

VYSOKÁ ŠKOLA EKONOMIE A MANAGEMENTU

DIPLOMOVÁ PRÁCE

2011

ING. KAREL JIRÁNEK

VYSOKÁ ŠKOLA EKONOMIE A MANAGEMENTU

Nárožní 2600/9a, 158 00 Praha 5

DIPLOMOVÁ PRÁCE

MASTER OF BUSINESS ADMINISTRATION

Vysoká škola ekonomie a managementu

+420 841 133 166 / info@vsem.cz / www.vsem.cz

VYSOKÁ ŠKOLA EKONOMIE A MANAGEMENTU

Nárožní 2600/9a, 158 00 Praha 5

NÁZEV DIPLOMOVÉ PRÁCE

Datová kvalita, integrita a konsolidace dat v BI společnosti PREdistribuce, a.s..

TERMÍN UKONČENÍ STUDIA A OBHAJOBA (MĚSÍC/ROK)

říjen 2011

JMÉNO A PŘÍJMENÍ / STUDIJNÍ SKUPINA

Karel Jiránek / MBA 23

JMÉNO VEDOUCÍHO DIPLOMOVÉ PRÁCE

Doc. Ing. Jan Pour, CSc.

PROHLÁŠENÍ STUDENTA

Prohlašuji tímto, že jsem zadanou diplomovou práci na uvedené téma vypracoval/a samostatně a že jsem ke zpracování této diplomové práce použil/-a pouze literární prameny v práci uvedené.

Datum a místo:

podpis studenta

PODĚKOVÁNÍ

Rád bych tímto poděkoval vedoucímu diplomové práce za metodické vedení a odborné konzultace, které mi poskytl při zpracování mé diplomové práce.

Vysoká škola ekonomie a managementu

+420 841 133 166 / info@vsem.cz / www.vsem.cz

VYSOKÁ ŠKOLA EKONOMIE A MANAGEMENTU

**Datová kvalita, integrita a konsolidace
dat v BI společnosti PREdistribuce, a.s.**

Data quality, integrity and consolidation

in BI data for PREdistribuce company

Autor: Ing. Karel Jiránek.

Souhrn

Summary

Klíčová slova:

Keywords:

JEL Classification:

Obsah

1 Úvod.....	1
2 Teoretická část.....	3
2.1 Problémy datové kvality, integrity a konsolidace.....	3
2.2 Master Data Management.....	3
2.3 Data Governance	4
2.4 Business Intelligence	7
3 Návrh systému řízení dat PREdistribuce.....	8
3.1 Procesní analýza PREdistribuce z pohledu dat.....	11
3.1.1	11
3.2 Vytvoření zjednodušeného datového modelu PREdistribuce.....	15
3.2.1 Metadata	16
3.3 Návrh pravidel Master Data Managementu do procesů PREdistribuce	18
3.4 Samotná metodika	19
A Chyba! Záložka není definována.	
3.5 Realizace opatření.....	30
3.6 Vytvoření vstupů pro implementaci BusinessObjects	31
3.6.1 BusinessObjects.....	31
4 Závěr.....	37
Literatura.....	38
Přílohy.....	1

Seznam zkratk

PES	Prvek elektrické soustavy; místo v síti; nemá účetní hodnotu
PTE	Prvek technické evidence (trafo, pojistka, skříň, ...); má výrobní (evidenční) číslo; je na místě v síti nebo ve skladu
TM	Technické místo, objekt v SAP PM, ekvivalentem k PES.
Vybavení	objekt v SAP PM, který je ekvivalentem k PTE.
MDM:	Master Data Management
CDI	Customer Data Integration
PIM	Product Information Management

Seznam tabulek

Seznam grafů

Seznam obrázků

1 Úvod

Společnost PREdistribuce, a.s. (dále jen PREdistribuce), která je provozovatelem distribuční soustavy elektrické energie na území hlavního města Prahy, má svá klíčová data v mnoha systémech - kmenová data sítě jsou uložena jako technická evidence prvků a vybavení v ERP systému SAP (v modulu PM), stejně jako data o zákaznících, smlouvách a zákaznických měřeních (v modulech IS-U a CRM). Topologie distribuční soustavy, tedy umístění jednotlivých prvků a jejich normálové (výchozí) zapojení, je primárně uloženo v systému GIS (od společnosti ESRI). Data potřebná pro řízení distribuční soustavy, jako jsou aktuální hodnoty měření nebo aktuální zapojení, včetně dat získávaných z dálkově řízených prvků pomocí systému SCADA (Supervisory control and data acquisition TELEGYR 8000) jsou uložena v databázi dispečerského řídicího systému NETAN (vytvořeného na míru společností MENTAR). Tyto tři klíčové zdroje dat (SAP, GIS, NETAN) jsou samozřejmě doprovázeny celou řadou menších databází zprostředkovávajících různé doplňkové informace.

Protože jak samotných dat, tak jejich zdrojů je velké množství, lze u nich evidovat všechny obvyklé problémy s daty:

- nepřesnosti (mnohá data nejsou vhodně strukturována)
- neúplnost (data byla historicky naplněna z různých zdrojů, mnoho informací chybí)
- nejedinečnost (některá data jsou duplicitní)
- nekonzistence (stejně objekty mají v různých systémech různě vyplněné atributy)

Tyto problémy značně ztěžují využití dat pro další systémy, zejména pak systémy pro podporu rozhodování managementu (Business Intelligence).

V souvislosti s přípravou na přechod na nový systém pro podporu rozhodování, BusinessObjects od společnosti SAP, které PREdistribuce připravuje v druhé polovině roku 2011, by bylo vhodné zrevidovat data společnosti, zejména ta používaná v manažerských sestavách, určit jejich primární zdroje a zajistit, aby informace byly pokud možno trvale pravdivé a relevantní.

Cílem této práce je tedy připravit na základě teoretických poznatků a analýzy různých metodik Master Data Managementu návrh vhodných abstraktních datových struktur a pravidel pro jejich řízení..

Největším problémem je velké množství dat, se kterými se v PREdistribucestribuci pracuje. Nutnou podmínkou pro další práci je proto abstrakce reálného datového modelu na obecnější datové entity a celky, se kterými bude možné dále pracovat jak v rámci této práce, tak následně v systémech Business Intelligence.

Následným krokem je stanovení důležitosti jednotlivých datových entit. Toho by se mohlo dosáhnout rozdělením dat do celků, odpovídajících tomu, které procesy data vytvářejí, spravují a hlavně využívají. Manažeři odpovědní za tyto procesy by potom měli sami rozhodnout, jak jsou pro ně ta která data důležitá.

Dalším problémem je udržitelnost dat z pohledu jejich pravdivosti a relevance. Z existujících metodik Master Data Managementu by měla být vybrána pravidla, která je možné aplikovat na procesy pořizování, zpracování a správy dat v PREdistribucestribuci a tato pravidla by měla být promítnuta do vnitropodnikových norem, které se daty zaobírají.

2 Teoretická část

2.1 Problémy datové kvality, integrity a konsolidace

2.2 Master Data Management

Master Data Management (dále jen MDM) se zabývá konsolidací klíčových (referenčních, vzorových) dat tak, aby byla jednotná, správná, spolehlivá a aktuální vždy, když jsou používána, kdekoliv v celé společnosti. MDM vznikl jako vyšší forma data warehousingu (datových skladů), ovšem se zaměřením na kvalitu dat a reálnou využitelnost integrovaných dat v operativních úlohách.

MDM se realizuje pomocí:

- technických prostředků
- procesů, lidí, jejich organizace
- znalostí (metodiků, know-how)

Na rozdíl od jiných způsobů správy dat přináší konsolidaci dat s ohledem na jejich obsah. Cílem zavádění MDM v organizaci je:

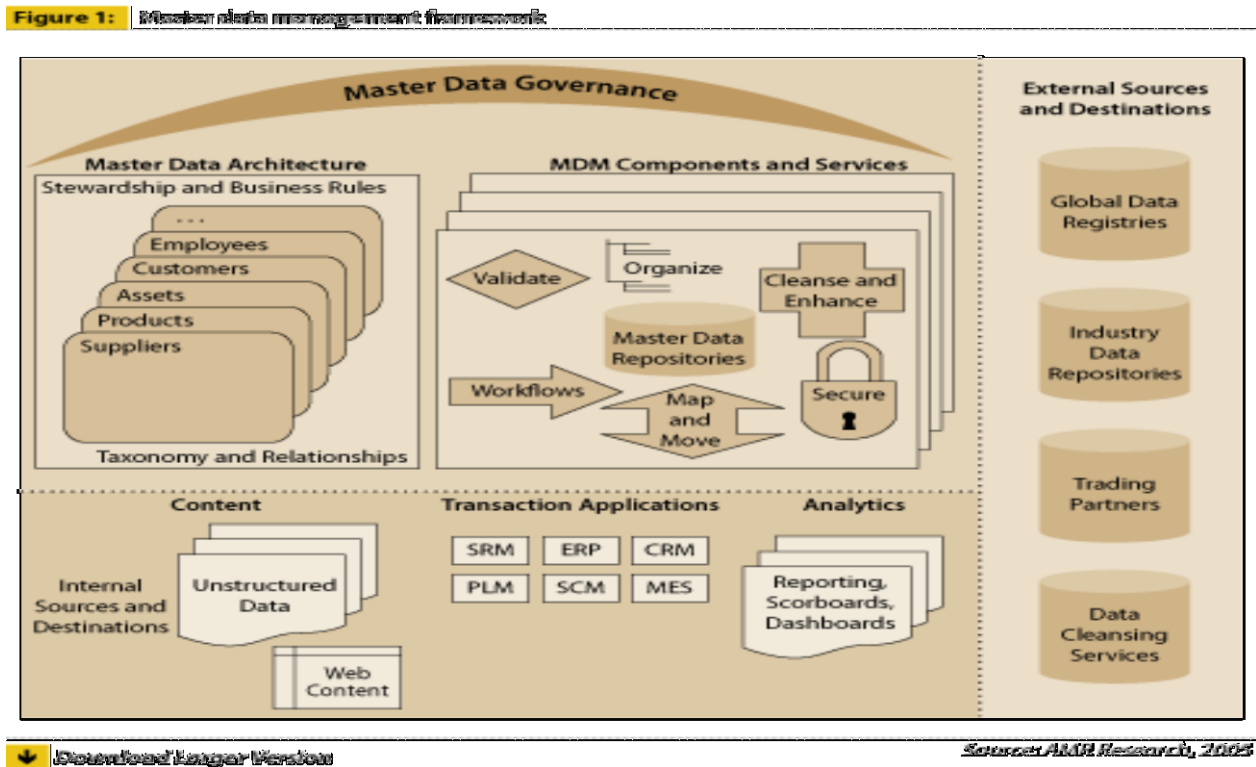
- konsolidace kritických podnikových informací,
- zajištění integrity dat, jejich správnosti, platnosti a úplnosti,
- zjednodušení správy komplexních dat,
- zavedení jednoznačné odpovědnosti za data,
- zvýšení efektivity business rozhodování,
- spolehlivé měření výkonnosti celé organizace.

