

Téma 2

Kvalita dat, projekty Data governance

1.	Kvalita dat.....	2
1.1	Hledisko mikroúrovně.....	2
1.2	Hledisko makroúrovně	2
1.3	Hledisko uživatelské úrovně	2
2.	Řízení kvality dat (procesy QA/QC)	3
3.	Funkce kontrolující kvalitu dat	5
3.1	Kontroly prostorových i popisných dat	5
3.2	Příklad kontrol popisných dat	5
3.3	Příklady funkcí kontrolujících kvalitu prostorových dat	5
4.	Příčiny nekvalitních dat a možná řešení	12
5.	Principy Data Governance	13
5.1	Co je Data Governance?.....	13
5.2	Data Governance a IT Governance	13
5.3	Co lze pomocí Data Governance dosáhnout?	14
5.4	Typické oblasti pro Data Governance.....	14
5.4.1	Data Governance se zaměřením na politiku, standardy a strategii	14
5.4.2	Data Governance se zaměřením na kvalitu dat	14
5.4.3	Data Governance se zaměřením na Privacy/Compliance/Security.....	15
5.4.4	Data Governance se zaměřením na Architekturu/Integraci	15
5.4.5	Data Governance se zaměřením datové sklady a business intelligence	15
5.4.6	Data Governance se zaměřením podporu managementu (management alignment)	16
5.4.7	Přehled aktivit v jednotlivých oblastech	16
5.4.8	Zaměření vašeho programu Data Governance	17
5.5	Životní cyklus Data Governance	18
5.6	Komponenty frameworku DGI Data Governance	18
5.6.1	Komponenta 1: Mise (poslání) a vize	19
5.6.2	Komponenta 2: Cíle, metriky, míry úspěchu, strategie financování	19
5.6.3	Komponenta 3: Definice dat a pravidel.....	20
5.6.4	Komponenta 4: Rozhodovací práva	20
5.6.5	Komponenta 5: Zodpovědnosti.....	20
5.6.6	Komponenta 6: Řízení, regulace a kontroly	21
5.6.7	Komponenta 7: Data stakeholders	21
5.6.8	Komponenta 8: Data Governance Office.....	21
5.6.9	Komponenta 9: Data Stewards.....	22
5.6.10	Komponenta 10: Procesy Data Governance	22
5.7	Výzvy.....	22
6.	Další zdroje informací.....	23

1. Kvalita dat

Kvalitu dat lze posuzovat z několika hledisek.

1.1 Hledisko mikroúrovně

Kvalitu dat lze posoudit lokálně – analýzou jednoho objektu a jeho okolí. Příklady kritérií kvality této úrovně:

- Přesnost
 - Polohová a výšková – přesnost určení souřadnic (očekávaná odchylka souřadnic od skutečné polohy)
 - Časová přesnost – aktuálnost
 - Tematická přesnost - přesnost určení atributů (hodnota atributu odpovídá číselníku nebo se nachází v zadaném intervalu, atributy splňují jiné požadované vlastnosti)
- Konzistence a integrita (jednoznačné relace mezi daty)
- Měřítko - rozlišení dat (nejmenší reprezentovatelná jednotka, nejmenší objekt zobrazený na mapě)

1.2 Hledisko makroúrovně

Kvalitu dat lze posoudit pouze analýzou celé datové sady. Příklady kritérií kvality této úrovně:

- Úplnost
 - úplnost pokrytí území
 - úplnost klasifikace (jak kvalitně zvolená klasifikace reprezentuje data)
 - úplnost verifikace dat
- Definice dat (existence a kvalita metadat), rodokmen dat (historie dat - zdroje a postup zpracování dat)
- Jazyková kvalita
- Projekce - kartografické zobrazení

1.3 Hledisko uživatelské úrovně

Kvalita dat z pohledu uživatele. Příklady kritérií:

- přístupnost dat
 - privátní data, veřejné databáze, ochrana údajů
 - interoperabilita dat
- přímá cena dat
- nepřímá cena dat (nepřímá cena dat způsobený vynuceným přepracováním nekvalitních dat může být vyšší než cena kvalitních dat)

2. Řízení kvality dat (procesy QA/QC)

QA - Quality assurance (zabezpečení kvality) je proces, kterým je měřena kvalita dat a ve kterém jsou implementovány prostředky pro zajištění kvality dat

QC - Quality control (řízení kvality) je proces, který zabezpečuje kvalitu dat v uživatelem očekávané úrovni

Pro zajištění procesu QA/QC disponují systémy kontrolními funkcemi pro kontrolu prostorových a popisných dat. Systémy umožňují provádění těchto kontrol

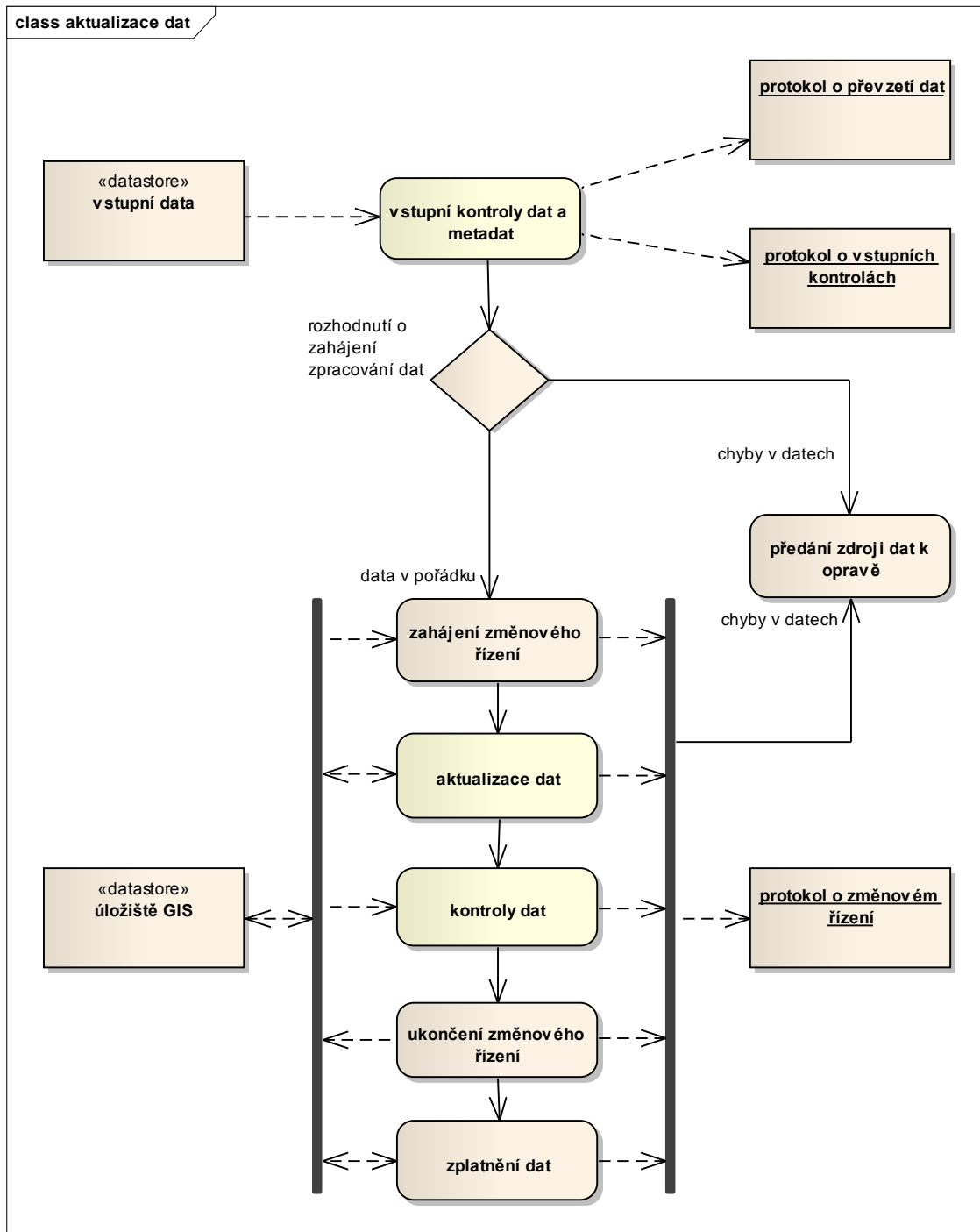
- při pořízení dat (speciálním případem jsou dynamické kontroly dat ala kontrola pravopisu a podtrhávání chyb v textových editorech
- při ukončení transakce
- dávkově v pravidelných intervalech nebo na vyžádání uživatele

Průběh kontrolních funkcí je řízen parametrickými předpisy, výsledky kontrolních funkcí jsou vizualizovány (vytváří se datové množiny obsahující chybové objekty nebo značky definující polohu chyby). Kontrolní funkce zásadně nemodifikují žádná data. Funkce, které upravují chybná data, se spouštějí v samostatných transakcích. Před zplatněním transakcí opět proběhnou kontroly dat.

