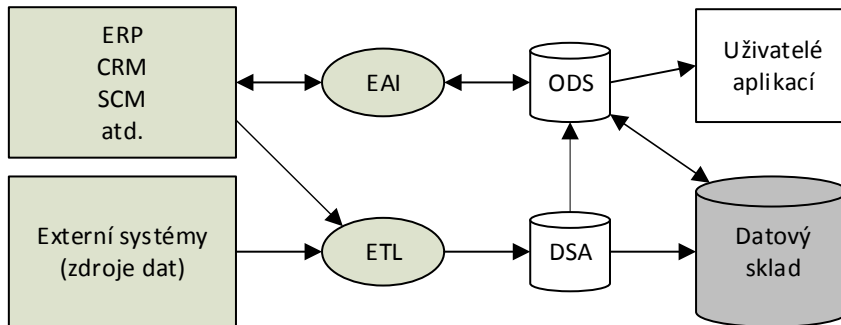
- Dočasné úložiště dat (Data Staging Area, DSA)
 - dočasné uložení extrahovaných dat z produkčních databází s cílem zajistit jejich přípravu a potřebnou kvalitu před vstupem do datového skladu;
 - data jsou:
 - detailní;
 - často nekonzistentní;
 - neagregovaná;
 - bez časové dimenze.
 - DSA obsahuje pouze aktuální data, tj. po jejich zpracování v DSA a přenosu do datového skladu nebo tržiště se z DSA odstraní;
 - transformaci dat provádějí tzv. ETL (Extraction, Transformation, Load) mechanismy, které jsou velmi důležitou součástí celé řady systémů a využívají se v různých okamžicích při převodu dat mezi systémy nebo jejich částmi.

Komponenty BI – dočasná úložiště dat



- Operativní úložiště dat (ODS – Operational Data Store)
 - vrstva, která slouží jako zdroj pro dotazování z konsolidovaných agregovaných dat se snahou minimalizovat dobu odezvy
 - například databáze základních číselníků;
 - rozhraní pro komunikaci se zákazníky;
 - databázi s možností rychlých jednoduchých dotazů nad malým množstvím analytických dat.
- Rozdíl mezi DSA a ODS
 - ODS je přístupné uživatelům a obsahuje konsolidovaná, konzistentní nebo i agregovaná data a jsou součástí analytického procesu.

Komponenty BI – operativní úložiště dat





- ODS může plnit následující role:
 - Datové úložiště pro operativní procesy;
 - Datové úložiště kmenových dat;
 - Sdílená cache více systémů (datově integrovaná nebo neintegrovaná);
 - Podpůrná vrstva datového skladu (stage vrstva);
 - Datové úložiště pro operativní reporting;
 - Náhrada za legacy aplikace (ODS poskytuje data zaniklých systémů).
- Z pohledu frekvence plnění a aktualizace dat:
 - Ad-hoc plnění nebo plnění na vyžádání;
 - Měsíční plnění;
 - Denní plnění;
 - Plnění vícekrát za den;
 - Plnění téměř v reálném čase;
 - Plnění v reálném čase.