Klíčové organizační koncepty v MDM jsou:

- implementace pravidel zacházení s podnikovými daty (Data Governance)

- výkonný mechanismus procesu vlastnictví, sdílení a péče o data (Data Stewardship)

Obrázek 1: Framework MDM



Zdroj: AMR Research

2.3 Data Governance

2.3.1 Popis Data Governance

Data Governance (dále jen DG) je akční program pro řízení datové kvality:

- na globální – celopodnikové úrovni,
- se zahrnutím zástupců všech dotčených systémů,
- procházející napříč liniemi, útvary, projekty a iniciativami.

DG je soubor rolí, jejich zodpovědností a procesů, které jsou potřebné pro řízení informační hodnoty uchovávaných dat.

DG pokrývá oblasti:

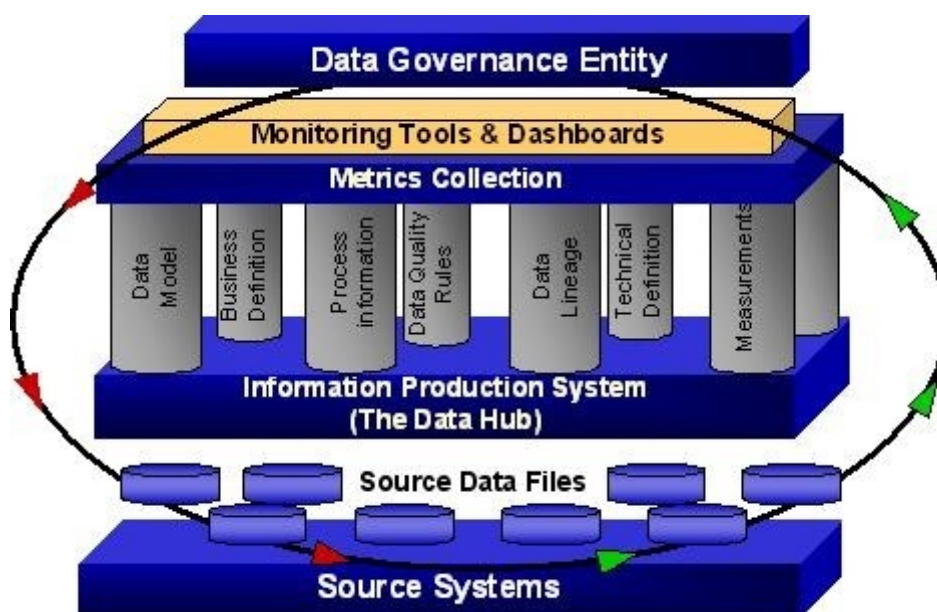
- správy datové architektury z pohledu business potřeb a její provázání na IT architekturu,
- datové integrace,
- správy metadat,
- řízení datové kvality (Data Quality Management).

DG představuje komplexní řízení dostupnosti, použitelnosti, integrity a bezpečnosti dat používaných v organizaci. Spolehlivý DG zahrnuje řídicí výbor nebo komisi a definuje sadu procedur a plán jejich implementace.

Prvním krokem při implementaci DG je určení vlastníků datových aktiv v organizaci. Musí být vytvořena pravidla, která specifikují, kdo je odpovědný za různé množiny a aspekty dat, zahrnující správnost (přesnost), dostupnost, konzistenci, úplnost a aktuálnost.

Definované procesy se musí zabývat jak jsou data skladována, archivována, zálohována a chráněna proti nehodám, krádeži či útoku. Sada pravidel musí stanovit jak data mohou používat oprávněné osoby. Nakonec, sada kontrolních a auditních procedur musí zabezpečit trvalé naplňování legislativních podmínek a regulačních opatření.

Obrázek 2: Schéma DG



Zdroj:

2.3.2 Aktivity programu DG

- Přesahují horizont samotné databáze CCD
- Jejich úspěch závisí na efektivním pokrytí celého životního cyklu dat od okamžiku pořízení v primárních systémech, přes zpracování v CCD, až po prezentaci na front-endu JOK PK a zpětné promítnutí do PS

Požadavky na společnost zavádějící DG

- Zasazení vysoko v organizační struktuře (ředitelé, GŘ)
- Existence sponzora na exekutivní úrovni
- Zavedení procesů, pokrývajících veškeré informační systémy

Klíčové součásti

- Sponzor
- Organizační schéma
- Procesy

Podle tohoto pojetí DG jsou rozlišovány následující klíčové role:

- Sponzor programu DG
- Vlastník procesů DQM (Data Quality Manager)
- Vlastníci dat
- Správci dat (datoví stevardi)
 - Korporátní datový správce (stevard)
 - Doménoví (linioví) datoví správci (stevardi)
 - Systémoví datoví správci (stevardi)
- Odběratelé (konzumenti, beneficianti) dat

2.3.3 Popis skupin DG

Vlastníci dat

- Oddělení nebo skupiny, které produkují data a poskytují je organizaci
- Vlastní „obsah“ dat
- Jsou zodpovědní za definici a kvalitu tohoto obsahu

Datoví správci (stevardi)

- Pověření zástupci, kteří jménem organizace spravují data a řídí související procesy
- Zajišťují plnění SLA na rozsah a kvalitu datového obsahu

Odběratelé dat (beneficienti)

- Konzumenti dat, kteří jejich používáním získávají určitou měřitelnou hodnotu
- Koncoví uživatelé
- Navazující systémy a procesy
- Externí partneři

2.4 Business Intelligence

3 Návrh systému řízení dat PREdistribuce

Společnost PREdistribuce, a.s. (dále jen PREdistribuce), se vyčlenila z Pražské energetiky, a.s. (dále jen PRE), aby se jako její 100% dceřiná společnost stala od 1. 1. 2006 regulovaným provozovatelem distribuční soustavy na území hlavního města, k čemuž je držitelem licence Energetického regulačního úřadu (dále jen ERÚ) na distribuci elektřiny na území hlavního města Prahy, města Roztok a obce Žalov.

Distribuční soustava PREdistribuce je tvořena 202 km vedení na napěťové hladině velmi vysokého napětí 110 kV (dále jen VVN), 3 829 km vedení vysokého napětí 22 kV (dále jen VN) a 7 750 km vedení nízkého napětí 0,4 kV (dále jen NN), což spolu s 22 transformovny, 329 rozpínacími stanicemi a 4 449 transformačními stanicemi dohromady představuje majetek ve výši cca 42 mld. Kč. Prostřednictvím této sítě dodala v roce 2010 PREdistribuce svým odběratelům do 750 000 odběrných míst více než 6 077 GWh elektrické energie (PREdistribuce 2011, s. 7).

Distribuce elektřiny zahrnuje veškeré procesy a činnosti spojené s řízením soustavy, připojováním a odpojováním odběratelů i výroben, rozvojem a obnovou síťových aktiv, provozováním, údržbou a opravami sítě, prováděním odečtů a měření a prodejem distribuce. Tomu všemu se ve společnosti věnuje něco málo přes 500 pracovníků.

V současné době čelí PREdistribuce mnoha novým výzvám:

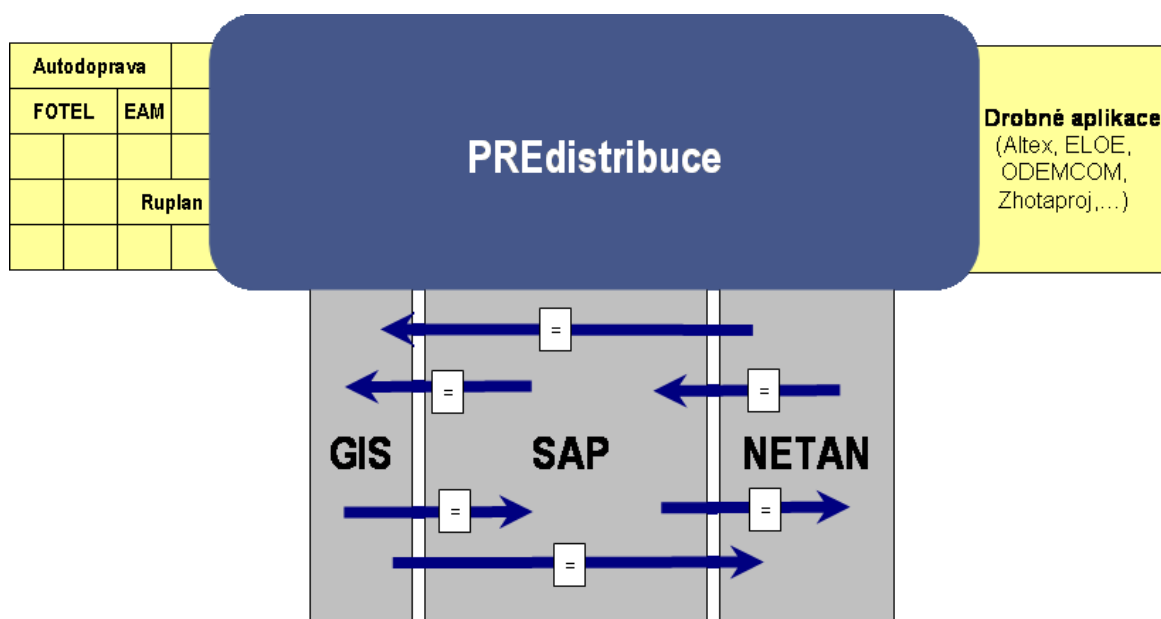
- implementaci nového DŘS
- pilotnímu projektu AMM
- postupnému zavádění myšlenek SmartGrid
- tlaku na zefektivňování investic.