Základní pracovní postup provádění kontrol obsahuje kroky:

- výběr dat, na která se mají aplikovat kontroly
- výběr typů kontrol, které se mají provést a případné nastavení parametrů
- výběr kontextových dat pro některé typy kontrol
- vyhodnocení výsledků kontrol
- případná oprava dat určených kontrolním nástrojem jako chybová data

Typický proces aktualizace dat, jehož součástí jsou kontroly dat



3. Funkce kontrolující kvalitu dat

3.1 Kontroly prostorových i popisných dat

- Kontroly primárních klíčů (constraint primary key)
- Kontroly unikátních klíčů (constraint unique key)
- Kontroly not NULL položek (constraint not null)
- Kontroly existence cíle pro cizí klíč (constraint foreign key)
- Kontroly hodnot jednotlivých atributů – datové typy
- Kontroly vztahu atributů v rámci jednoho záznamu (např. vztah mezi krajními body časových intervalů platnosti)
- Kontroly vztahu atributů v rámci několika záznamů jedné entity (např. po dvojicích disjunktní časové intervaly platnosti záznamů o jednom objektu)
- Kontroly vztahu atributů v rámci různých entit (vztahy mezi entitami neřešené constraintem)

Je vhodné mít k dispozici i kontroly, které kontrolují vlastnosti dat zabezpečené datovým modelem (datové typy, constrainty PK, UK, FK, NOT NULL, apod.). Důvody:

- constrainty mohou být z nějakého důvodu v databázovém schématu znepřístupněny
- při dávkovém vstupu dat (prvotní plnění, dávková aktualizace) je vhodné v pracovním schématu constrainty nepoužívat a kontrolovat vlastnosti celé dávky aplikačně – můžeme tak získat v jednom běhu úplný seznam chyb v datech

3.2 Příklad kontrol popisných dat

Příklad na datovém modelu entit Pojistenec – Zamestnani – Zamestnavatel (Zamestnani tvoří asociaci mezi entitami Pojistenec a Zamestnavatel).

- V tabulce Pojistenec nejsou různé instance se stejnou hodnotou Pojistenec.rodneCislo
- V tabulce Pojistenec nejsou různé instance se stejnou hodnotou Pojistenec.evidencniCisloPojistence
- Pokud Pojistenec.rodneCislo = NULL, pak Pojistenec.evidencniCisloPojistence není NULL. Pokud Pojistenec.rodneCislo není NULL, pak Pojistenec.evidencniCisloPojistence = NULL
- V tabulce Pojistenec neexistují dvě instance a,b takové, že Pojistenec(a).rodneCislo = Pojistenec(b).evidencniCisloPojistence
- V tabulce Zamestnavatel existuje pro skupinu instancí se shodnou hodnotou Zamestnavatel.identifikacniCisloOrganizace nanejvýš jedna instance s hodnotou Zamestnavatel.platnostDo = NULL
- Pro všechny instance v tabulce Zamestnani platí Zamestnani.zacatekVztahu <= Zamestnani.konecVztahu nebo Zamestnani.konecVztahu = NULL
- Cizí klíč Datovachranka.PojisteneclD je uspokojen
- Ve skupinách definovaných stejnou hodnotou Zamestnavatel.zdrojZamID, je na největší hodnotu Zamestnavatel.platnostDo (NULL je největší) navázána DS, pokud pro tuto největší hodnotu neexistuje záznam v DobaZam

3.3 Příklady funkcí kontrolujících kvalitu prostorových dat

Příklady kontrolních funkcí jsou převzaty z dokumentace ZABAGED.

Kontrola nadbytečného rozdělení liniových objektů

Význam kontroly:

Vyhledává liniové objekty stejného feature typu, které jsou rozděleny na více napojujících se linií, aniž je důvod k jejich rozdělení podle definovaných pravidel (v místech, kde dochází ke změně hodnoty atributu nebo v místě, kde se v lomovém bodě stýkají více jak 2 liniové objekty - "T napojení prvků").

Kontrolované objekty:

Kontrolují se specifikované liniové objekty. Předpokládá se, že se budou kontrolovat všechny feature typy, které mají geometrický typ definovaný jako "LINE".

Bude možné pustit kontroly pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Vzdálenost pro prostorové ztotožnění 2 bodů.
- Výčet kontrolovaných feature typů.

Ke každému kontrolovanému feature typu pak lze dále definovat:

- Množinu atributů jejich sémantické tabulky, jejichž hodnoty musí být shodné proto, aby se rozdělení linií vyhodnotilo jako nadbytečné.
- Množinu feature typů liniových objektů, které se mohou napojovat ve styčném bodě kontrolovaných linií, aniž by to znamenalo důvod k rozdělení kontrolovaných linií. Bude možno uvést i negaci seznamu.
- Množinu feature typů bodových objektů, které mohou ležet ve styčném bodě kontrolovaných linií, aniž by to znamenalo důvod k rozdělení kontrolovaných linií. Bude možno uvést i negaci seznamu.

Výsledek kontroly:

Kontrola umístí do chybové množiny značku chyby umístěnou v místě napojení nadbytečně rozdělených linií.

Kontrola nutného rozdělení liniových objektů

Význam kontroly:

Vyhledává liniové objekty, které v jejich průběhu kdekoliv kříží (i bez společného vrcholu) jiný liniový objekt daný množinou feature typů, případně ve vnitřním vrcholu leží nějaký bodový objekt, což je důvodem k rozdělení kontrolované linie v daném vnitřním vrcholu. Jsou definovány vztahy mezi jednotlivými feature typy, které určují pravidla pro vyhodnocení nutnosti "rozdělení".

Kontrolované objekty:

Kontrolují se specifikované liniové objekty. Bude možné pustit kontroly pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Vzdálenost pro prostorové ztotožnění 2 bodů.
- Výčet kontrolovaných feature typů.

Ke každému kontrolovanému feature typu pak lze dále definovat:

- Množinu atributů jejich sémantické tabulky, jejichž hodnoty musí být shodné proto, aby se rozdělení linií vyhodnotilo jako nadbytečné.
- Množinu feature typů liniových objektů, které se mohou napojovat ve styčném bodě kontrolovaných linií, aniž by to znamenalo důvod k rozdělení kontrolovaných linií (bude možné uvést i "negaci" seznamu).
- Parametr určující, zda důvodem k rozdělení kontrolované linie je napojení jiné linie vnitřními vrcholy nebo jen napojení jiné linie počátečním resp. koncovým vrcholem.
- Množinu feature typů bodových objektů, které mohou ležet ve styčném bodě kontrolovaných linií, aniž by to znamenalo důvod k rozdělení kontrolovaných linií (bude možné uvést i "negaci" seznamu).

Výsledek kontroly:

Kontrola umístí do chybové množiny značku chyby umístěnou v místě, kde má být provedeno rozdělení linie.

Kontrola volných konců liniových objektů

Význam kontroly:

Vyhledává liniové objekty, na jejichž počáteční nebo koncový vrchol není napojen jiný povolený liniový objekt, případně povolený bodový objekt.

Za napojení lze ve specifikovaných případech považovat i napojení do vnitřního vrcholu jiné linie.

Kontrolované objekty:

Kontrolují se specifikované liniové objekty. Bude možné pustit kontroly i pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Vzdálenost pro prostorové ztotožnění dvou bodů.
- Výčet kontrolovaných feature typů.
- Výčet feature typů, které definují objekt "použitelné" pro napojení kontrolovaného prvku.
- Příznak určující, zda kontrolovaná linie může/nemůže končit ve vnitřním vrcholu jiné linie.
- V případě možnosti ukončení ve vnitřním vrcholu, pak výčet liniových feature typů, ve kterých se může kontrolovaný objekt napojovat do vnitřního vrcholu.