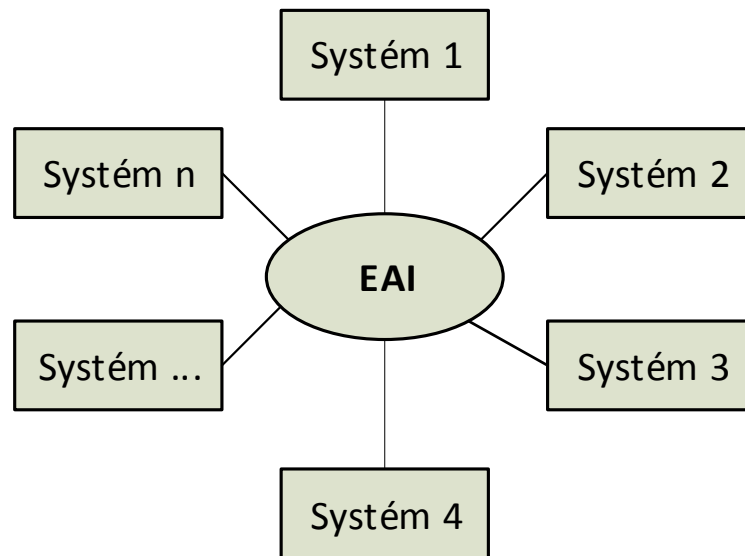
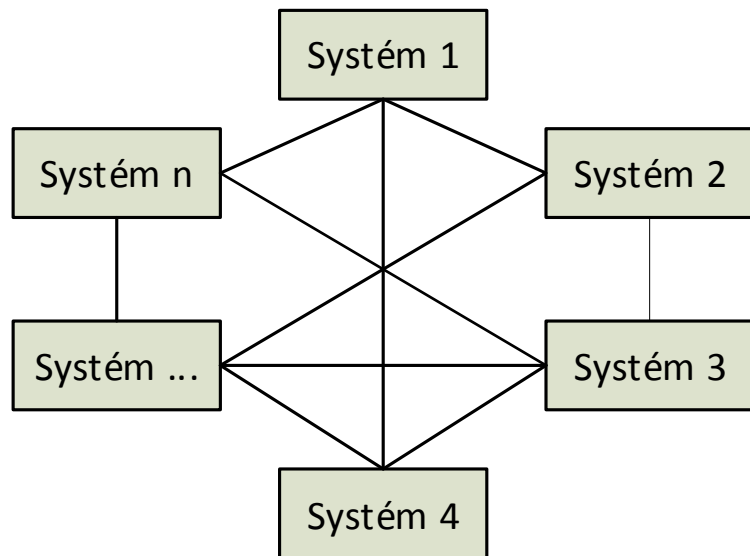
- ODS - způsoby získávání dat
 - Replikace tabulek z primárních systémů;
 - Replikace transakcí z primárních systémů;
 - Integrace transakcí z primárních systémů;
 - Derivace dat z datového skladu.
 - Z hlediska logické architektury
 - Datové úložiště;
 - Transformační vrstva;
 - Komunikační vrstva;
 - Nástroje pro správu a vývoj datového úložiště;
 - Nástroje pro vývoj transformačních a komunikačních metadata;
 - Nástroje pro správu ostatních metadat a dokumentace.
-

Komponenty BI – EAI



- EAI - Enterprise Application Integration
 - Koncept s cílem jedinečné integrace
 - na úrovni dat;
 - na úrovni aplikací.
 - EAI
 - umožňuje přenos dat do různých typů datových úložišť v reálném čase;
 - podpora rozvoje nové generace datových skladů označovaných jako Real-time Data Warehouse;
 - umožňuje vyšší nezávislost na nižších vrstvách informačního systému (operační systémy, databáze, SW jiných dodavatelů);
 - Využívání universálních technických standardů, jazyků a protokolů (XML, BPEL, SOAP, WSDL, JDBS, JMS, CORBA a další);
 - využívání universálních integračních technologií a řešení (J2EE, Bi-zTalk a další).
-