3.1 Popis stávajícího stavu

PREdistribuce společně se svoji matkou Pražská energetika, a.s. (dále jen PRE) má implementovány 3 klíčové systémy. SAP, GIS a Dispečerský řídicí systém (dále jen DŘS).

Tyto systémy se využívají na různých stupních řízení obou firem. Zatímco pro PRE je klíčový SAP, PREdistribuce využívá data ze všech systémů (viz Obrázek 3).

Obrázek 3 Schéma aplikací



Zdroj: Vlastní tvorba autora

Ve společnosti panují obavy, že data uložená napříč systémy jsou v řadě případů zdvojená, nekonzistentní a špatně provázaná. Výsledkem tedy je, že v řadě případů panují pochybnosti o jejich aktuálnosti a relevanci

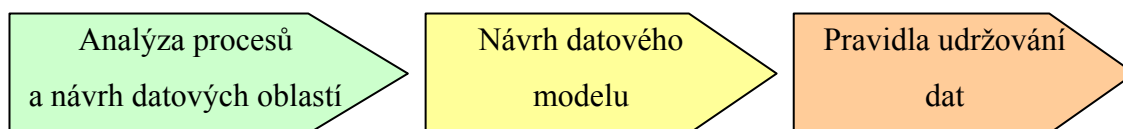
Za tohoto stavu by bylo vhodné vytvořit nad těmito daty metadata (vlastník, důležitost, požadavky na bezpečnost, kvalita, využití, přidaná hodnota, atd.), která by mohla různým stupňům PRE a PREdistribuce zlepšit pohled na přenosovou soustavu a zkvalitnit jejich služby

Cílem praktické části této práce je:

- Navrhnout koncepci řízení dat
- Zohlednit a popsat požadavky na kategorizaci dat (účel použití, priorita atp)
- Specifikovat/navrhnout zobecněnou datovou strukturu metadat
- Specifikovat/navrhnout obecný datový model, který:
 - Bude otevřený pro záznam stávajících i budoucích typů dat
 - Umožní vzájemnou synchronizaci dat mezi klíčovými systémy
 - Bude otevřený k budoucím dynamické práci nad zdrojovými daty dle požadavků jednotlivých částí PRE, PREdistribuce dle zásad unbundlingu a Programu opatření
- Navrhnout systém práce s touto datovou strukturou tak, aby v hlavních systémech byla data vždy aktuální, relevantní a důvěryhodná (respektovat, že se jedná o desítky TB dat)
- Definovat rozhraní a způsob výměny dat s ohledem na skutečnost, že data budou synchronizována nejen mezi klíčovými systémy PREdistribuce, jednotlivými vnitřními útvary, ale v budoucnu mohou být poskytována i třetím stranám
- Minimalizovat vznik duplicit
- A v neposlední řadě najít argumenty pro nebo proti vybudování paralelního datového skladu dublujícího stávající 3 systémy

Smyslem této práce je udržet náskok PREdistribuce (potažmo PRE) v kvalitě poskytovaných služeb a efektivnosti při vynakládání budoucích investic a připravit se na výše popsané blížící se výzvy.

Obrázek 4 Postup řešení



Zdroj: Vlastní tvorba autora

3.2 Procesní analýza PREdistribuce z pohledu dat

Tabulka 1 Procesní a datové oblasti

Správa sítí	Řízení DS		Řízení aktiv	Prodej distribuce	
		Údržba sítí	Výstavba sítí	Obsluha zákazníků	Měření a odečty
Kmenová data sítě	Data reálného času	Data údržby	Investiční data	Zákaznická data	Data měření

Zdroj: vlastní tvorba autora

Na základě analýzy procesů bylo definováno 12 datových oblastí:

- Ekonomické informace
- Kmenová data sítě (statická)
- Zapojení a řízení sítě
- Data měření
- Majetkoprávní vztahy
- Provozní a údržbová data
- Rozvoj sítě
- Výstavba
- Zákaznické informace
- Legislativa a normy
- Jiné datové domény
- Správa sítí VN/NN, RS, TS a ochran

Role a kompetence v řízení dat

Je třeba vyjasnit pojetí rolí „vlastník dat“ (manager) a „správce dat“ (datový správce). V úvahu připadají dvě pojetí jejich kompetencí a vzájemné spolupráce a následně jejich

spolupráce s rolemi vlastníků procesů (resp. představitel business potřeb, liniové řízení) a správci aplikací (metodik aplikace, infromatický správce):

Varianta 1

Manažer definuje požadavky na data (kvalitu); správce provádí naplnění požadavků definováním postupů práce s daty. Toto je přístup vlastníků dat jako zákazníků, správce dat jako dodavatel.

Varianta 2

Manažer schvaluje a odpovídá za realizaci požadavků businessu (dodržení kvality), pro dílčí data deleguje pravomoci k zajištění adekvátních postupů a akcí na datového správce; zákazníkem je business, „dodavatelem“ je manager („odpovědným“) a správce („výkonným“). Ve společnosti PREDistribuce je preferováno je toto pojetí, proto bude dále rozpracováno.

- Naplnění „kritérií kvality“ dat realizuje správce dat:
- Definováním postupů prací s daty (pro vznik, editaci, čtení, poskytování a likvidaci dat)
- Kontrolou (auditem) plnění těchto postupů (pravidel)
- Požadavky na procesy (business) na realizaci postupů či organizační změny
- Požadavky na realizaci v IS, resp. projekty infromatické podpory
- Správce dat odpovídá za dokumentaci:

Definice dat Jednoznačné zachycení metadat v modelu (popis struktur)

Postup vzniku dat	kdo a jak pořizuje (obsah)
Postup editace	kdo a jak edituje (obsah)
Postup čtení	kdo a kdy čte (obsah)
Postup likvidace	kdo, jak a kdy likviduje (obsah)

Pro celkový přehled o datech se ukazuje jako vhodné ustavit roli správce datového modelu (logický přehled, nenahrazuje datového architekta) – viz dále.

Stále se jeví jako vhodné definovat pravomoci a odpovědnosti k „datům“ odděleně od kompetencí v oblasti řízení informačních systémů, přestože u „neproblémových dat“

uložených pouze v jedné aplikaci pravděpodobně bude role datového správce totožná s metodikem aplikace

3.2.1 Životní cyklus dat

Výše uvedené se krystalizovalo též v průběhu interview, kdy přišla řeč na životní cyklus dat – byly identifikovány fáze vznik, pořízení, editace, využití a likvidace a zvažováno, z jaké z identifikovaných fází je vhodné určovat správce dat a zda nadřízený „vlastník“ musí být nezbytně jeho liniově nadřízený.

V jednotlivých fázích vystupují role:

Název	Popis
původce dat	prvotní vznik dat
vstupní kontrolor dat	kontrola správnosti (kvality) dat
pořizovatel dat	pořízení do informačního systému (IS)
editátor dat – modifikátor	mění již zadaná data
editátor dat – doplňovač	přidává další atributy k již existujícím datům
uživatel dat – uživatel	čte a využívá data
poskytovatel dat	předání dat z IS externímu uživateli
likvidátor dat	odstraňuje data

3.2.2 Řízení požadavků na nová data

a) Na základě interního nebo externího (legislativního) podnětu vzniká požadavek na nová data (nový atribut dat).

b) Vlastník procesu (vedoucí zaměstnanec) předá požadavek správci datového modelu, který zjišťuje zda požadovaný atribut se již nevyskytuje v datovém modelu.

i. Atribut se v modelu nevyskytuje: správce modelu navrhuje, kdo bude vlastníkem nového atributu...

ii. Atribut se v modelu vyskytuje: správce modelu předá požadavek správci dat zasažené oblasti, který s metodikem aplikace v níž se atribut vyskytuje a s metodikem aplikace v níž je atribut požadován, analyzuje možné způsoby řešení požadavku a náklady.

c) Správce dat předá návrh řešení požadavku vlastníkovi dat ke schválení, resp. k projednání s managerem businessu..

3.2.3 Posouzení rozsahu a kvality dat

5.1 Realizovaný postup

Nejprve byla navržena možná kritéria pro posuzování kvality dat a jejich význam. Ani z pohledu teorií a zvyklostí firem totiž neexistuje zcela ustálená terminologie. Na sledované úrovni podrobnosti a bez vazby na informační systémy však nebylo možné aplikovatelnost kritérií vyzkoušet.

Na úrovni konceptuálního modelu bylo provedeno pocitové ohodnocení důležitosti a problémovosti dostupných dat a to formou přidělení omezeného počtu bodů k příslušným informačním blokům. K vybraným datům byl doplněn popis problémů.

Pro hrubý přehled ohodnocení důležitosti a problémovosti informačních bloků je možné použít pocitové ohodnocení formou přidělování bodů. K nejdůležitějším a nejproblémovějším datům je vhodné doplnit zdůvodnění, případně návrh řešení nekvality dat.