Výsledek kontroly:

Kontrola umístí do chybové množiny liniový objekt a značku chyby v místě volného konce.

Kontrola prostorových vztahů mezi objekty

Význam kontroly:

Kontroluje prostorový vztah (prostorovou souvztažnost) objektů oproti předepsaným pravidlům. Především se jedná o kontrolu prostorových vztahů bodových nebo liniových objektů vůči liniovým objektům. Pokud daný typ bodového nebo liniového objektu neleží na specifikovaném výčtu typů liniových objektů, pak je vyhodnocen jako chybový. Typ požadovaného prostorového vztahu je určen parametrickým předpisem. V rámci vyhodnocování chyb se zohledňuje definované okolí grafických prvků. Součástí této kontroly může být také kontrola shodnosti hodnot určených atributů pro objekty v definovaném prostorovém vztahu.

Kontrolované objekty:

Kontrolují se specifikované bodové nebo liniové objekty. Bude možné pustit kontroly i pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Vzdálenost pro prostorové ztotožnění 2 bodů.
- Výčet kontrolovaných feature typů.

Pro každý kontrolovaný feature typ lze dále určit:

- Výčet feature typů liniových objektů, na kterém může být umístěn (v případě, že se kontroluje liniový objekt pak musí být umístěn všemi jeho vrcholy a počet vrcholů v souběhu musí souhlasit).
- Výčet dalších feature typů liniových objektů, na kterých musí mít umístěný alespoň 1 vrchol (v případě, že se kontroluje liniový objekt).
- Příznak určující, zda kontrolovaný objekt může/nemůže ležet na vnitřním vrcholech nebo jen koncových vrcholech liniových prvků.
- Pokud se má kontrolovat hodnota popisného atributu, pak se určuje název atributu kontrolovaného feature typu a název atributu feature typu, který vytváří daný prostorový vztah.

Výsledek kontroly:

Kontrola umístí do chybové množiny bodové nebo liniové objekty, které nesplňují požadovaný prostorový vztah k liniovému objektu.

Kontrola chybného překrývání malých plošných objektů

Význam kontroly:

Vyhledává prvky mající charakter menších plošných objektů a které se svojí obvodovou hranicí vzájemně překrývají. Charakter kontrolovaných prvků je dán také jejich geometrickým vyjádřením prvků, a to uzavřeným polygonem (který je pořízen jako 1 grafický prvek).

Kontrola také vyhledává umístění jednoho plošného útvaru v druhém plošném útvaru ("budova v budově").

Kontrolované objekty:

Kontrolují se specifikované plošné prvky, které mají charakter menších plošných objektů.

Parametrizace:

- Výčet kontrolovaných feature typů.
- Pro každý kontrolovaný feature typ lze určit výčet jiných feature typů plošných prvků, které bud překrývají / jsou uvnitř kontrolovaného prvku a přesto se tato situace nebude považovat za chybovou (např. uvnitř budovy může být obvod vnitrobloku).

Výsledek kontroly:

Kontrola umístí do chybové množiny takové prvky, které vykazují chybu.

Kontrola uzavřenosti plochy a jednoznačného přiřazení centroidu

Význam kontroly:

Vyhledává liniové objekty reprezentující hranice definovaných ploch, které netvoří uzavřené polygony. Vyhledává také liniové objekty, které se kříží s jinými liniovými objekty za předpokladu, že tyto liniové objekty reprezentující hranice definovaných ploch. Následně kontroluje počet centroidů v uzavřených polygonech seskládaných z jednotlivých liniových objektů. V případě, že se kontrola provádí v rámci změnového řízení a během aktualizace dat byly modifikovány objekty reprezentující hranice ploch nebo centroidy, pak dojde nejdříve k aktualizaci (přegenerování) ploch a aktualizaci příslušných topologických struktur. Výsledné přegenerované plochy bude možné zobrazit pomocí standardních funkcí klientského prostředí, což umožní i případnou vizuální kontrolu ploch.

Kontrolované objekty:

Kontrolované feature typy budou ty, které reprezentují všechny objekty reprezentující hranice ploch nebo centroidy ploch.

Parametrizace:

Příznak určující jaké podtypy kontrol se budou provádět:

- Vyhledání liniových objektů, které netvoří uzavřené polygony.
- Vyhledání ploch bez centroidu.
- Vyhledání ploch s více centroidy.

Výsledek kontroly:

Kontrola umístí do chybové množiny liniové objekty, které netvoří uzavřené polygony. Do chybové množiny se budou také ukládat plochy, které nemají jednoznačně přiřazený centroid.

Kontrola vyhledání sousedních ploch stejného typu (významu)

Význam kontroly:

Vyhledává sousedící plochy stejného feature typu resp. plochy stejného významu daného určeným atributem. Kontrola se de facto provádí nad liniovými objekty, které tvoří hranice ploch.

Kontrolované objekty:

Kontrolují se specifikované plošné objekty, jejichž hranice jsou "seskládané" z jednotlivých liniových objektů.

Předpokládá se, že se budou kontrolovat všechny feature typy, které budou tvořit plošnou topologii.

Bude možné pustit kontroly i pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Výčet kontrolovaných feature typů.
- Pro každý kontrolovaný feature typ lze určit atributy popisných dat, jejichž hodnota představuje význam plochy.
- Výčet feature typů objektů reprezentujících hranice ploch, které mohou dělit plochy stejného typu a přesto tato situace nebude považována za chybovou.

Výsledek kontroly:

Kontrola umístí do chybové množiny značku liniové objekty, které tvoří hranice sousedních ploch stejného významu.

Kontrola nepovolených prostorových kombinací objektů

Význam kontroly:

Vyhledává nepovolené kombinace prostorového umístění bodových, liniových nebo "malých" plošných objektů uvnitř daných ploch. Plochy zde budou reprezentovány plochami generovanými z jednotlivých liniových hranic.

Kontrolované objekty:

Kontrolují se specifikované plochy. Bude možné pustit kontroly i pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Výčet kontrolovaných feature typů reprezentující plochu.
- Pro každý kontrolovaný feature typ bude možné určit výčet feature typů bodových, liniových nebo "malých" plošných objektů, které nesmí ležet uvnitř kontrolované plochy a pro každý takový feature typ bude také určeno, zda je za chybu považováno to, že leží uvnitř plochy celý nebo je za chybu považováno to, když leží uvnitř plochy jen jeho část.

Výsledek kontroly:

Kontrola umístí do chybové množiny bodové, liniové nebo "malé" plošné objekty vyhodnocené jako chybové.

Kontrola návaznosti liniových objektů včetně jejich atributů

Význam kontroly:

Vyhledává liniové objekty, které jsou stejného typu a jsou vzájemně napojeny, ale u kterých se liší hodnota specifikovaného popisného atributu. Současně se může kontrolovat i fakt, zda navazující linie se stejnými hodnotami atributů tvoří uzavřený útvar (např. pro hranici chráněného území).

Kontrolované objekty:

Kontrolují se specifikované liniové objekty. Bude možné pustit kontroly i pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Vzdálenost pro prostorové ztotožnění 2 bodů.
- Výčet kontrolovaných feature typů.
- Pro každý kontrolovaný feature typ bude možné určit množinu atributů popisných dat, jejichž hodnoty mají být v rámci návaznosti linií stejné.
- Příznak určující, zda kontrolované linie mají tvořit uzavřený útvar.

Výsledek kontroly:

Kontrola umístí do chybové množiny liniové objekty, které vykazují chybu. Pravidla pro změnu hodnoty popisného atributu nelze vždy přesně algoritmicky predikovat, proto mnohdy půjde jen o "potenciální" chybu.

Kontrola prostorově vzájemně souvisejících objektů

Význam kontroly:

Kontrola vyhledává liniové objekty, které mohou prostorově souviset s jinými typy liniových objektů. Prostorovou souvislostí je zde myšlen přesný souběh linií nebo částí linií.

Kontrolované objekty:

Kontrolují se specifikované liniové objekty.

Parametrizace:

- Vzdálenost pro prostorové ztotožnění 2 bodů.
- Výčet kontrolovaných feature typů.
- Výčet feature typů, které reprezentují potenciálně prostorově souvisejícím liniové objekty.
- Vzdálenost, ve které se vyhledávají prostorově související objekty.