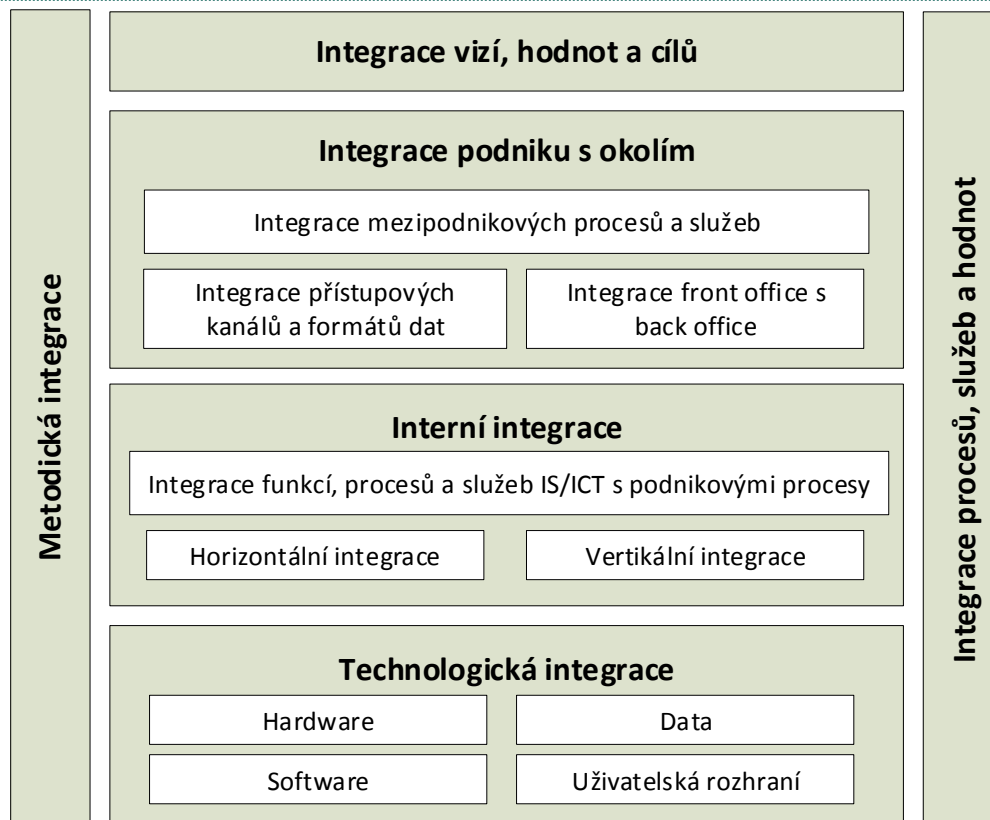
Komponenty BI – EAI



Komponenty BI – Koncepce integrace IS/ICT v podniku



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ



- Datový sklad (Data Warehouse)
 - integrovaný, subjektově orientovaný, stálý a časově rozlišený souhrn dat uspořádaný pro účely efektivního zpracování dat a jejich přeměny na informace pro podporu rozhodování.
- Data jsou uložena podle typu bez vazby na aplikace, ve kterých byla pořízena;
- Centrální úložiště dat z celého podniku, která se nemění a která jsou časově označena;
- Ve vztahu k datovým tržištím je mnohdy datový sklad pojímán jako množina datových tržišť.

- OLAP (Online Analytical Processing)
 - technologie uspořádání rozsáhlých dat formou krychlí tak, aby k nim byl zajištěn co nejrychlejší přístup za účelem tvorby datových analýz a kontingenčních objektů (tabulky, grafy);
 - součástí OLAP je i podpora analytických nástrojů;
 - zdrojovými daty technologie OLAP jsou databáze OLTP (Online Transactional Processing) (běžně ukládány v datových úložištích).

- ROLAP (Relační OLAP)
 - Agregovaná data jsou uložena ve speciálních tabulkách v rámci relační databáze. Analýza je náročná, protože požadované struktury se vytvářejí pomocí SQL dotazů, které vybírají příslušná data z relacemi propojených tabulek. Pracuje nad relačními databázemi.
 - Výhoda: flexibilita
 - Nevýhoda: nároky na výkon databáze, pomalejší analýza.
 - MOLAP (Multidimenzionální OLAP)
 - Vícerozměrný spreadsheet, datová krychle, hierarchické dimenze, schéma STAR, SNOWFLAKE, FACT CONSTELLATION. Významně rychlejší než ROLAP. Pracuje nad datovými sklady nebo tržišti.
 - Výhoda: rychlost analýzy
 - Nevýhoda: nároky na diskový prostor, malá flexibilita (data se musí transformovat do datového skladu).
-

- HOLAP (Hybridní OLAP)
 - Analýza probíhá nad relačními databázemi, ale tabulky s agregovanými daty se ukládají do datových skladů.
 - Výhody: přístup k rozsáhlým datům při současně rychlé agregaci.
 - Nevýhody: nutnost udržovat data na dvou místech.
- DOLAP (Dynamický OLAP)
 - Multidimenzionální kostka je během zpracování vytvořena v paměti RAM.
 - Výhoda: flexibilita
 - Nevýhoda: RAM paměť dostatečně velká a rychlá (vysoké nároky); nutnost opětovného vytváření kostky pro každou analýzu znovu