Případně je možné použít týmové hodnocení (zejména na úrovni řídicího výboru, případně vedení PREDistribuce) a společné hledání konsensu o výběru klíčových a zároveň nepřiliš kvalitních dat pro další aktivity.

Na této úrovni datového modelování, resp. členění informačních bloků nelze očekávat podklady pro posouzení efektivnosti míry detailu sledování některých objektů, resp. míry podrobnosti a rozsahu příslušných dat.

Pro objektivní ohodnocení rozsahu a zejména kvality dat by bylo nutné jít do mnohem většího detailu (tabulek a atributů) v jednotlivých informačních systémech. Z hlediska efektivnosti je reálné pouze u opravdu klíčových dat a je potřeba do projektu zapojit správce aplikací či jiné IT odborníky, případně využít externích služeb a specializovaných nástrojů pro posuzování kvality dat (a často též následnou konsolidaci obsahu).

3.3 Vytvoření zjednodušeného datového modelu PREDistribuce

Datový model

- Byly probrány různé pohledy na datový model.
- Byl vybrán pohled dle „vlastníků“ dat.

Na základě identifikovaných podkladových dat byl postupně vytvářen přehledový model se zachycením podstatných vazeb mezi jednotlivými daty, resp. jejich skupinami.

Obrázek 5 Konceptuální datový model

Zdroj: vlastní tvorba autora

Použitý postup a výsledný model je při vědomí nutné míry abstrakce a zjednodušení možné hodnotit jako vyhovující. Domapování dat pro zbývající oblasti distribuce by nemělo rozsah modelu nějak násobně zvětšit, protože na této úrovni podrobnosti jsou mnohé datové objekty společné všem napěťových hladinám a v modelu jsou již tedy zachyceny.

Vzhledem k rozsahu datového modelu na konceptuální úrovni byl pro zpracování jako dostačující nástroj Visio, nicméně pro soustavnější práci a zejména hlídání skutečných vazeb mezi objekty na detailnější úrovni by bylo třeba využít sofistikovanější prostředí:

- specializovaný modelovací nástroj s databázovou evidencí údajů (typu CASE nebo prostředí ARIS), což by ale vyžadovalo vzhledem k ceně a složitosti prakticky centralizovanou správu modelu určeným specialistou s možným vytvořením úzkého hrdla, brzděním ostatních aktivit, obtížným praktickým využitím a tím udržením aktuální údajů,
- metadata repository, resp. aplikaci pro správu dat o datech – pravděpodobně též na trhu, ale pokud by byla vytvořena na míru, mohla by mimo jiné umožnit sledování odpovědných rolí k datům (s vazbou na personální systém) a zakomponování vazeb na dosavadní podpůrné nástroje IIS

(evidence aplikací ve STATIIS a DEPO). Vhodným nastavením oprávnění by umožnila široký přístup a sdílení mezi metodiky dat, včetně adresné odpovědnosti za aktuálnost evidovaných metadat a dalších souvisejících údajů.

3.3.1 Metadata

Obecně používané pojmy (u různých teorií či dodavatelů se interpretace může lišit, resp. termíny se významem překrývají, ale jsou pojmenovány odlišně):

- přesnost
- správnost
- validita
- platnost
- syntaxe&formát
- korektnost
- úplnost (kompletnost)
- konzistence
- integrita
- aktuálnost
- jedinečnost
- včasnost
- dostupnost
- bezpečnost

V angličtině např. Data Quality Metrics (podle SLA for Data by Sid Frank, Published: June 6, 2006):

- Accuracy and Precision
- Completeness
- Reliability
- Availability
- Timeliness/Freshness

- Consistency and Uniqueness
- Formats and Delivery Packaging

Tabulka 2 Kritéria datové kvality (potenciální nové atributy dat PREdistribuce)

Požadavky na data		Kvalitu určuje	
Název	Popis	Náplň (obsah dat)	Systém poskytující data (organizační, IS)
identifikovatelnost	jednoznačnost objektu (primární klíč)	X	(X)
vazby na ostatní data	cizí klíče	X	X
bezpečnost	přístup k datům – oprávnění		X
dostupnost	přístup k datům v čase (+ případné náhradní postupy)		X
umístění	Přístup k datům v místě (v aplikaci) (kde to chce mít uživatel k dispozici)		X
naplněnost	je to vůbec pořízeno do IS?	X	
přesnost	hloubka, detailizace	X	
aktuálnost	odpovídající data skutečnosti po dobu požadované platnosti	X	
formát	struktura dat, S/N	X	X
náklady - pořízení	interní zdroje		X
náklady - údržba	interní zdroje		X
náklady - likvidace	interní zdroje		X
historické údaje	hloubka zpětné sledovatelnosti údajů	X	X
auditovatelnost změn	sledovatelnost pořizovačů a editátorů	X	X
(legislativní potřeba)	sledují se data z externích důvodů		
doba archivace	nezbytná doba archivace		X
potřebnost	hodnota dat (souhrnné ohodnocení přínosu dat)	Ohodnocuje „zákazník dat“ (business manager)	
náklady	souhrnné ohodnocení, kolik nás stojí „údržba“ takto kvalitních dat (náklady na IS, ale též na procesy pořizování, kontroly atd.)	Ohodnocuje „vlastník (metodik) dat“	

Zdroj: Vlastní šetření autora.

3.4 Návrh pravidel Master Data Managementu do procesů PREdistribuce

Návrh systému řízení dat, zejména působnosti rolí, včetně obsazení

V rámci přípravy byly posouzeny různé přístupy k řešení správy dat, pojetí rolí a možného rozdělení jejich pravomocí a odpovědností. Zvažována byla nejvhodnější varianta zejména kvůli vazbě hierarchie řízení dat na liniovou organizační strukturu a řízení rozvoje informačních systémů.

V rámci identifikace a klasifikace dat byly navrženy odpovědné útvary za zastřešení správy datových oblastí, případně některých dílčích dat. Protože však nejsou řešeny reálné kompetence, bude muset být přiřazování rolí konkrétním datům řešeno po přípravě metodiky.

Na úrovni řídicího výboru a v rámci případné realizace posoudit navrženou metodiku a její životaschopnost. Zvážit formu jejího publikování a míru závaznosti a vymahatelnosti uvedených ustanovení. S tím souvisí i výběr a způsob určení/jmenování pracovníků do „datových“ rolí.

Stanovené povinnosti pravděpodobně nebude (stejně) efektivně realizovat u všech dat, resp. předpokládá se postupné zavádění správy dat do praxe dle prioritních potřeb. Pro ověření schválené metodiky jmenovat role ve vztahu k několika klíčovým datům a odzkoušet aplikovatelnost v praxi, včetně zastřešení řídicím výborem.

Mezi další metodiky, které byly analyzovány, patří:

- Master Data Management (MDM) – stručný popis viz Příloha č. 4
- Customer Data Integration (CDI) – přístup k integraci zákaznických dat
- Data Governance (DG) – stručný popis viz Příloha č. 5
- Data Stewardship
- Enterprise Knowledge Management, vč. Data Ownership

Dalšími možnými zdroji se mohou stát:

- Total Information Quality Management (TIQM) – koncept komplexního řízení informační kvality, který se skládá z šesti principů vedoucích k řízení

informační kvality a vybudování firemní kultury potřebné k trvalému udržení informační kvality:

1. Proces hodnocení architektury a struktury dat.
 2. Proces měření kvality a čistoty dat.
 3. Proces měření finančních ztrát z nekvalitních dat.
 4. Proces čištění a transformace dat.
 5. Proces zvýšení informační kvality napříč organizací.
 6. Proces stabilizace kvality informačního prostředí.
- Data Management Body od Knowledge (DMBOK) – komplexní přehled celkem 9 funkčních oblastí řízení dat zpracovaný DAMA International Foundation
 - Information Lifecycle Management (ILM)

3.5 Samotná metodika

3.5.1 Výklad pojmů a zkratk

Pojem Význam

Data, (informace, znalosti) Pro potřeby této metodiky pojem data zastupuje různé formy zachycených dat a informací (případně znalostí), strukturované i nestrukturované povahy. Přestože většina důležitých dat zpravidla bývá/měla by být uložena v informačních systémech a metodika je takto primárně směřována, resp. formulována, je nutné zdůraznit, že správa dat je obecně uplatnitelná i na data zachycená jiným způsobem s tím, že některá ustanovení vztažená k IS/ICT (databázím, aplikačnímu SW apod.) se potom pro ně vykládají přiměřeně nebo se nepoužijí.

Datová oblast, datová doména, informační blok, ...

datová entita, atribut Různé termíny usnadňující pojmenování různého seskupování dat či zdůraznění určitého stupně abstrakce/detailu sledovaných dat.

Je-li to účelné a efektivní, je možné role správy dat přiřadit ke kterékoliv takovéto množině či elementu dat (při obsazování role samozřejmě s jednoznačným výkladem rozsahu dat svěřených do správy)..

Datový model Prostředek pro zachycení datových oblastí, domén, entit či atributů a jejich vzájemných vazeb. Může být realizován grafickou a/nebo textovou formou ve specializovaném modelovacím nástroji nebo jiném vhodném prostředí (např. metadata repository).

Konceptuální / logický datový model Model dat, který popisuje data pohledem businessu a mj. sjednocuje různé pojmy pro stejné datové objekty a zajišťuje jednotný výklad významu dat.

Fyzický datový model Popisuje datové struktury v informačních systémech.

Zdrojový systém Informační systém, ve kterém primárně vznikají určitá data (určená na úrovni souboru dat, datového záznamu nebo atributu).

Cílový systém Informační systém zpracovávající nebo zpřístupňující data, která vznikla, byla doplněna, opravena či jinak editována v jiném systému.