Výsledek kontroly:

Kontrola umístí do chybové množiny liniové objekty, které jsou vyhodnoceny jako chybové.

Kontrola duplicity liniových objektů

Význam kontroly:

Vyhledává liniové objekty, které jsou prostorově totožné (leží na sobě), případně jejichž některá část je prostorově totožná s částí jiného liniového objektu.

Kontrolované objekty:

Kontrolují se specifikované liniové objekty.

Předpokládá se, že se budou kontrolovat všechny feature typy, které mají geometrický typ definovaný jako "LINE". Bude možné pustit kontroly i pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Vzdálenost pro prostorové ztotožnění 2 bodů.
- Výčet kontrolovaných feature typů.
- Příznak určující, zda "duplicitní pár" musí tvořit objekty se stejným feature typem nebo zda jej mohou tvořit i objekty s různým feature typem (ve druhém případě se jako další parametr přidává výčet takových feature typů).
- Příznak určující, zda prostorově duplicitní linie musí/nemusí mít shodnou orientaci.
- Parametr určující u každého kontrolovaného liniového feature typu, jakým způsobem se vyhodnocuje prostorová duplicita:
- Celek - duplicita je vyhodnocena jen v případě, že celý průběh liniových objektů je totožný.
- Více segmentů - duplicita je vyhodnocena v případě, že 2 a více segmentů (úseček) liniových objektů jsou totožné.
- 1 Segment - duplicita je vyhodnocena již v případě, že 1 segment (úsečka) liniových objektů je totožný.

Výsledek kontroly:

Kontrola umístí do chybové množiny duplicitní liniové objekty a značku chyby umístěnou v místě těch vrcholů linií, které jsou prostorově totožné u obou linií.

Kontrola duplicity bodových objektů

Význam kontroly:

Vyhledává bodové objekty, které jsou prostorově totožné (leží na sobě).

Kontrolované objekty:

Kontrolují se specifikované bodové objekty. Předpokládá se, že se budou kontrolovat všechny feature typy, které mají geometrický typ definovaný jako "POINT". Bude možné pustit kontroly i pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Vzdálenost pro prostorové ztotožnění 2 bodů.
- Výčet kontrolovaných feature typů.
- Příznak určující, zda "duplicitní pár" musí tvořit objekty se stejným feature typem nebo zda jej mohou tvořit i objekty s různým feature typem.

Výsledek kontroly:

Kontrola umístí do chybové množiny značku chyby v místě vztažného bodu bodového objektu.

Kontrola minimální povolené velikosti objektů

Význam kontroly:

Vyhledává liniové a plošné objekty, jejichž velikost (délkový rozměr nebo plocha) je menší než povolená hodnota. U liniových objektů a plošných objektů, jejichž hranice jsou tvořeny liniovými objekty lze kontrolovat i minimální povolenou velikost jednotlivých segmentů (úseček).

Kontrolované objekty:

Kontrolují se specifikované liniové a plošné objekty. Bude možné pustit kontroly i pro jednotlivě definované skupiny feature typů.

Parametrizace:

- Výčet kontrolovaných feature typů.
- Parametr určující, zda se kontroluje plocha, celková délka linie nebo délka jednotlivých úseček, případně povolený poměrový vztah plocha vs. minimální velikost hranice plochy.
- Parametr obsahující minimální povolenou hodnotu velikosti (poměru).

Výsledek kontroly:

Kontrola umístí do chybové množiny takové objekty, které vykazují chybu.

Kontrola hodnot popisných atributů

Význam kontroly:

Kontrola je zaměřená na vyhledání objektů, jejichž hodnoty jsou mimo definovaný rozsah hodnot nebo jejichž hodnoty nejsou vyplněné.

U některých popisných atributů bude zadání hodnoty v rámci povoleného rozsahu případně zadání povinného atribut řízeno přímo definicí metamodelu projektu nebo na úrovni definice struktury databázových objektů.

Kontrolované objekty:

Kontrolují se specifikované atributy s popisnými daty u daných feature typů.

Parametrizace:

Programově budou kontroly zřejmě rozděleny na dva základní typy, kterým bude poskytnut jiný výčet parametrů.

Pro kontrolu na vyhledání atributů, jejichž hodnoty nejsou vyplněné:

- Výčet kontrolovaných feature typů a jejich atributů, které nesmí být prázdné (musí být vyplněné).

Pro kontrolu na vyhledání atributů, jejichž hodnoty jsou mimo povolený rozsah:

- Výčet kontrolovaných feature typů a jejich atributů, které se mají kontrolovat.
- Příznak určující, zda prázdná hodnota je pro hodnotu atributu akceptovatelná.
- Parametr určující typ operátoru pro vyhodnocení hodnot (=, >, <, !=, IN, BETWEEN, LIKE).
- Hodnoty (hodnoty) určující výčet nebo rozsah povolených hodnot.

Výsledek kontroly:

Datový výběr obsahující prostorová a popisná data pro objekty vyhodnocené jako chybové.

4. Příčiny nekvalitních dat a možná řešení



Příčina	Příklady	Řešení
Nevhodný datový model a databáze	Aplikačně realizované constrainty	Vytvořit vhodný DM a použít robustní databázi
	Datový model je příliš obecný, takže neumožňuje využít constrainty	
	Nevhodný prostředek pro ukládání dat (soubory, nevhodná databáze)	
Chybné aplikace	Nedostatečné kontroly dat při jejich změnách	Důkladné testování aplikací
Špatné postupy při prvotním plnění dat	Chyby ve vstupních datech	Kvalitní podpora procesů prvotního plnění dat
	Chyby v procesech prvotního plnění	
	Nedostatečné kontroly	
	Záměr dodavatele skrýt informace o chybných datech	
	Záměr uživatele skrýt informace o chybných datech	
Špatné postupy při aktualizaci dat	Chyby ve vstupních datech	Kvalitní podpora procesů aktualizace dat
	Chyby v procesech aktualizace	
	Nedostatečné kontroly	

Kvalitní data pomohou zajistit projekty data governance.

5. Principy Data Governance

5.1 Co je Data Governance?

Krátká definice: Data Governance (DG) je autorita, která je pověřena výkonem rozhodování v oblastech souvisejících s daty.

Trochu delší definice: Data Governance je systém rozhodovacích práv a zodpovědností v procesech, které souvisejí s daty a informacemi a které jsou prováděny podle dohodnutých modelů. Tyto modely popisují:

- **KDO:** Kdo může pracovat s informacemi
- **CO:** Jaké akce může s informacemi provádět
- **S ČÍM:** S jakými informacemi může tyto akce provádět
- **KDY:** Kdy může tyto akce provádět
- **KDY:** Za jakých okolností může tyto akce provádět
- **JAK:** Jakými metodami může tyto akce provádět

Data Governance se vztahuje na organizační složky, pravidla, rozhodovací práva a odpovědnosti lidí a informačních systémů, které vykonávají procesy vztahující se k informacím.

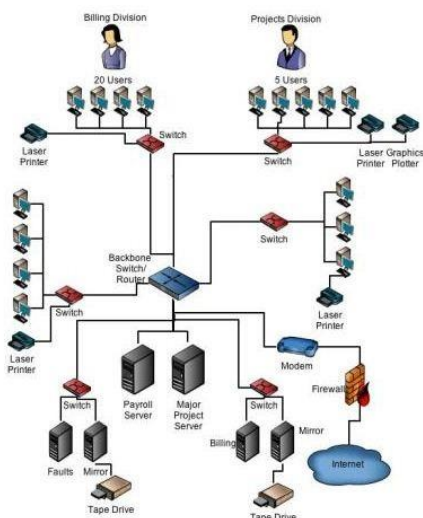
Data Governance samozřejmě nikdy nenahradí řízení. Jen je doplňuje. Governance přichází do hry, když jednotliví manažeři zjišťují, že nemohou - nebo by neměli - provádět nezávislá rozhodnutí. Governance sdružuje mezioborové týmy, které vytvářejí vzájemně propojená pravidla, řeší problémy a poskytují služby stranám zainteresovaným na datech (data stakeholders¹).

Členové těchto mezioborových týmů - datoví správci (Data Stewards) a/nebo datoví guvernérů (Data Governors) - obecně pocházejí z business oblasti organizace. Nastavují politiku, podle které IT skupiny a datové skupiny pracují: stanovují architekturu, implementují vlastní osvědčené postupy a adresují požadavky. Za governance může být považován celkový proces provádění této práce.