Komponenty BI – reporting



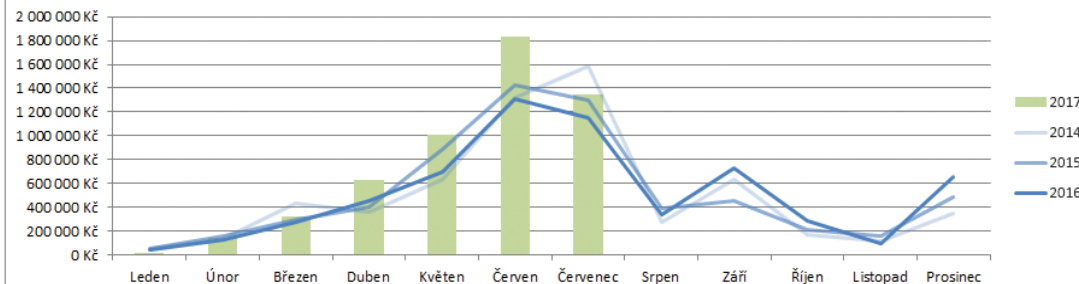
- Reporty
 - základní výstupy BI;
 - vytvářejí se na základě dotazů prostřednictvím analytických aplikací;
 - reporty mohou být
 - standardní - většinou předem definované, které se v daných časových intervalech opakují;
 - ad-hoc reporty - které představují jednorázové požadavky.
 - reporty mohou být určeny
 - pro vnitřní prostředí (zaměstnanci);
 - pro vnější prostředí (externí subjekty nebo cílové skupiny - spolupracující podniky, dodavatelé, odběratelé, banky jako věřitelé, státní orgány, apod.).
 - reporty - vesměs grafy, tabulky, číselné seznamy, souhrny, apod.
-

Komponenty BI – reporting*

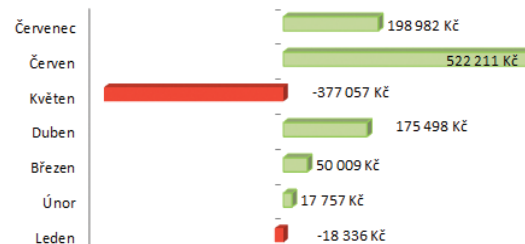


**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVÍNĚ

Vývoj tržeb

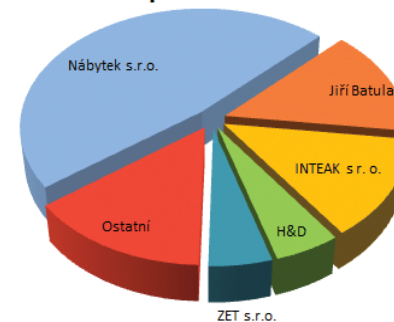


Plnění plánu 2017



Zisk	2012							Celkem
	M1	M2	M3	M4	M5	M6	M7	
ACO nábytkové prvky s.r.o.	1 951 Kč	13 171 Kč	73 832 Kč	146 342 Kč	1 912 Kč	13 171 Kč	77 695 Kč	172 685 Kč
AK - Media a. s.					15 002 Kč	49 440 Kč	15 002 Kč	49 440 Kč
Argo, a. s.			24 281 Kč	48 451 Kč			24 281 Kč	48 451 Kč
H&D	2 232 Kč	13 440 Kč	49 038 Kč	49 440 Kč	69 727 Kč	148 320 Kč	97 229 Kč	211 200 Kč
INTEAK spol. s r. o.			25 270 Kč	49 440 Kč	125 528 Kč	494 400 Kč	150 797 Kč	543 840 Kč
Jaromír Novák - Nábytek	2 824 Kč	25 267 Kč	45 696 Kč	92 947 Kč	1 374 Kč	12 634 Kč	49 894 Kč	130 848 Kč
Jiří Batula			101 079 Kč	197 760 Kč	150 531 Kč	395 520 Kč	251 611 Kč	593 280 Kč
Nábytek s.r.o.	4 781 Kč	26 880 Kč			605 992 Kč	1 928 160 Kč	610 773 Kč	1 955 040 Kč
Ralf Schneider			25 269 Kč	49 440 Kč	-28 202 Kč	2 050 Kč	-2 933 Kč	51 489 Kč
Vladimír Šimek - JIPO					23 128 Kč	96 902 Kč	23 128 Kč	96 902 Kč
ZET s.r.o.	4 440 Kč	26 880 Kč			44 016 Kč	175 200 Kč	48 456 Kč	202 080 Kč
neuveдено	5 381 Kč	38 707 Kč	0 Kč	0 Kč	0 Kč	0 Kč	5 381 Kč	38 707 Kč
Celkový součet	21 610 Kč	144 346 Kč	320 697 Kč	633 820 Kč	1 009 008 Kč	3 315 797 Kč	1 351 315 Kč	4 093 963 Kč