(Charakter systému je úzce svázán s danými daty, pro jedny může být zdrojovým systémem, pro jiné cílovým.)

Program DMD Dlouhodobý program (soubor projektů a dalších aktivit) podporovaný vedením PREdistribuce s vizí „Zajistit a udržet kvalitu informací používaných v procesech distribuce“.

Zkratka Význam

DMD Datový model distribuce, Data Management distribuce

IS/ICT Informační systém / informační a komunikační technologie

3.5.2 Úvod

(1) Data a informace jsou cenným bohatstvím (aktivem) každé společnosti, o které je nezbytné se pečlivě starat stejně jako o jiný podnikový majetek. Řízení a správa dat je v informačním věku důležitou podnikovou funkcí.

(2) Kvalita dat je nezbytnou podmínkou pro získání kvalifikovaných informací. Kvalitní data tak představují jeden z kritických faktorů ovlivňujících správné rozhodování, efektivní fungování podnikových procesů a dlouhodobé udržení podnikatelské úspěšnosti.

a) Kvalitní data přináší společnosti snižování nákladů na špatně vyhodnocené informace a následná chybná každodenní i zásadní strategická rozhodnutí.

b) Používání správných a validních dat ve styku se zákazníky a externími partnery přispívá také k dobrému jménu společnosti.

c) V neposlední řadě mají spolehlivá a auditovatelná data pro interní statistiky či externí výkaznictví vliv na zajištění souladu s legislativními a regulatorními požadavky. Přitom neplnění či podcenění některých požadavků může mít zásadní ekonomické dopady spojené s uplatněním sankcí a v řadě oblastí se jedná až o trestně právní zodpovědnost.

d) Konsolidace a hlavně zvyšování datové kvality vede také k významnému posunu ve využití stávajících informačních systémů.

(3) Cílem metodiky správy dat distribuce je ustanovit zejména procesně-organizační systém a pomocí jeho naplňování dospět minimálně u klíčových dat ke stavu, kdy:

a) data jsou sjednocena, a to nejen jejich obsah, ale také jejich chápání a interpretace,

b) data vznikají definovanými postupy ve spolehlivých systémech a tím jsou kvalitní a důvěryhodná,

c) data jsou dostupná vždy, když jsou potřebná, a všem, kteří je potřebují a mají právo s nimi pracovat.

3.5.3 Role v oblasti správy dat (Data Governance/ Data Management)

Dále uvedené působnosti a pravidla představují obecné zásady pro výkon vyjmenovaných rolí. Při obsazování dané role ve vztahu ke určeným datům (datové oblasti, doméně, entitě a atributu či jinak určené skupině) je možné práva a povinnosti konkretizovat či doplnit.

3.5.3.1 Sponzor správy dat

- (1) Sponzor správy dat určuje celkový směr řízení kvality a správy dat a poskytuje na ně zdroje.
- (2) Sponzor správy dat distribuce prosazuje Program DMD a podporuje řízenou správu dat distribuce.

3.5.3.2 Datový řídicí výbor (Data Governance Committee)

- (1) Datový ŘV je složen z vybraných zástupců businessu a vedoucího IIS. Zástupci businessu (distribuce) jsou určeni ředitelem (PREdistribuce) a obvykle se jedná zároveň o vlastníky klíčových dat.
- (2) Datový ŘV schvaluje obecné procesy, pravidla a politiky pro správu dat.
- (3) Datový ŘV rozhoduje o prioritách řešení datových problémů a rozhoduje o rozdělení rozpočtu, je-li na problematiku správy dat vyčleněn.
- (4) Datový ŘV schvaluje projekty z oblasti zlepšování kvality dat, úprav souvisejících procesů a změn či vývoje IS/ICT. I v rámci rozvoje prosazuje řešení aspektů datové kvality a standardizace dat.
- (5) Datový ŘV určuje obsazení rolí manažera dat, vlastníků dat a potvrzuje metodiky dat.

3.5.3.3 Manažer (kvality) dat (Data Quality Manager)

- (1) Manažer dat (distribuce) je nositelem know-how v oblasti správy dat a vlastníkem procesů řízení správy a kvality dat (distribuce).

(2) Manažer dat připravuje společné metodiky a další průřezové dokumenty, vykládá pravidla ostatním zaměstnancům a podporuje jejich dodržování.

(3) Manažer dat metodicky řídí a zastřešuje tvorbu (konceptuálního a logického) datového modelu, kontroluje existenci a aktuálnost popisů dat z pohledu struktur (metadat) i pravidel pro práci s daty. Rozhoduje o způsobu zařazení nových dat do datového modelu a přiřazení za ně odpovědných osob.

(4) Manažer dat připravuje souhrnné zprávy pro Datový ŘV o fungování systému řízení dat, aktuálním stavu a trendech datové kvality, probíhajících „datových“ aktivitách a navrhuje opatření a projekty pro zlepšování dat. Manažer přitom vychází z hlášení metodiků a vlastníků dat, vlastního zkoumání, případně zjištění nezávislých datových auditorů.

3.5.3.4 Datový architekt (Data Architect)

(1) Datový architekt průřezově zastřešuje technologické aspekty uložení dat v informačních systémech.

(2) Datový architekt v úzké spolupráci se správcem dat, resp. správcem aplikací řeší zejména požadavky na datovou integraci, technologickou konsolidaci a sdílení klíčových dat mezi více systémy.

(3) Není-li vedoucím IIS určeno jinak, roli datového architekta vykonává IT architekt (zaměstnanec G 34 402).

3.5.3.5 Vlastník dat (Data Owner)

jako představitel majitele dat – „morální vlastnění“

(1) Vlastník dat je určen jako zástupce konkrétní společnosti, která je z pohledu legislativy majitelem „obsahu dat“ (přestože data může zpracovávat jiná společnost).

(2) Roli vlastníka dat určuje na základě návrhu manažera dat Datový ŘV. Zpravidla je to vedoucí sekce nebo oddělení. Pokud je to možné, doporučuje se určit vlastníka dat z nadřazených zaměstnanců pořizovačů a editátorů dat.

(3) Vlastník dat odpovídá za způsob nakládání s daty vzhledem k požadavkům společnosti a nárokům vyplývajícím z příslušné legislativy.

(4) Vlastník dat určuje pravidla pro přístup k datům a minimálně jednou ročně schvaluje inventuru přidělených oprávnění do aplikací pracujících se svěřenými daty.

a) Pravomoc schválit inventuru může písemně delegovat na vlastníka aplikace, resp. vlastníky aplikací. Takovéto přenesení odpovědnosti se však nedoporučuje, pokud vlastník aplikace je z jiné společnosti než vlastník dat a zejména pokud se jedná o data citlivá z pohledu unbundlingu.

b) Pravomoc schvalovat během roku požadavky na konkrétní přístupy a přidělení oprávnění v příslušných aplikacích pro přístup k datům obvykle vykonává vlastník aplikace nebo metodik aplikace, pokud vlastník dat nestanoví, že je chce schvalovat osobně nebo jinou osobou.

jako garant dat – „faktická vykonatelnost“

(1) Pro reálnou vykonatelnost garance kvality dat se může ukázat praktické, přenést roli vlastníka dat (pro výkon v této kapitole dále uvedených působností) na osobu, která v souvislosti se společnými procesy (distribuce a obchodu), resp. zajišťováním služeb (na základě Smlouvy o poskytování služeb mezi společnostmi v rámci Skupiny PRE), má přenesenou odpovědnost i za používané informační systémy a zpracování dat.

(2) Není-li vlastník dat určen, potom jeho pravomoci a odpovědnosti vykonává vlastník aplikace, resp. společně vlastníci všech aplikací, ve kterých jsou data uložena. Vedoucí úlohu by přitom měl hrát vlastník aplikace, která slouží pro daná data jako zdrojový systém. (Ten by měl být také prvním uvažovaným kandidátem při hledání vhodného vlastníka dat.)

(3) Aby se zajistilo vnímání kvality dat jako primárně business a nikoliv informatické záležitosti, zaměstnanec sekce IIS by měl být vlastníkem/garantem dat pouze ve výjimečných a zvláště odůvodněných případech.

(4) Vlastník dat je zodpovědný za dodržování stanovené kvality dat, případně určené zlepšování, zejména ve vztahu k obsahu dat.

a) Pro praktické zajištění plnění kritérií kvality konkrétních (dílčích) dat ze svěřené datové oblasti nebo jinak specifikované skupiny dat určuje obsazení role metodika, resp. metodiků dat.

b) Pokud určí více metodiků úzce souvisejících nebo dokonce překrývajících se dat, zodpovídá za jejich koordinaci.

(5) Vlastník dat schvaluje (nebo doporučuje Datovému ŘV ke schválení) investice do řešení datové kvality svěřených dat a to jak z procesně-organizační, tak obsahové stránky (např. čištění dat).

(6) Vlastník dat se z pohledu datové kvality a konsolidace dat vyjadřuje k návrhu investic do příslušných aplikací (např. změny struktur klíčových dat v IS, tvorba rozhraní pro přenos či sdílení dat).

3.5.3.6 Metodik dat (Business Data Steward)

(1) Metodik dat je výkonnou složkou zajišťující zodpovědnost za kvalitu dat pro vlastníka dat.

(2) Není-li metodik dat určen, potom jeho pravomoci a odpovědnosti vykonává metodik aplikace, resp. společně metodici všech aplikací, ve kterých jsou data uložena. Vedoucí úlohu by přitom měl hrát metodik aplikace, která slouží pro daná data jako zdrojový systém. (Ten by měl být také prvním uvažovaným kandidátem při hledání vhodného metodika dat.)

(3) Metodik dat pomocí dále uvedených činností a konkretizováním obecných pravidel pro správu svěřených dat zajišťuje nastavení adekvátních postupů a srozumitelných pravidel k dosažení stanovené úrovně kvality dat, zejména z pohledu datového obsahu.