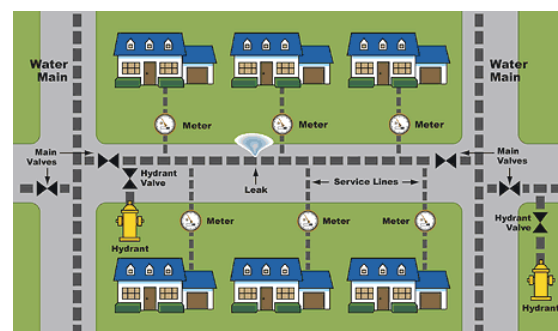
5.2 Data Governance a IT Governance

Pokud chceme vědět, jaký je rozdíl mezi Data Governance a IT Governance, zkusme odpovědět na jinou otázku: jaký je rozdíl mezi pojmem data nebo informace a pojmem informační technologie? Pro odpověď zvolme analogii s rozvodem vody: informační technologie hraje roli systému potrubí, čerpadel, ventilů, šoupat, zásobníků, data mají roli vody, která teče těmito zařízeními.

**Informační technologie a data
KVALITA DAT**



**Rozvod vody a voda
KVALITA VODY**



¹ <http://www.businessvize.cz/strategie/kdo-jsou-to-vlastne-stakeholders-a-proc-a-jak-se-o-ne-zajimat>

Obáváte se, že voda, která vaším potrubím teče, je otrávena nebo znečištěna? Zavoláte instalatéra? Samozřejmě, že ne! Instalatéři jsou specialisté na potrubí, čerpadla a zásobníky, a ne na to, co jimi teče. Zavoláte si odborníky, kteří vědí, jak testovat kvalitu vody - specialisty, kteří poznají rozdíl mezi pitnou vodou a dalšími typy čirých kapalin.

Velké organizace musejí dělat mnoho rozhodnutí o svých informačních systémech a datech, která obsahují a která jimi proudí. Mnoho z těchto rozhodnutí je přijímáno ve skupině IT zaměřené na IT problémy, jako je rozhodování, kdy je na čase nasadit novou aplikaci se všemi novými trubkami, čerpadly a nádržemi. Nicméně některé zájmy, obavy, starosti si do takové skupiny nikdy nenajdou cestu. Otázky týkající se toho, co „protéká trubkami“ vyžadují mezioborovou skupinu, která přináší odborné znalosti potřebné k rozhodnutím týkajících se dat.

5.3 Co lze pomocí Data Governance dosáhnout?

Implementací konkrétního programu Data Governance můžete očekávat dosažení následujících univerzálních cílů.

1. Umožnit lepší rozhodování
2. Snížit neshody a konflikty mezi lidmi
3. Chránit potřeby stran zainteresovaných na datech (data stakeholders)
4. Proškolit vedení a zaměstnance, aby akceptovali společné přístupy k problematice dat
5. Vybudovat standardní, opakovatelné procesy a zajistit transparentnost těchto procesů
6. Snížit náklady a zvýšit efektivitu práce týkající se dat

Čeho lze ještě dosáhnout? Nejběžnějším cílem programů Data Governance je standardizace definic dat v celé organizaci nebo alespoň v její části: v rámci nějaké aktivity nebo projektu. Další cíle závisí na zaměření konkrétních programů.

5.4 Typické oblasti pro Data Governance

5.4.1 Data Governance se zaměřením na politiku, standardy a strategii

Tento typ programu je vhodný, pokud některé skupiny v rámci organizace potřebují podporu pro vedení mezioborových týmů. Například společnosti přecházející od vývoje typu „silo“ na vývoj enterprise systémů mohou mít vývojové týmy, které se vzpírají vedení datových architektů a modelářů. Formální politiky Data Governance podporované mezioborovými Stewardsy mohou dodat váhu pozicím architektů.

Ze zaměření programu Data Governance na politiku, standardy a strategii mohou těžit také další enterprise iniciativy, jako je Enterprise Data Management (EDM), Business Process Reengineering (BPR), standardizace platform a získávání (akvizice) datových sad a systémů. Často tyto typy programů začínají tím, že se soustředí na soubory master dat a/nebo metadat.

Charta² tohoto typu programu vede účastníky Data Governance k odpovědnosti vůči:

- *hodnocení, schvalování, sledování politiky*
- *sběru, výběru, přezkoumání, schvalování, sledování norem*
- *konsolidaci sady politik a standardů*
- *příspěvkům k business pravidlům*
- *příspěvkům k datovým strategiím*
- *identifikaci data stakeholderů, nastavení rozhodovacích práv a zpřesnění odpovědností*

5.4.2 Data Governance se zaměřením na kvalitu dat

Tento typ programu je používán v případě problémů spojených s kvalitou, integritou a použitelností dat. Může být sponzorován skupinou Data Quality nebo business týmem, který potřebuje data lepší kvality. Často je úsilí o kvalitu dat v počátku nasměrováno na master (kmenová) data. Tyto typy programů téměř vždy zahrnují Data Quality software. Mohou začít se zaměřením na celý podnik, jedno oddělení nebo projekt.

Charta tohoto typu programu vede účastníky Data Governance k odpovědnosti vůči:

² Základní dokument projektu, Ustavující listina projektu.

- *nastavení směřování na kvalitu dat*
- *sledování kvality dat*
- *reportování stavu iniciativ zaměřených na kvalitu*
- *identifikaci data stakeholderů, nastavení rozhodovacích práv a zpřesnění odpovědností*

5.4.3 Data Governance se zaměřením na Privacy/Compliance/Security

Tento typ programu je typicky používán, pokud existují obavy o Data Privacy, Access Management/Permissions, řízení informační bezpečnosti, dodržování regulačních, smluvních nebo interních požadavků. Program může být sponzorován business odděleními, IT, nebo být přímým důsledkem programu Governance Risk, and Compliance (GRC) (Řízení rizik a shody). Pro tento typ programu často poskytuje mandát vrcholové vedení.

Tyto programy obvykle začínají v rozsahu celého podniku, ale často jsou omezeny na konkrétní typy dat. Téměř vždy zahrnují technologie pro vyhledávání citlivých dat, k ochraně dat a/nebo ke správě politik nebo kontrol.

Charta tohoto typu programu vede účastníky Data Governance k odpovědnosti vůči:

- *zajištění ochrany citlivých dat prostřednictvím Access Managementu a bezpečnostních požadavků*
- *konsolidace frameworků a iniciativ*
- *pomoci zhodnotit rizika a určit kontroly pro řízení rizik*
- *pomoci vynutit dodržování požadavků z pohledu právního, smluvního, architektonického*
- *identifikaci data stakeholderů, nastavení rozhodovacích práv a zpřesnění odpovědností*

5.4.4 Data Governance se zaměřením na Architekturu/Integraci

Tento typ programu je typicky vhodný ve spojení s pořízením, vývojem nebo aktualizací hlavního systému, který vyžaduje nové úrovně mezioborového rozhodování a odpovědnosti. Jiný driver takového programu je přechod na Service Oriented Architecture (SOA), s jeho potřebou dobře spravovaných dat, nebo nové zaměření na metadata, Master Data Management (MDM) nebo Enterprise Data Management (EDM).

Charta tohoto typu programu vede účastníky Data Governance k odpovědnosti vůči:

- *zabezpečení konzistentních definic dat*
- *podpoře architektonických zásad a norem*
- *podpoře programů Metadata, SOA, Master Data Management, Enterprise Data Management*
- *integračním výzvám z mezioborového pohledu*
- *identifikaci data stakeholderů, nastavení rozhodovacích práv a zpřesnění odpovědností*

5.4.5 Data Governance se zaměřením datové sklady a business intelligence

Tento typ programu je typicky vhodný ve spojení s konkrétním datovým skladem, datovým tržištěm (data mart) nebo BI nástrojem. Tyto snahy vyžadují závažná rozhodnutí o datech a organizace často implementuje governance, aby pomohla počátečním rozhodnutím, která mají podpořit rozhodnutí následná a prosazení standardů a pravidel, až se nový systém stane funkčním.

Rozsah může být zpočátku omezen na pravidla, role a odpovědnosti za nový systém, ale někdy tento typ programu slouží jako prototyp pro celopodnikový program Data Governance.