Top 5 zákazníků 2016



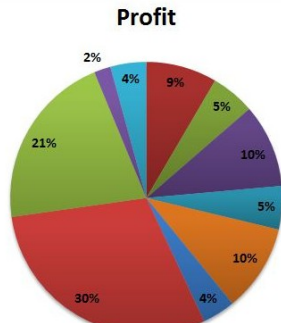
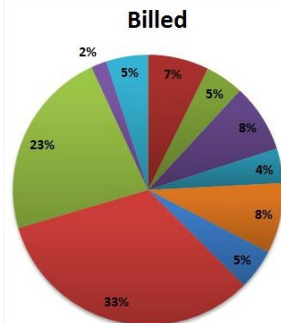
*<https://www.stormware.cz/pohoda/business-intelligence/>

Komponenty BI – reporting*

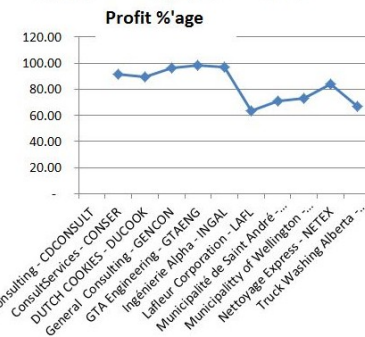
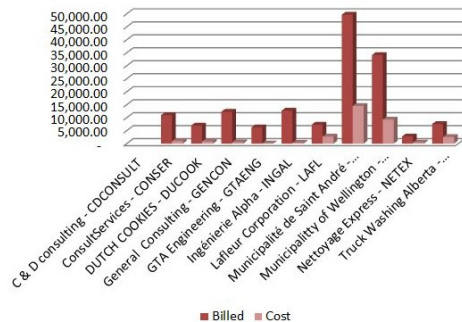


R35

All Trans Years by Month



	Billed	Cost	Profit	Profit %
C & D consulting - CDCONSULT				
ConsultServices - CONSER	10,957.93	958.25	10,011.32	91.36
DUTCH COOKIES - DUCOOK	6,934.40	827.37	6,194.93	89.34
General Consulting - GENCON	12,266.80	533.88	11,785.41	96.08
GTA Engineering - GTAENG	6,217.00	103.09	6,120.44	98.45
Ingénierie Alpha - INGAL	12,664.75	349.54	12,323.65	97.31
Lafleur Corporation - LAFL	7,292.94	2,712.05	4,637.84	63.59
Municipalité de Saint André - STANDR	49,665.79	14,517.29	35,337.63	71.15
Municipality of Wellington - WELLING	34,106.24	9,241.41	24,929.64	73.09
Nettoyage Express - NETEX	2,767.50	434.00	2,333.50	84.32
Truck Washing Alberta - TRUCKWA	7,560.00	2,556.47	5,064.80	66.99
All Client Projects	150,433.35	32,233.35	118,739.16	78.93



*<http://www.bi-metrix.com/isvabak.php>

- EIS - Executive Information System
 - analytické aplikace - cílem je poskytování informací pro podporu rozhodování, vesměs v podobě reportů;
 - obsahují nástroje pro monitorování podnikových procesů a měření jejich efektivity i efektivity výkonnosti celého podniku;
 - vstupem jsou data z datového skladu;
 - požadavek na
 - uživatelskou přívětivost těchto aplikací;
 - snadnou editaci vstupních parametrů jednotlivých dotazů.
 - standardem je grafické uživatelské rozhraní;
 - postupem času se tyto aplikace přesouvají z nejvyšších úrovní řízení na úroveň operativ-ní s primárním využitím pro specialisty v jednotlivých oborech (marketing, poradenství, apod.).
-