(4) Metodik dat odpovídá za definici svěřených dat – popis terminologie, významu dat i logických struktur a vazeb z pohledu businessu – a její dokumentaci, resp. zachycení v datovém modelu.

(5) Metodik dat podporuje sladění business požadavků na data. Vytváří business pravidla pro datovou standardizaci, která reflektuje potřeby uživatelů dat.

(6) Metodik dat v rámci projektů změn či vývoje IS/ICT sleduje příležitosti pro opětovné použití či zlepšení existujících dat a rozhoduje o potřebě doplnění či vytvoření nového datového elementu.

(7) Metodik dat odpovídá za stanovení postupů pro pořizování a editaci dat (včetně výmazu) tak, aby bylo zajištěno dosahování stanovených kritérií kvality dat. V rámci těchto pravidel může být účelné určit např. kontrolní body pro podklady, ze kterých se data pořizují (např. sběrné formuláře nebo externí elektronická data), a pro pořizování dat do informačního systému navrhnout a implementovat dostatečné validační kontroly (organizační i programové).

(8) Pokud zásady pro vznik a modifikaci dat nejsou stanoveny a prakticky realizovány, nedoporučuje se investovat do dávkového zkvalitňování (opravy) dat (ať automatizovaně, tak ručně), protože nelze spolehlivě garantovat, že kvalitu vyčištěných dat lze udržet.

(9) Při pouhém čtení dat sice nemůže dojít k poškození obsahu dat, přesto je vhodné, aby metodik dat stanovil, resp. měl zmapované i způsoby takového využívání dat, aby poznal potřeby všech uživatelů dat a také zamezil například chybné interpretaci či špatnému následnému zpracování dat v cílových systémech.

(10) Metodik dat rozpracovává pravidla pro přístup k datům, resp. pro přidělování přístupů a určení oprávnění, případně stanovuje další opatření či požadavky ve vztahu k bezpečnosti dat (dostupnost, zálohování, archivace atp.).

(11) Zvláště u dat citlivých (chráněných, osobních, utajovaných atp.) metodik dat dále stanoví pravidla a způsoby jejich likvidace a to nejen v informačních systémech.

(12) Metodik dat definuje způsob a četnost pravidelných kontrol nejen aktuálního stavu datové kvality, ale především dodržování stanovených postupů a plnění pravidel. Kontroly provádí zpravidla sám nebo po dohodě pověří jinou vhodnou osobu (např. metodika aplikace, správce dat nebo nadřízeného zaměstnance pořizovačů). V odůvodněných případech může požádat o provedení mimořádného nezávislého auditu či jiné komplexnější prozkoumání svěřených dat.

(13) V případě zjištění neshod metodik dat navrhuje nejen způsob nápravy obsahu dat, ale i další vhodná opatření ke zlepšení stavu a to:

- a) návrhem doškolení či sankcí pro zaměstnance, kteří pravidla porušují, nebo
- b) při vyhodnocení obtížné aplikovatelnosti stanovených pravidel úpravou postupů a zásad,
- c) případně požadavkem na zásadnější redesign procesů či organizační změny.

(14) Pokud očekávané kvality dat nelze dosáhnout pouze procesně-organizačními opatřeními na straně uživatelů dat, navrhuje metodik dat požadavky na realizaci úprav příslušných aplikací, resp. jejich datových struktur, vazeb, automatizovaných kontrol, čištění, přenosu a sdílení dat.

3.5.3.7 (Informatický/technologický) správce dat (System Data Steward)

- (1) Správce dat je informatickým partnerem metodika dat.
- (2) Správce dat posuzuje realizovatelnost potřeb a požadavků metodika dat a zajišťuje systémové a technologické řešení správy svěřených dat.
- (3) Správce dat má přehled o způsobu uložení dat v informačních systémech a odpovídá za popis dat na úrovni fyzického datového modelu.
- (4) Správce dat spolupracuje s metodikem dat při identifikaci datových chyb a návrzích metod pro jejich opravy.
- (5) Zásadní a průřezové otázky posuzují správci dat společně s datových architektem.
- (6) Není-li vedoucím sekce IIS určeno jinak, roli správce dat vykonává (informatický/ technologický) správce aplikace, resp. společně správci aplikací, které s příslušnými daty pracují.

3.5.3.8 Producent dat (Data Provider)

- (1) Producenti dat (pořizovači a editátoři) musí cítit odpovědnost za kvalitu svých „produktů“ (dat/informací) vůči ostatním uživatelům dat.

(2) Producent dat je odpovědný za přesnost a správnost konkrétních dat v rámci pravidel a zásad stanovených metodikem dat. Pořizovač dat obvykle odpovídá za kvalitu (shoda s pravidly i zachycovanou realitou) jím pořizovaného záznamu, případně za to, že nezadal opakovaně stejný záznam. Editátor odpovídá za správnost změněného atributu (pokud systém neumožňuje zjistit, jaký údaj modifikoval, může být odpovědný za kvalitu celého záznamu).

(3) Obsazení role producenta (pořizovače, editátora) dat se zpravidla stanoví hromadně v rámci pravidel pro přístup k datům a přiděluje se při schvalování a nastavení oprávnění do zdrojového systému.

(4) Pokud je producentem dat externí subjekt, je vhodné stanovit a aplikovat kontrolní kritéria pro převzetí dat a požadavky na datovou kvalitu zakotvit do smlouvy. Vůči systému správy dat nese odpovědnost (za kvalitu dat) osoba, která provádí či zajišťuje pořízení dat do interního systému.

3.5.3.9 Uživatel dat (Data Beneficient)

(1) Pro adekvátní správu dat je třeba důkladně zmapovat uživatele dat, resp. jejich skupiny a porozumět jejich požadavkům na kvalitu dat. Uživatelé dat (prostřednictvím svých zástupců) případně mohou přímo přispívat do rozpočtu správy dat a to podle toho, jakou hodnotu jim poskytovaná data (v dané kvalitě) přinášejí.

(2) Obsazení role uživatele dat se zpravidla stanoví hromadně v rámci pravidel pro přístup k datům a přiděluje se při schvalování a nastavení oprávnění příslušné aplikace. (Pochopitelně role producenta a uživatele dat se často bude překrývat.)

(3) Datům a informačním systémům s přímými externími uživateli je třeba z pohledu kvality dat a zabezpečení přístupových oprávnění, resp. zajištění zpřístupnění či poskytnutí pouze adekvátních záznamů věnovat zvláštní pozornost.

(4) V roli uživatele dat kromě přímých uživatelů aplikace mohou být též zprostředkovaně např. příjemci papírových sestav či adresáti elektronických zpráv a to i v podobě externího subjektu či informačního systému. Požadavky na datovou kvalitu pak mohou vyplývat i z legislativy či smluvních podmínek.

3.5.4 Datová architektura

Problém konsolidace a kvality dat v informačních systémech se řeší v zásadě od okamžiku, kdy začaly být nasazovány specializované systémy na jednotlivé agendy. Každý z těchto systémů má obvykle svoji databázi, svoje datové struktury, svůj datový model, svoji logiku – vzniká tedy více reprezentací jednoho reálného subjektu. V okamžiku, kdy je třeba získat o dané entitě všechny informace, je třeba získat data z více systémů, pokud ovšem přesně víme, které záznamy se daného subjektu (např. zákazníka) týkají.

- (1) Každá data by tedy měla ideálně vznikat a být editována pouze v jednom zdrojovém systému.
- (2) Existuje-li pro určitá data více zdrojových systémů:
 - a) musí být stanovena „silná“ procesně-organizační pravidla pro jejich vzájemnou konsolidaci. Dodržování těchto pravidel musí být pravidelně kontrolováno a efektivnost postupů revidována, aby se zajistilo kvalitní a úplné sledování všech relevantních dat, případně jejich reduplikace;
 - b) je vhodné posoudit důvody a případně navrhnout technologické změny směřující ke zjednodušení pořizování a editace dat, případně vzájemného sdílení konsolidovaných datových záznamů či doplňování atributů.
- (3) V cílových systémech je využíváno:
 - a) zobrazení data přímo z databází zdrojových systémů prostřednictvím on-line rozhraní. (V cílovém systému tedy nejsou vůbec uložena nebo jen v technicky nezbytné míře, např. identifikátory a cizí klíče.);
 - b) v odůvodněných případech replikace (kopírování) dat dle určených pravidel. Je nutné si uvědomit, že toto vícenásobné uložení totožných dat zvyšuje nároky na diskovou kapacitu datových úložišť.

(4) Pokud nastává datová změna ve zdrojovém informačním systému, musí o ní být v dostatečném předstihu informováni správci všech navazujících systémů, aby ji adekvátním způsobem zohlednili.

(5) Nedílnou součástí péče o data v informačních systémech musí být řešení problematiky bezpečnosti dat. Dle charakteru dat z pohledu obsahu i nároků na způsob jejich využívání dle požadavků metodika a vlastníka dat je nutné věnovat pozornost zejména:

- a) nastavení přístupových oprávnění a přidělování práv jednotlivým uživatelům,
- b) fyzickému omezení přístupu k systémům, diskům a datovým mediím, případně šifrování dat na jejich úložištích,
- c) zajištění dostupnosti dat, resp. garantované doby provozu informačního systému,
- d) řešení zálohování a obnovy.

3.6 Realizace opatření

Návrh návazných projektů (realizačních)

V rámci interview s vlastníky dat zaznělo několik dílčích problémů s dostupností konkrétních dat a dokumentů, ale vesměs se jednalo o problém procesně-organizačního charakteru, než primárně o datovou nekvalitu, přitom se problém nezdál natolik zásadní, aby byl iniciován samostatný nápravný projekt. Pár záležitostí souvisejících s realizací v informačních systémech je již obecně známo a jsou předmětem řešení v rámci probíhajících či naplánovaných IT projektů.