Charta tohoto typu programu vede účastníky Data Governance k odpovědnosti vůči:

- *stanovení pravidel pro užívání dat a pro definice dat*
- *identifikaci data stakeholderů, nastavení rozhodovacích práv a zpřesnění odpovědností*
- *identifikaci kroků DG v systems development life cycle (SDLC)*
- *vyjasnit hodnoty datových aktiv a projektů, které souvisejí s daty*

5.4.6 Data Governance se zaměřením podporu managementu (management alignment)

Tento typ programu je vhodný, pokud manažeři jen obtížně dělají "běžná" rozhodnutí vztažená k datům, protože tato rozhodnutí mají velký potenciální dopad na chod podniku nebo dodržování předpisů v podniku. Manažeři si uvědomují, že se potřebují sejit, aby rozhodnutí udělali společně, ale buď neznají všechny zainteresované strany, které do rozhodování musí zahrnout, nebo existuje nějaká překážka k jejich seznání. V takových případech formální program Data Governance může dát manažerům větší jistotu při rozhodování.

Někdy se tyto programy skládají především z rad (councils), které se scházejí k analýze vzájemných závislostí, k rozhodování a k vydávání politik. Jindy má program DG několik oblastí, jako je podpora řízení a zabezpečení dodržování předpisů.

Charta tohoto typu programu vede účastníky Data Governance k odpovědnosti vůči:

- měření hodnoty dat a s daty souvisejících prací
- konsolidace frameworků a iniciativ
- identifikaci data stakeholderů, nastavení rozhodovacích práv a zpřesnění odpovědností
- identifikaci kroků DG v systems development life cycle (SDLC)
- sledování a vykazování projektů souvisejících s daty
- podpoře vydávání zpráv a stanovisek souvisejících s daty

5.4.7 Přehled aktivit v jednotlivých oblastech

Aktivita	Oblast	Politika, standardy, strategie	Kvalita dat	Privacy, Compliance, Security	Architektura, integrace	Datové sklady a BI	Podpora managementu
Hodnocení, schvalování a sledování politiky		+					
Sběr, výběr, přezkoumání, schvalování a sledování norem		+					
Konsolidace politik a standardů		+					
Příspěvky business pravidlům		+					
Příspěvky k datovým strategiím		+					
Identifikace data stakeholderů, nastavení rozhodovacích práv a zodpovědností		+	+	+	+	+	+
Nastavení směru pro kvalitu dat			+				
Sledování kvality dat			+				
Reportování stavu iniciativ zaměřených na kvalitu			+				
Zajištění ochrany citlivých dat prostřednictvím Access Managementu a bezpečnostních požadavků				+			

Aktivita	Oblast	Politika, standards, strategie	Kvalita dat	Privacy, Compliance, Security	Architektura, integrace	Datové sklady a BI	Podpora managementu
Konsolidace frameworků a iniciativ				+			
Pomoc při zhodnocení rizik a určení kontrol pro řízení rizik				+			
Pomoc při vynucení dodržování požadavků z pohledu právního, smluvního, architektonického				+			
Zabezpečení konzistentních definic dat					+		
Podpora architektonických zásad a norem					+		
Podpora programů Metadata, SOA, Master Data Management, Enterprise Data Management (EDM)					+		
Cross-functional pozornost k integračním výzvám					+		
Stanovení pravidel pro užívání dat a definice dat						+	
Identifikace kroků Data Governance v systems development life cycle (SDLC)						+	+
Vyjasnění hodnot datových aktiv a s daty souvisejících projektů						+	+
Konsolidace frameworků a iniciativ							+
Sledování a vykazování projektů souvisejících s daty							+
Podpora vydávání zpráv a stanovisek souvisejících s daty							+

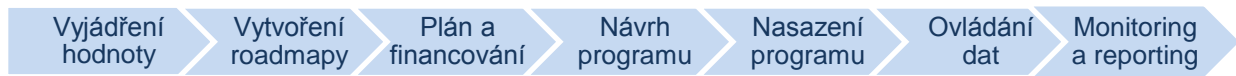
5.4.8 Zaměření vašeho programu Data Governance

Na co bude váš Data Governance program zaměřen? Rozhodnutí bude vyplývat z toho, jaké zájmy a potřeby mají účastníci vašeho programu Data Governance. To bude ovlivňovat složení data stakeholderů zapojených do vašich akcí a rozhodování a množství práce vyžadované vašimi stakeholdery.

Kdo je data stakeholder? Je to každý jednotlivec nebo skupina, která by mohla ovlivňovat nebo být ovlivněna diskutovanými daty. Někteří stakeholderi jsou zřejmí - obchodní skupiny, IT týmy, datoví architekti a databázoví administrátoři. Ostatní stakeholderi nemusí být pro rozhodování v konkrétní situaci tak zřejmí. Vědět, které stakeholdery přivést ke stolu k jednání a kdy to udělat, je odpovědností týmu Data Governance.

5.5 Životní cyklus Data Governance

Všechny programy mají životní cyklus. Životní cyklus Data Governance má následujících sedm fází.



Všimněte si, že Data Governance nezačíná návrhem programu.

- *Než začnete rozhodování o tom, kdo bude v jakém výboru, měli byste mít jasno o hodnotách, které váš program přinese.*
- *Měli byste vytvořit roadmapu odsouhlasenou a sdílenou se stakeholdery.*
- *Než se rozhodnou pro podporu vašeho programu, budou chtít tyto subjekty znát odpovědi na otázky: KDO/CO/KDY/KDE/PROČ/JAK. Musíte předvídat jejich otázky. Budete potřebovat znát předběžné odpovědi, i když do doby, než uděláte aktuální návrh programu, to budou jen předpoklady.*

Realizace prvních tří fází životního cyklu Data Governance je obtížná. Všichni jsme slyšeli říkat, že je těžké vidět les pro stromy. Co když jste jedním ze stromů?

Při provádění činností potřebných k získání podpory a financování, nezapomeňte, že váš program může plánovat řešení více zájmových oblastí. Každý nový program nebo programová oblast by měla být zavedena pomocí sedmi kroků životního cyklu.

Poznámka k závěrečné fázi životního cyklu Data Governance: pokaždé, když uvažujete o novém souboru činností, předvídejte očekávání stakeholderů týkající se průběžného sledování prací, měření úspěšnosti, a reportování stavu. Vaše schopnost dodávat standardní metriky, které splňují očekávání stakeholderů, může způsobit právě ten rozdíl, který rozděluje programy na chronicky lopotné a neúspěšné a programy, které se staly rutinou.

5.6 Komponenty frameworku DGI Data Governance

Framework DGI Data Governance se skládá z deseti komponent.

Pravidla a povinnosti (Rules of Engagement)

1. Poslání a vize
2. Cíle, metriky governance a míry úspěchu, strategie financování
3. Datová pravidla a definice
4. Rozhodovací práva
5. Zodpovědnosti
6. Řízení a kontroly