Komponenty BI – dolování dat



- Dolování data – Data mining
 - procesy získávání předem neznámých nebo uživatelem nedefinovaných informací z datových skladů nebo rozsáhlých databází;
 - orientuje na obsah dat a jeho výstupy jsou zcela nové dosud neznámé informace a závislosti;
 - využitelný jak pro získání nových deskriptivních informací, tak i pro prediktivní informace;
 - slouží pro
 - podporu identifikace klíčových faktorů podnikání;
 - ověřování resp. testování hypotéz;
 - objevování korelací mezi různými typy dat;
 - apod.
 - využívá celou řadu technik matematických a statistických metod (neuronové sítě, genetické algoritmy, rozhodovací stromy, clustering, klasifikace apod.).
-

Komponenty BI – dolování dat



- Data mining (DM) se používá především v oblastech, kde se sbírá velké množství dat. Typickými příklady obrovských datových souborů jsou například:
 - údaje o klientech;
 - pohyby na účtech (bankovníctví);
 - údaje o volání (telefonní operátoři);
 - informace o tom, jak lidé nakupují (obchodní řetězce a internetové obchody);
 - pohyb uživatelů na internetu, datové informace o expresi genů (genetika);
 - provozní záznamy zaznamenávají průběh provozních parametrů (průmysl).
 - DM není výsadou pouze velkých společností, ale tyto postupy mají nemalý potenciál i v malých a středních podnicích.
-

Komponenty BI – oborová znalost



- BI systémy jsou implementovány správně, pokud se přesně ví, co se od nich očekává;
- Očekávání přitom mohou být různá v různých oborech lidské činnosti;
- Je proto zcela nezbytné, aby systém BI byl vždy implementován na míru výsledným potřebám;
- S rozvojem v této i jiných oblastech samozřejmě existuje a neustále vzniká celá řada použitelných vzorů vyžadujících stále menší úpravy.

Komponenty BI – kvalita dat



- Předpokladem kvalitních datových výstupů je kvalita vstupních dat;
- Nutné vlastnosti dat pro tvorbu analýz:

Vlastnost dat	Úkol nástrojů připravujících data
Úplnost	Nalézt chybějící data a vyřadit data nepoužitelná.
Soulad	Zajistit jednotný formát dat.
Konzistence	Zpracovat data, která jsou v rozporu (konfliktu).
Přesnost	Data musí být aktuální.
Unikátnost	Zajištění minimální nebo lépe nulové duplicity (neredundantní).
Integrita	Zpracování dat postrádajících důležité vztahy s ostatními daty.

Komponenty BI – kvalita dat



- Kvalita dat v datových skladech a BI je dána
 - kvalitou dat z nichž se data čerpají (transakční systémy);
 - kvalitou dat v externích souborech a číselnících převzatých z jiných organizací a institucí (ČSÚ, ČNB, aj.);
 - kvalitou přenosových a transformačních procesů (ETL), včetně jejich procesního zabezpečení.
 - Permanentními potenciálními narušovatelí jsou změnová řízení ať již formou projektů nebo jen formou dílčích uživatelských požadavků (i například jen opravy v transakčních systémech nebo číselnících).
 - Typy chyb:
 - nestejná granularita;
 - překlepy při pořizování dat;
 - chybná metadata;
 - implicitní a explicitní hodnoty „Null“;
 - smíšení formátu struktur;
 - transformační chyby;
 - překrývání (overloaded) atributů.
-

- Metadata
 - informace o jiných datech (katalogový lístek v knihovně);
 - jsou přímo vázaná na principy, funkce a obsah.
- V rámci systémů
 - metadata popisující veškeré části integrovaných systémů a jejich obsahu.
- V rámci BI představují popis:
 - datových modelů;
 - funkcí;
 - transformačních pravidel;
 - reportů;
 - dalších.