Vzhledem k malé míře detailu identifikovaných dat (časový rámeček projektu) a absenci inženýrského pohledu (složení projektové týmu) nebylo možné provést detailní posouzení dat v systémech, jak z hlediska fyzického modelu, datové architektury a rozhraní, tak z hlediska kvality obsahu dat. Proto nebylo možné stanovit ani konkrétní návazné projekty, potažmo odhadnout jejich cenu.

Doporučení = Lépe dokumentovat v rámci projektových schůzek identifikované problémy a zabezpečit, že budou posouzeny, resp. využity jako argumenty pro výběr

prioritních projektů a dalších aktivit. Pro lepší motivaci projektového týmu i propagaci projektu realizovat „quick wins“.

Pro zásadní rozhodnutí o náročnějších a nákladnějších realizačních projektech je však nutné provést mnohem podrobnější zmapování (alespoň u klíčových) dat, a mít k dispozici detailnější analýzu problémů, pokud možno objektivní důkazy a metriky o reálných datech. To je však možné pouze při zapojení IT specialistů spravujících příslušné informační systémy, případně za přispění sofistikovaných nástrojů pro kontrolu kvality dat. Teprve poté je možné odhadnout očekávané přínosy, porovnat je s předpokládanými náklady a navrhovat konkrétní návazné realizační projekty pro opravu a zlepšování dat.

3.7 Vytvoření vstupů pro implementaci BusinessObjects

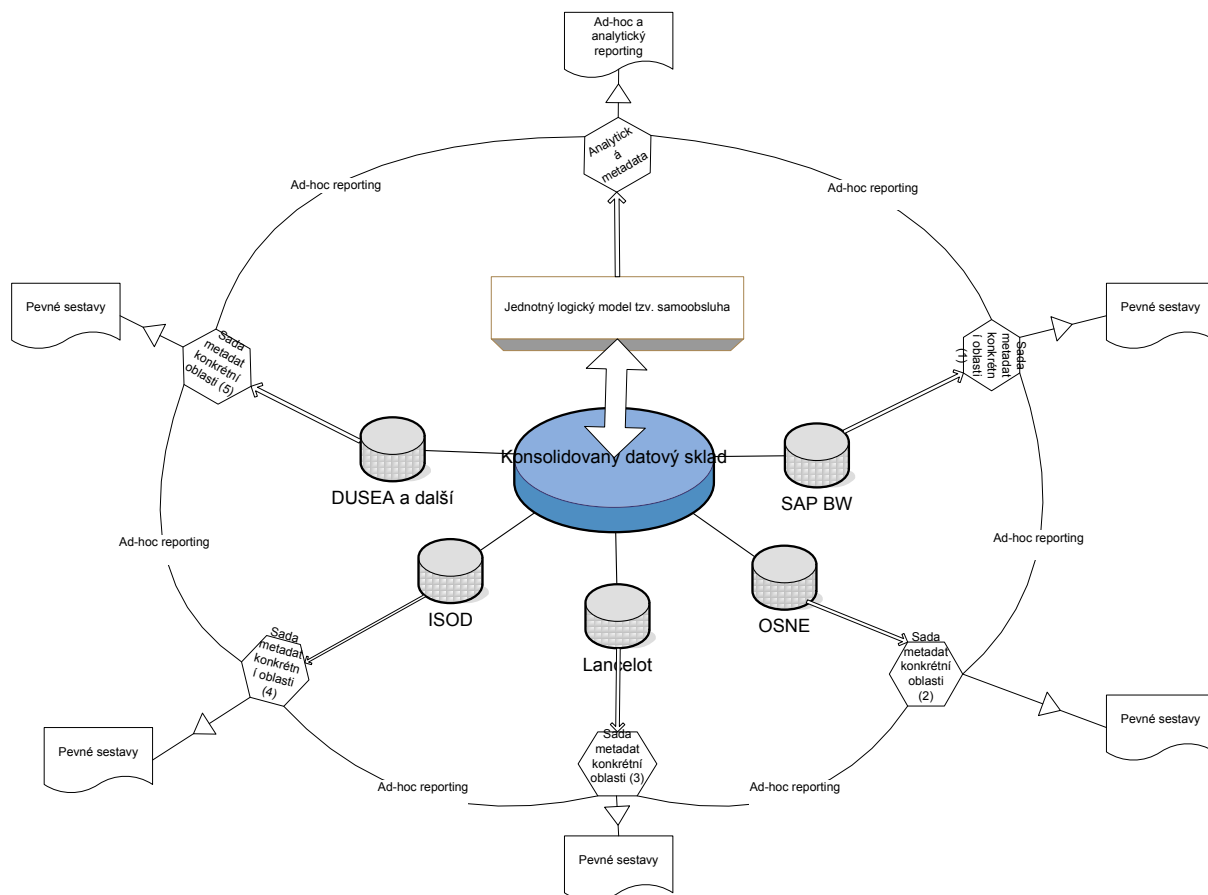
3.7.1 BusinessObjects

Balíček SAP BusinessObjects (SAP BusinessObjects Package Business Intelligence Package (BI))

- Crystal Reports 2008
- SAP BusinessObjects Enterprise Premium Edition (includes Live Office)
- SAP BusinessObjects Mobile
- SAP BusinessObjects Explorer
- SAP BusinessObjects Advanced Analysis (MS Office edition)
- SAP BusinessObjects Web Intelligence
- SAP BusinessObjects Xcelsius Enterprise
- SAP Xcelsius Enterprise Interactive Viewing
- SAP BusinessObjects Integration for Oracle E-Business Suite
- SAP BusinessObjects Integration for PeopleSoft Enterprise
- SAP BusinessObjects Integration for JD Edwards EnterpriseOne
- SAP BusinessObjects Integration for Siebel
- SAP Business Objects Data Federator
- SAP Business Objects Metadata Management

Model 1 – Konsolidovaný datový sklad navržený na míru

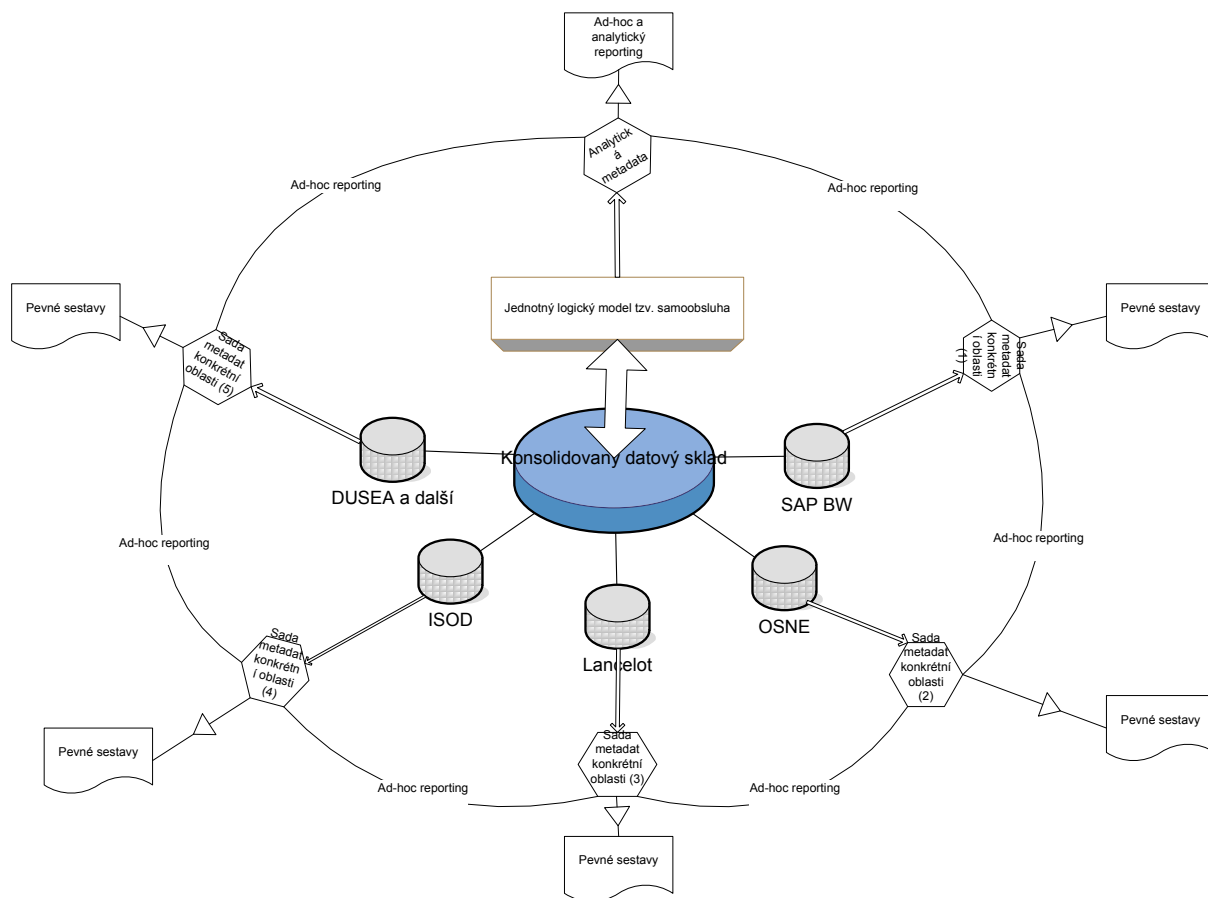
- Využití relačního datového skladu navrženého na míru
- Využití datové pumpy SAP Business objects Data Integrator