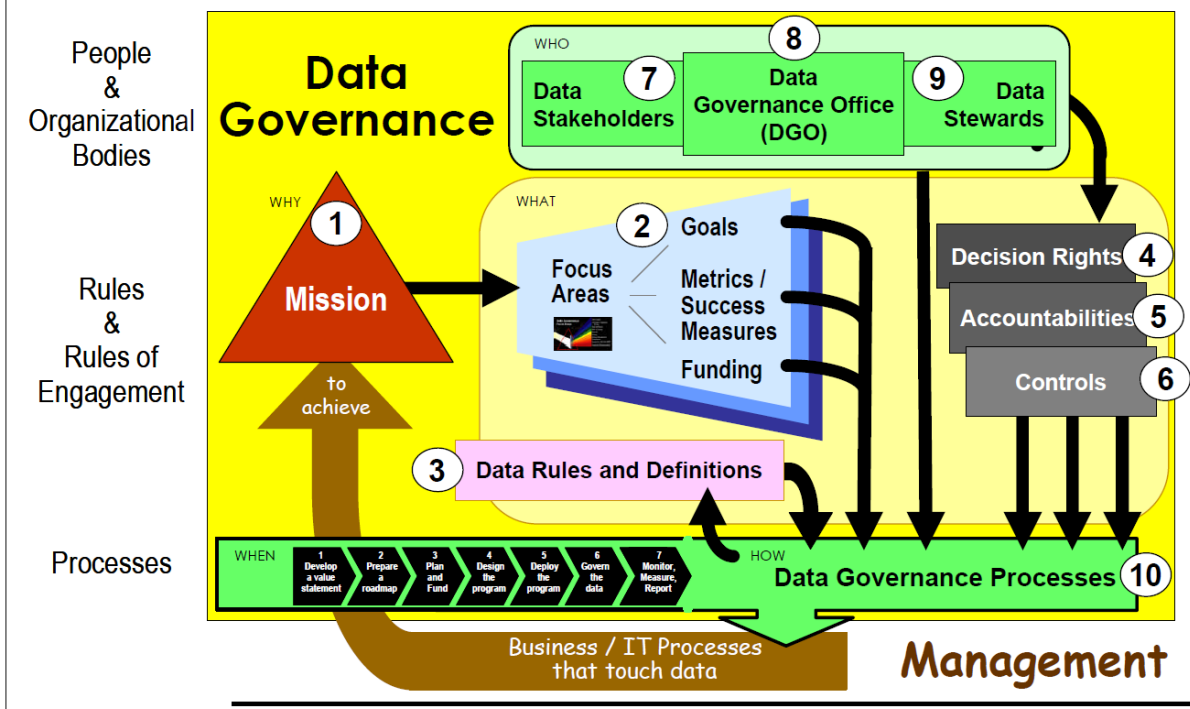
Lidé a organizační složky

7. Data stakeholders
8. Data Governance office
9. Data Stewards

Procesy

10. Proaktivní, reaktivní a průběžné procesy Data Governance

10 Components of a Data Governance Program



5.6.1 Komponenta 1: Mise (poslání) a vize

Na nejvyšší úrovni má obvykle mise Data Governance tři části:

- 1) Aktivně definovat a/nebo sladit pravidla.
- 2) Zajistit trvalou ochranu dat a služeb pro stakeholdery, která překlenuje hrance mezi organizačními jednotkami.
- 3) Reagovat a řešit otázky vyplývající z nedodržování pravidel.

Je to podobné jako s funkcemi, které poskytují tři složky politické správy. Jedna složka vlády - moc zákonodárná - stanoví pravidla ve formě zákonů. Další složka - výkonná složka - spustí tato pravidla a zároveň poskytuje služby. Třetí složka - soudní – se zabývá porušováním pravidel, a pokud je to nutné, interpretuje zákony a řeší rozpory mezi různými soubory pravidel a předpisů.

Není to o tom, že váš program Data Governance potřebuje tři větve, tři části. Struktury vašeho businessu a řízení IT již fungují v "exekutivní složce". Často stejná skupina Data Stewardů, která stanoví pravidla, bude také řešit konflikty.

Nicméně, jak budete budovat skupinu pro Data Governance, definovat role a odpovědnosti, je vhodné si uvědomit, že senátoři a soudci požadují podpůrný personál. Chcete-li být úspěšní při stanovení a vymáhání politiky, standardů a dalších typů pravidel, váš program Data Governance bude muset poskytnout tento typ podpory.

Ujistěte se, že jste vytvořili jasnou vizi. Jak by vaše organizace měla vypadat s vyspělou Data Governance? Jak asi bez ní?

Poznámka: Sdělení o poslání mohou být suchá. Na druhé straně, jazyk, kterým sdělujete své vize, by měl být bohatý, evokující, přesvědčivý. Vaše vize by měla mít možnost inspirovat zúčastněné strany, pomoci jim představit možnosti, podpořit cíle související s daty.

5.6.2 Komponenta 2: Cíle, metriky, míry úspěchu, strategie financování

Některé z cílů vašeho programu mohou vyústit do "soft" výsledků, které jsou neoficiální nebo které lze těžko měřit. Další by měly být SMART: Specific, Measurable, Achievable, Relevant and Timely (konkrétní, měřitelné, dosažitelné, relevantní a aktuální). Jak se rozhodnout, kterým cílům byste se

měli věnovat? Začněte tím, že budete předvídat účinek governance na "4P": programy, projekty, profesní obory, people (lidi jako jednotlivci). Ptejte se, jak by vaše úsilí by mohlo pomoci v podnikových programech (nebo high-profil projektech):

- *zvýšit příjmy a hodnotu*
- *spravovat náklady a složitost*
- *zaručit přežití*

Ptejte se, jak by váš program mohl podpořit práci architektů, kvalitu, vývoj aplikací nebo jiné profesní disciplíny. Ptejte se sami sebe, co si přejí nebo co bolí klíčové pracovníky a co by mohl splnit silný program Data Governance.

A nezapomeňte se podívat na data samotná. Jak můžete ovlivnit množství, kvalitu nebo ochranu dat a metadat? Ptejte se, jak lze měřit takový účinek.

Metriky - stejně jako cíle - by měly být SMART. Každý, kdo je zapojen do Data Governance, by měl vědět, jak vypadá úspěch a jak se měří. Zvažte vytvoření hodnoty podle vzorce:

Jestliže uděláme A, měli bychom očekávat B s výsledkem C. V opačném případě bychom měli očekávat D s výsledkem E.

Tak zřejmé vymezení hodnoty pomáhá nalézt možnosti financování pro váš program. S vašimi klíčovými stakeholdery vyjasněte:

- jak byste mohli financovat Data Governance Office (nebo její ekvivalent)
- jak byste mohli financovat čas datových analytiků a architektů potřebný k pomoci definovat pravidla, definovat data, a výzkumné problémy, které musí být vyřešeny
- jak byste mohli financovat aktivity Stewardship
- jaké protokoly je třeba stanovit pro business a IT pracovníky, kteří
 - pomohou při definici dat
 - budou analyzovat datové problémy
 - pomohou při řešení datových problémů

5.6.3 Komponenta 3: Definice dat a pravidel

Tato komponenta se týká politik, norem, compliance požadavků, obchodních pravidel souvisejících s daty a datových definic. V závislosti na oblastech zaměření může váš program pracovat na:

- vytvoření nových pravidel/definic
- shromáždění stávajících pravidel/definic
- vymezení mezery a přesahy
- konsolidace a stanovení priorit konfliktních pravidel definic
- vytvoření nebo formalizace pravidel pro situaci, kdy jsou aplikovány určité definice

5.6.4 Komponenta 4: Rozhodovací práva

Před tím, než je vytvořeno pravidlo nebo uděláno rozhodnutí související s daty, je třeba řešit a priori to, kdo bude dělat rozhodnutí a jakým procesem. Toto je odpovědností programu Data Governance. Cílem je usnadnit shromažďování rozhodovacích práv (a někdy také jejich dokumentaci a uchovávání). Rozhodovací práva jsou vlastně metadata o rozhodnutích souvisejících s daty.

Rozhodovací práva pro programy compliance-based lze často jednoduše definovat. Například: mělo by být rozhodnutí o tom, zda je rozhodnutí v souladu se zákonem, ponecháno na hlasování těch, kteří budou muset akceptovat s rozhodnutí? Samozřejmě že ne, rozhodnutí musí být provedeno správnou radou (představenstvem) organizace na základě vstupu z právního oddělení.

U ostatních typů programů může stanovení rozhodovacích práv vyžadovat více jednání. Například: kdo by měl rozhodnout o délce datového pole v novém systému? Možná by rozhodnutí mělo být provedeno datovým architektem. Ale možná to vyžaduje vstup z mnoha zúčastněných stran. Možná, že jedna z nich přijde s omezením, kterému je třeba podřídit rozhodnutí.

5.6.5 Komponenta 5: Zodpovědnosti

Jakmile je vytvořeno pravidlo nebo je uděláno rozhodnutí související s daty, organizace musí být připravena jednat podle tohoto rozhodnutí. Kdo by měl jednat a kdy? U činností, které nejsou úhledně

mapovány na povinnosti jednotlivých oddělení, můžeme očekávat, že program DG definuje zodpovědnosti, které mohou být zahrnuty do každodenních procesů v organizaci.

To platí zejména pro programy se zaměřením na compliance (dodržování předpisů). Organizace někdy bojuje s přiřazením odpovědností pro compliance aktivity. Proč? Zprvu proto, že compliance - jako governance - je obecně iniciativa, která překračuje hranice organizačních útvarů a která vyžaduje konsolidaci mezioborových aktivit. Za druhé, mnoho manažerů, kteří jsou zbláhli v plánování pro řízení jejich specifické oblasti, má méně zkušeností s požadavky, které přicházejí s Post-Compliance paradigmatem. Toto nové paradigma říká, že práce na compliance požadavku nekončí tím, že to uděláš. Končí, až pokud to uděláš, zkontroluješ, zdokumentuješ a prokážeš shodu.