Komponenty BI – správa metadat



- Typy metadat

Typ metadat	Charakteristika
Metadatum zdrojových systémů	Data vázaná ke zdrojovým datům (význam vstupních dat, využití vstupních dat v rámci BI systémů, pochopení a využití procedur pro zpracování dat a funkcionalitu zdrojových systémů).
Metadatum zdrojových komponent	BI nástroje jsou orientované na obsah. Na druhé straně ale existují možnosti jejich dalšího rozšíření a rozvoje, které by vždy mělo vykazovat charakteristiku komplexního, integrovaného a adekvátním způsobem funkčního celku.
Metadatum datové pumpy	Jsou vázaná a popisují původ dat.
Metadatum uživatelské vrstvy	Definice reportů a příslušných komponent uživatelského interface a agregace původu jednotlivých dat.

Komponenty BI – decision support system (DSS)



- DSS
 - systémy používané na středních nebo nižších úrovních řízení pro rozhodování při běžných činnostech;
 - vývoj DSS započal na počátku 70. let;
 - nabízejí uživatelům rozhodnutí na základě vytvořených modelů;
 - nepracují pouze s daty, které načtou z datových skladů, ale umožňují uživatelům vytvářet i vlastní rozhodovací modely;
 - poskytují uživateli nabídky řešení a případně kladením dotazů usměrňují jeho postup;
 - výstupem jsou obvykle periodické a speciální zprávy či výsledky matematických simulací;
 - DSS systémy prošly vývojem, ve kterém lze identifikovat 5 generací.

Komponenty BI – decision support system (DSS)

- Charakteristika jednotlivých generací vývoje DSS

Generace DSS	Charakteristika
První	Systemy využívající jednoduché multikriteriální metody, jednoduchá uživatelská rozhraní, jednoduché struktury databází a vesměs šlo o systémy orientované na jeden konkrétní problém.
Druhá	Systemy obsahující zdokonalený návrh databází a jejich řízení (SŘBD – Systém řízení báze dat, anglicky DBMS – Database Management System).
Třetí	Systemy s podporou komunikačního rozhraní mezi systémem a uživateli.
Čtvrtá	Modulární systémy s novými přístupy k jejich řízení.
Pátá	Systemy využívající prvky umělé inteligence pro zlepšení komunikace a konzultací.

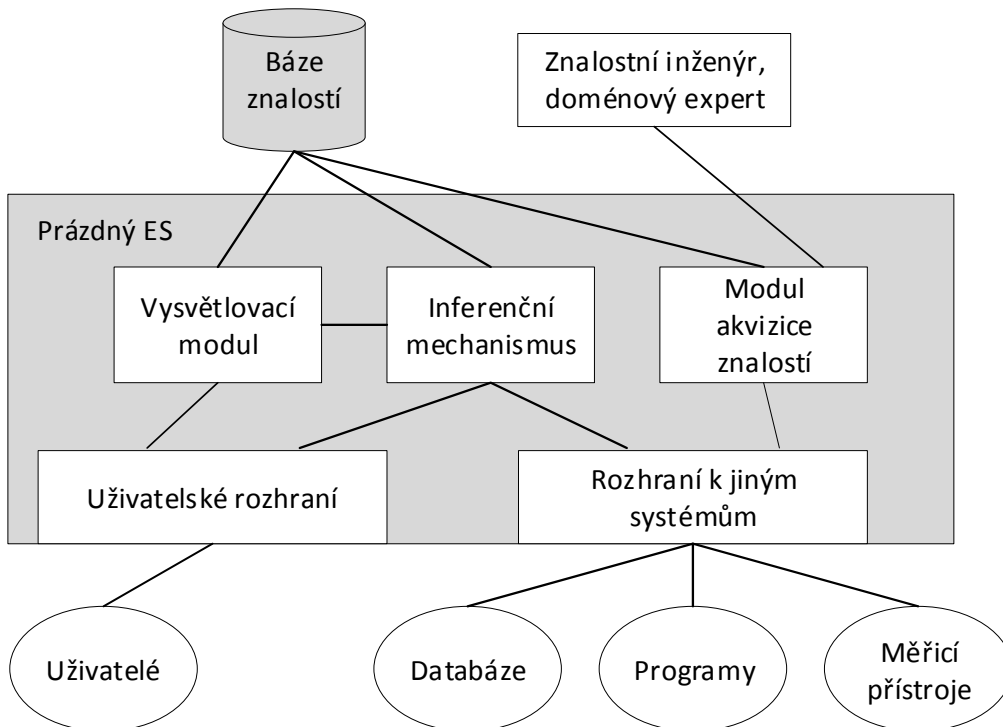
Komponenty BI – expertní systémy (ES)