Model 2 – Konsolidovaný datový sklad Business Warehouse

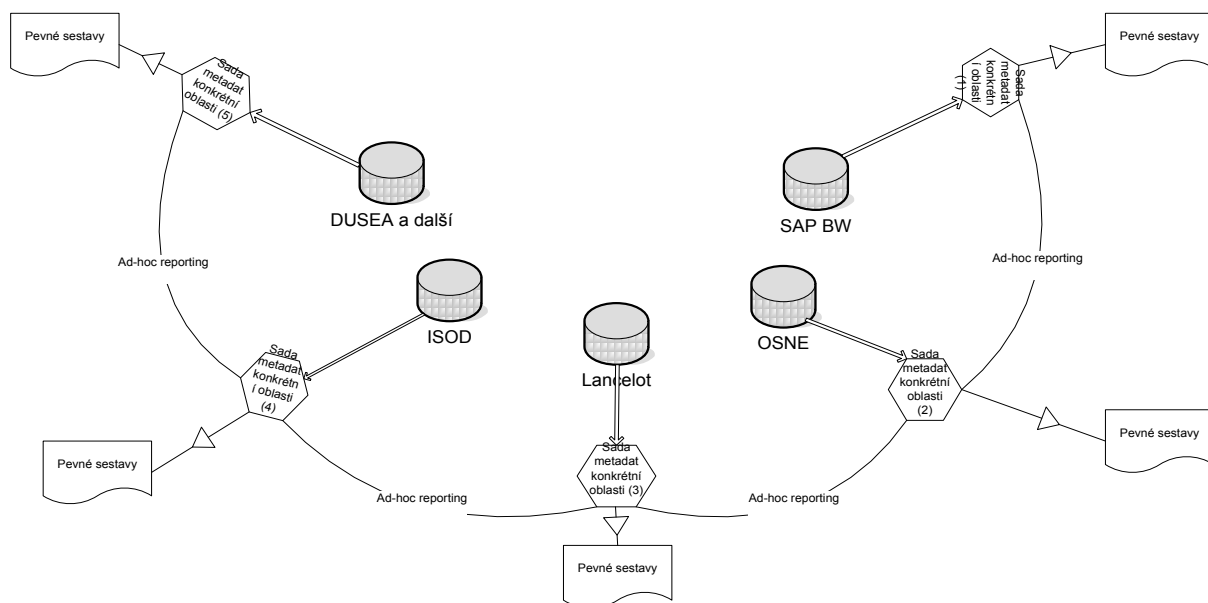
- Využití technologie BW. Architektura bude koncipována na bázi DSO objektů v nulté vrstvě data store
- Konsolidované BW může být dedikovaný stroj anebo další vrstva v již stávajícím BW

- Integrace pro vrstvu reporting bude použita komponenta Business Objects Data Federator



Model 3 – Bez konsolidace

- Využití pouze stávající infrastruktury
- Integrovaná vrstva pouze reportingová platforma Business Objects
- Více logických modelů s propojením až na úrovni jednotlivých reportů



Tabulka kladů a záporů z hlediska koncové práce s daty a implementace

	Model 1	Model2	Model 3
Sjednocená data	+	+	-
Integrace napříč SAP	+	-	+
Analytický ad hoc reporting	+	+	-
Konsolidovaný datový model	+	+	-
Cena	+	-	-
Limity další manipulace s daty	-	+	-
Složitost samotného reportingu	+	+	-
Časová náročnost samotného reportingu	+	+	-
Nutnost asistence konzultační firmy pro reporting	+	+	-
Udržitelnost metadatového modelu v čase	+	+	-
Vazby mezi logickými modely se urdžují na reportu	+	+	-
Lze nasazovat ihned	-	-	+

Identifikované problémy:

Data Federator neumožní přístup na aplikační vrstvu R/3. Produkt komunikuje pouze s BW. Nemá se tak stát ani v následující verzi.

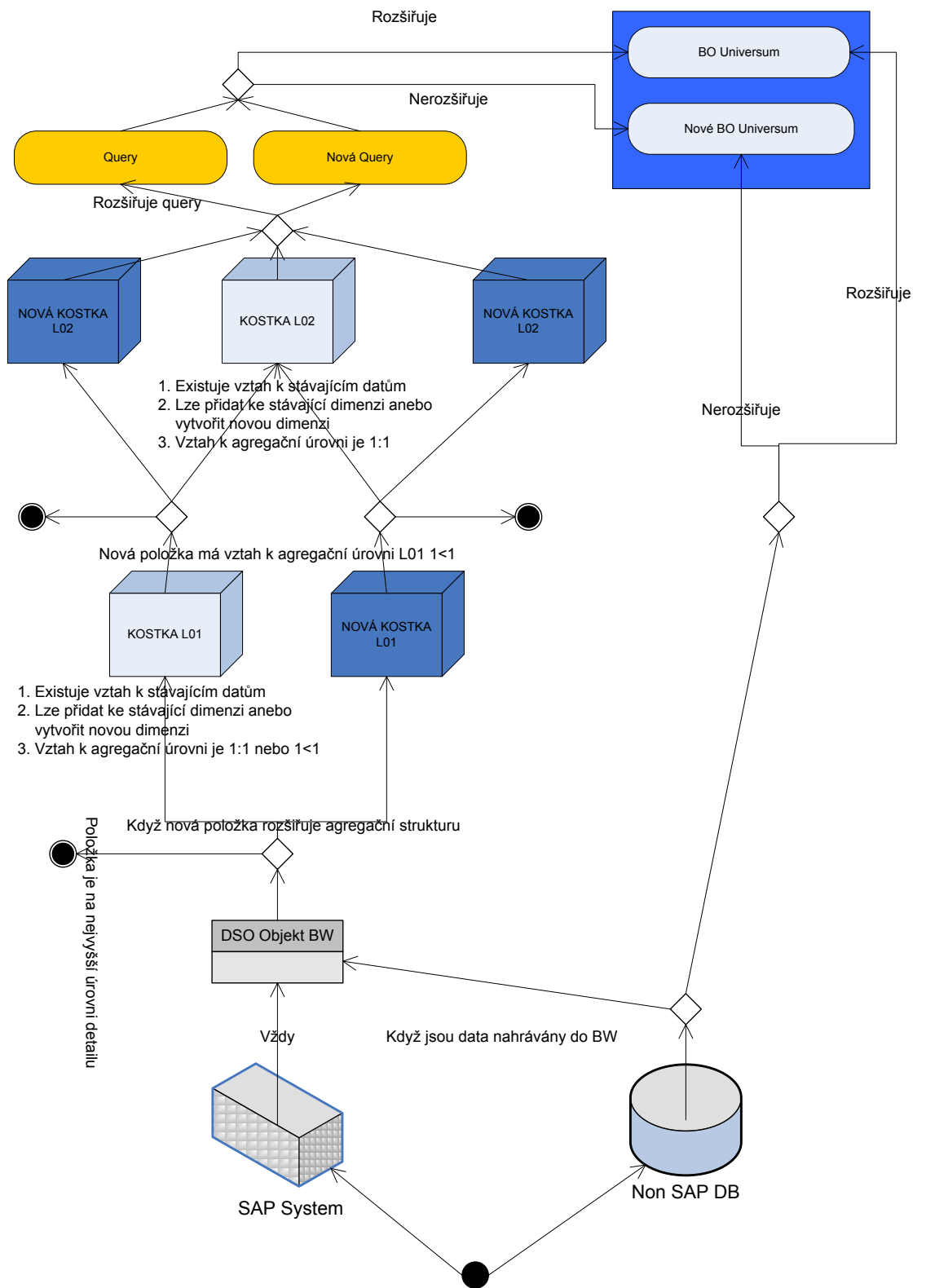
Je nutné ověřit, zda je reálné přenést repliku dat z R/3 do BW. Bude nutné vytipovat neobjemnější tabulky. Nicméně z principu vytváření nulté vrstvy vyplývá, že existují standardní extraktory, které přenáší data v nejmenším detailu, ale již z jen nutných tabulek a prováděné pomocí OUTER JOIN, což sníží celkové objemy přednášených dat.

V první polovině roku 2011 byla uvolněna nová verze BO (verze 4), která obsahuje množství změn. Mj. jsou do tvůrce metadatového modelu Designera implementovány některé funkce Data Federatoru – tj. bude umožněno mít multi data sourceové universum resp. metadatový model. Tím se zjednoduší práce koncového uživatele se reportingem nad více datovými zdroji a tím pádem v současnosti nad více metadatovými modely. Po novu by již pracoval s jedním metadatovým modelem. To bude významné plus pro všechny tři varianty, ale zejména pro variantu 3.

Revize datových zdrojů BO-BI

Byly zmapovány základní datové entity PRE a vazby mezi nimi. High-level logický datový model byl vytvořen v SW Visio.

Obr. Rozhodovací strom při začlenění nového objektu do metadatové struktury pro koncový reporting



4 Závěr

Literatura

Primární zdroje

...

Monografie

TVRDÍKOVÁ Milena. Aplikace moderních informačních technologií v řízení firmy: nástroje ke zvyšování kvality informačních systémů, 1. vyd. Praha: Grada Publishing, 2008. 173 s. ISBN 978-80-247-2728-8.

POUR Jan a kol. Podniková informatika: 2., přepracované a aktualizované vydání, 2. vyd. Praha: Grada Publishing, 2009, 496 s. ISBN 978-80-247-2615-1.

SHMUELI Galit. a kol. Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner, 2. vyd. New Jersey: John Wiley and Sons, 2010, 428 s. ISBN 978-0-470-52682-8.

CONRAD Swen a kol. IT Business Management Solutions from SAP: A Pocket Guide, 1. vyd. Zalrmobbel: Van Haren Publishing, 2010, 126 s. ISBN 978-90-8753-620-6.

Odborné knihy a časopisy

Internetové zdroje

SAP: BusinessObjects BI Solutions [online] SAP, 2011 [cit. 2011-02-10]. Dostupné z WWW: <<http://www.sap.com/solutions/sapbusinessobjects/large/business-intelligence/index.epx>>.

Přílohy

Příloha 1 Zápisy z interview s vlastníky dat