Jednotliví manažeři často nejsou připraveni identifikovat všechny úkoly a integrační body pro navrhování a provádění kontroly, dokumentace a auditovatelné důkazy shody. Opravdu, v dodržování životního prostředí (compliance environment), jednotliví manažeři nesmí interpretovat požadavky samostatně. Místo toho firmy často přecházejí k modelu, kdy centralizovaná skupina tyto požadavky rozvíjí a rozšiřuje je stakeholderům. Někdy je Data Governance požádána o pomoc s požadavky a odpovědnostmi souvisejícími s daty.

Poznámka: Ať už pracujete s požadavky na životní prostředí nebo s jinými předpisy, většina aktivit musí být dělána mezioborovými týmy. Vaši koordinátoři v oblasti Data Governance budou muset pochopit aktivity a řídit se protokoly platnými ve vaší organizaci.

5.6.6 Komponenta 6: Řízení, regulace a kontroly

Je dobře známo, že data jsou neustále v ohrožení. Narušení citlivých dat má důsledky pro ty, kterým byla tato data svěřena data. Je tedy zřejmé že data mohou představují riziko. Jak si poradíme s tímto rizikem? Budeme je kontrolovat, nejlépe budeme předcházet události, u kterých nechceme, aby se objevily. U těch událostí, u kterých si nejsme jisti, že jim dokážeme předejít, musíme alespoň zajistit odstranění problémů, pokud už události nastanou. Strategie řízení rizik v organizaci je prováděna pomocí prostřednictvím regulačních opatření.

Často je program DG požádán o doporučení, jaká regulační opatření k podpoře cílů Data Governance by měla být aplikována. To se týká více úrovní: sítě, operačních systémů, databází, aplikací, uživatelských postupů.

Data Governance může být také požádána o doporučení, jakým způsobem stávající obecná opatření (řízení změn, politiky, školení, modely životního cyklu a projektové řízení, a další) by měla být upravena pro podporu cílů Data Governance a podnikových cílů.

5.6.7 Komponenta 7: Data stakeholders

Data stakeholderi pocházejí z celé organizace. Patří do skupin, které vytvářejí data, které používají data a které vytvářejí pravidla a požadavky na data. Protože data stakeholderi mají jednak vliv na rozhodování a jednak jsou ovlivněni rozhodnutími vztaženými k datům, budou očekávat, že musí být zahrnuti do programu Data Governance.

Někteří z nich budou očekávat, že mají být zahrnuti do některých typů rozhodnutí souvisejících s daty. Někteří z nich budou očekávat, že budou přizváni ke konzultacím před tím, než budou rozhodnutí formalizována a další budou spokojeni s tím, že jsou informováni o rozhodnutích poté, co jsou udělána.

Často podmnnožina výkonných stakeholderů tvoří Data Governance Board, aby měli dohled nad programem, nad politikami a nad řešením problémů. Jindy je dohled nad Data Governance zajišťován stávajícími organizačními složkami, jako je Řídící výbor IT a Výkonný výbor IT.

5.6.8 Komponenta 8: Data Governance Office

Dříve jsme poznamenali, že vláda, senát a soudci musí mít podpůrný personál. Stejně tak potřebují podpůrný personál lidé ve vaší organizaci, kteří dělají rozhodnutí související s daty, kteří definují data, monitorují soulad s pravidly a řeší problémy.

Data Governance Office (DGO) usnadňuje a podporuje tyto governance aktivity. DGO sbírá hodnoty metrik pro měření úspěchu a reportuje o nich data stakeholderům. DGO poskytuje průběžnou péči stakeholderům v podobě komunikace, přístupu k informacím, vedení evidence, vzdělávání a poskytuje data stakeholderům i další podporu.

5.6.9 Komponenta 9: Data Stewards

Rada (výbor) pro Data Stewardship (nebo Governance) se skládá ze souboru Data Stakeholderů, kteří se setkávají, aby dělali rozhodnutí související s daty. Mohou stanovit politiku a stanovit standardy nebo mohou vytvořit doporučení, kterými se řídí vyšší úroveň vedení skupin Data Governance.

Někdy - zejména pro velké organizace – je jediná úroveň stewardů nedostatečná. V tomto případě může existovat hierarchie stewardů. Ve větších či menších organizacích mohou členové Rady pro Data Stewardship vytvořit týmy nebo pracovní skupiny, které řeší specifické problémy dat nebo rozhodnutí.

Data Governance programy s důrazem na kvalitu dat mohou zahrnovat i Data Quality Stewards. Tyto role (s odvozenou působností v Data Governance) typicky reportují týmu pro business funkce nebo týmu pro kvalitu dat. Stewardi zkoumají množiny dat podle kritérií na úplnost, správnost a integritu. Jiné typy stewardů mohou mít jiné konkrétní odpovědnosti související s daty.

5.6.10 Komponenta 10: Procesy Data Governance

Komponenty 1 – 6 z DGI Frameworku Data Governance se zabývají pravidly. Snaží popsat „pravidla angažovanosti“ (povinnosti, rules of engagement), která využívají komponenty 7 – 9 (lidé a organizační jednotky) během geovernance. Tato poslední složka - procesy – popisuje metody, kterými budou spravována (govern) data. Jedná se proaktivní, reaktivní a „ongoing“ procesy.

V ideálním případě by tyto procesy měly být standardizované, dokumentované a opakovatelné. Měly by být vytvářeny tak, aby podporovaly regulační požadavky a compliance požadavky na správu dat, privátnost, bezpečnost a správu přístupu.

Každá organizace bude rozhodovat, jak strukturovaný a formální bude proces vládnutí daty. Data Governance Institute (DGI) doporučuje (a běžně implementuje) formální, dokumentované, opakovatelné postupy pro:

1. Vyrovnávání (aligning) politik, požadavků a řízení
2. Zřízení rozhodovacích práv
3. Zřízení zodpovědností
4. Vykonávání Stewardship
5. Řízení změn
6. Definování dat
7. Řešení otázek a problémů
8. Specifikaci požadavků na kvalitu dat
9. Vestavění governance do technologie
10. Péči o stakeholdery
11. Komunikacei
12. Měření a vykazování hodnoty

5.7 Výzvy

Na první průmyslové Data Governance konferenci v Orlando, Florida USA v prosinci roku 2006, představitelé úspěšných programů Data Governance prohlásili, že podle jejich zkušenosti, Data Governance je mezi 80% a 95% o komunikaci! Všichni vyjádřili údiv nad množstvím písemné a ústní komunikace, která je nezbytná k úspěšnému vytvoření sdružení stakeholderů, aby dosáhli svých cílů.

Není komunikace vaše silná stránka? To je v pořádku. Plány komunikace a podpůrné prostředky lze vyvinout osobami s odbornými znalostmi a poté je administrovat zaměstnanci vašeho programu.

Například většinu pracovníků DGO zaměstnává segmentace zákazníků, aby porozuměli informačním potřebám stakeholderů, preferované terminologii, a speciálním zájmům. Rozvíjejí komunikaci v různých oblastech

- elevator speech (krátká, rychlá a jednoduchá definice služby, produktu, ...),
- value statements (prohlášení o hodnotě),
- impact statements (prohlášení o dopadech),
- prezentace a jiná dokumentace
- šablony pro e-mail a šablony pro prezentace a zprávy

takže jsou schopni doručit správné verze governance zpráv správným lidem se správnou úrovní detailu. Vytvářejí matici participace stakeholderů, aby se ujistili, že stakeholdeři nejsou přehlíženi, a že správní lidé dostávají zprávy ve správném pořadí.

6. Další zdroje informací

- International Association for Information and Data Quality (<http://www.iaidq.org/>)
- Data Governance Institute (<http://www.datagovernance.com/>)
- Kvalita dat a INSPIRE
http://inspire.jrc.ec.europa.eu/documents/INSPIRE_/JRC83209_Online_Data_quality_in_INSPIRE.pdf

Připomínky a dotazy k obsahu dokumentu pošílej, prosím, na adresu:

Rudolf Richter, richter@fi.muni.cz