- ES
 - expertní systémy jsou počítačové programy, simulující rozhodovací činnost experta při řešení složitých úloh a využívající vhodně zakódovaných, explicitně vyjádřených znalostí, převzatých od experta, s cílem dosáhnout ve zvolené problémové oblasti kvality rozhodování na úrovni expert;
 - jejich základem je báze znalostí;
 - v obecné rovině je cílem expertních systémů dosažení co nejlepší odezvy na vstupní, reálná data;
 - Základní charakteristiky:
 - oddělení znalostí a mechanismu jejich využívání;
 - rozhodování za stavu neurčitosti;
 - schopnost vysvětlování.
-

Komponenty BI – expertní systémy (ES)

- Základní složky ES



- **Báze znalostí:**
 - obsahuje znalosti z určité domény a specifické znalosti o řešení problémů v této doméně.
- **Báze faktů**
 - se vytváří v průběhu řešení konkrétního problému a obsahuje data k řešenému problému.
- **Inferenční mechanismus**
 - obsahuje obecné (doménově nezávislé) algoritmy schopné řešit problémy na základě manipulace se znalostmi z báze znalostí.

Komponenty BI – expertní systémy (ES)



- Podle obsahu báze znalostí můžeme expertní systémy rozdělit na:
 - problémově orientované
 - jejichž báze znalostí obsahuje znalosti z určité domény.
 - prázdné (shells)
 - jejichž báze znalostí je prázdná.
 - diagnostické
 - úkolem je určit, která z hypotéza z předem definované konečné množiny cílových hypotéz nejlépe koresponduje s daty týkajícími se daného konkrétního případu.
 - plánovací
 - obvykle řeší takové úlohy, kdy je znám cíl řešení a počáteční stav a je třeba s využitím dat o konkrétním řešeném případě nalézt posloupnost kroků, kterými lze cíle dosáhnout.
-

Komponenty BI – expertní systémy (ES)



- Typické kategorie způsobů použití expertních systémů:
 - Konfigurace
 - sestavení vhodných komponent systému vhodným způsobem;
 - Diagnostika
 - zjištění příčin nesprávného fungování systému na základě výsledků pozorování;
 - Interpretace
 - vysvětlení pozorovaných dat;
 - Monitorování
 - posouzení chování systému na základě porovnání pozorovaných dat s očekávanými;
 - Plánování
 - stanovení posloupnosti činností pro dosažení požadovaného výsledku.
-

Komponenty BI – expertní systémy (ES)



- Typické kategorie způsobů použití expertních systémů:
 - Prognózování
 - předpovídání pravděpodobných důsledků zadaných situací.
 - Ladění
 - sestavení předpisu pro odstranění poruch systému.
 - Řízení
 - regulace procesů (může zahrnovat interpretaci, diagnostiku, monitorování, plánování, prognózování a ladění).
 - Učení
 - inteligentní výuka při níž studenti mohou klást otázky např. typu proč, jak, co kdyby.
-

- NOVOTNÝ, O., POUR, J. a D. SLÁNSKÝ, 2005. *Business Intelligence – Jak využít bohatství ve vašich datech*. Praha: Grada. ISBN 978-80-247-6685-0.
- LABERGE, R., 2012. *Datové sklady – Agilní metody a business intelligence*. Praha: Computer Press. ISBN 978-80-251-3729-1.
- <http://www.adastra.cz/ict-reseni/ods>
- http://www.statsoft.cz/file1/PDF/newsletter/2014_02_26_StatSoft_Uvod_do_data_miningu.pdf
- https://is.mendelu.cz/eknihovna/opory/zobraz_cast.pl?cast=21856
- <http://www.uai.fme.vutbr.cz/~jdvorak/Opory/ExpertniSystemy.pdf>



Děkuji za pozornost

Otázky?
