



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA PODNIKATELSKÁ

FACULTY OF BUSINESS AND MANAGEMENT

ÚSTAV INFORMATIKY

INSTITUTE OF INFORMATICS

ZAVEDENÍ SYSTÉMU ŘÍZENÍ DAT

DATA GOVERNANCE SYSTEM IMPLEMENTATION

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Denis Szentkirályi

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Petr Sedlák

BRNO 2024

Zadání diplomové práce

Ústav: Ústav informatiky
Student: **Bc. Denis Szentkirályi**
Vedoucí práce: **Ing. Petr Sedlák**
Akademický rok: 2023/24
Studijní program: Informační management

Garant studijního programu Vám v souladu se zákonem č. 111/1998 Sb., o vysokých školách ve znění pozdějších předpisů a se Studijním a zkušebním řádem VUT v Brně zadává diplomovou práci s názvem:

Zavedení systému řízení dat

Charakteristika problematiky úkolu:

Úvod
Teoretická východiska práce
Analýza současného stavu
Vlastní návrhy řešení
Závěr

Cíle, kterých má být dosaženo:

Cílem této magisterské práce je zavést systém řízení dat s využitím poznatků ze zavedených postupů řízení dat. Studie se bude zabývat současnými problémy, navrhne klíčové komponenty a posoudí celkový ekonomický přínos lepšího řízení dat.

Základní literární prameny:

ČSN EN ISO/IEC 27001 Informační bezpečnost, kybernetická bezpečnost a ochrana soukromí - Systémy managementu informační bezpečnosti - Požadavky, 2023. [Praha]: Česká agentura pro standardizaci.

ČSN EN ISO/IEC 27002 Informační bezpečnost, kybernetická bezpečnost a ochrana soukromí - Opatření informační bezpečnosti, 2023. [Praha]: Česká agentura pro standardizaci.

DOUCEK, Petr; KONEČNÝ, Martin a NOVÁK, Luděk, 2019. Řízení kybernetické bezpečnosti a bezpečnosti informací. Praha: Professional Publishing. ISBN isbn978-80-88260-39-4.

SEDLÁK, Petr a KONEČNÝ, Martin, 2023. Přeměna ISMS v manažerské informatice. Brno: CERM, akademické nakladatelství. ISBN isbn978-80-7623-110-8.

SEDLÁK, Petr a KONEČNÝ, Martin, 2021. Kybernetická (ne)bezpečnost: problematika bezpečnosti v kyberprostoru. Brno: CERM, akademické nakladatelství. ISBN isbn978-80-7623-068-2.

Termín odevzdání diplomové práce je stanoven časovým plánem akademického roku 2023/24

V Brně dne 4.2.2024

L. S.

doc. Ing. Miloš Koch, CSc.
garant

doc. Ing. Vojtěch Bartoš, Ph.D.
děkan

Abstrakt

Diplomová práca sa venuje problematike správy dát. Cieľom práce je zhodnotiť súčasný stav správy dát vo vybranej spoločnosti a využitím poznatkov zo zavedených postupov následne zaviesť systém správy dát, ktorý povedie k zlepšeniu súčasnej situácie. Súčasný stav spoločnosti je analyzovaný pomocou analýzy medzier a pomocou Maturity Assessmentu a na základe ich výsledkov je navrhnutá implementácia systému správy dát, ktorá reflektuje špecifické požiadavky spoločnosti a je v súlade s najlepšími postupmi.

Abstract

The diploma thesis focuses on the issue of data governance. The aim of the thesis is to evaluate the current state of data governance in the selected company and to implement a data governance system that will lead to an improvement of the current situation. The current state of the company is analysed using Gap Analysis and Maturity Assessment and based on their results, the implementation of a data governance system is proposed that reflects the specific requirements of the company and is in line with best practices.

Kľúčové slová

dáta, riadenie dát, správa dát, dátová bezpečnosť, dátová kvalita, dátové vlastníctvo

Key words

data, data management, data governance, data security, data quality, data stewardship

Bibliografická citácia

SZENTKIRÁLYI, Denis. *Zavedení systému řízení dat* [online]. Brno, 2024 [cit. 2024-05-05]. Dostupné z: <https://www.vutbr.cz/studenti/zav-prace/detail/158857>. Diplomová práce. Vysoké učení technické v Brně, Fakulta podnikatelská, Ústav informatiky. Vedoucí práce Ing. Petr Sedlák.

Čestné prehlásenie

Prehlasujem, že predložená diplomová práca je pôvodná a spracoval som ju samostatne.
Prehlasujem, že citácia použitých prameňov je úplná, že som vo svojej práci neporušil autorské práva (v zmysle Zákona č. 121/2000 Sb., o právu autorskom a o právach súvisiacich s právom autorským).

V Brne dňa 5. 5. 2024

Bc. Denis Szentkirályi

autor

Pod'akovanie

Týmto by som sa rád poďakoval vedúcemu mojej diplomovej práce a školiteľovi Ing. Petrovi Sedlákovovi za odbornú pomoc a užitočné rady. Zároveň by som sa rád poďakoval tímu Data Governance pod vedením Ing. Martina Šťastného za konzultácie a praktické rady z praxe.

Pod'akovanie venujem tiež mojej rodine, priateľke, kamarátom a domácim miláčikom, ktorí mi s láskou boli veľkou oporou počas celého môjho štúdia.

OBSAH

ÚVOD A CIEĽ PRÁCE	12
1 TEORETICKÉ VÝCHODISKÁ PRÁCE	13
1.1. Základné pojmy	13
1.1.1. Dáta	13
1.1.2. Dáta a informácie	14
1.1.3. Dáta ako aktívum	15
1.1.4. Dáta a hodnota	15
1.1.5. Dáta a kvalita	16
1.1.6. Dáta a riziko	16
1.2. Data Management	17
1.3. DAMA Framework	18
1.3.1. Znalostné oblasti	19
1.4. Data Governance	21
1.4.1. Data Management a Data Governance	22
1.4.2. Data Stewardship	22
1.5. Metadata Management	24
1.6. Data Security	25
1.6.1. Big Data Security	26
1.7. Nástroje	26
1.7.1. Business Glossary	27
1.7.2. Data Catalog	28
1.7.3. Data Discovery	29
1.7.4. Data Lineage	30
1.7.5. CMS	31
1.8. Maturity Assessment	31

1.8.1.	Maturity Assessment založený na DMBOK2.....	32
1.8.2.	Maturity Assessment založený na PPTG.....	33
1.9.	Lewin.....	34
1.10.	PERT	35
1.11.	Analýza rizík	35
1.11.1.	Identifikácia rizík	36
1.11.2.	Vyhodnotenie rizík.....	36
1.11.3.	Mapa rizík	36
1.11.4.	Protiopatrenia a metódy znižovania rizika.....	36
1.12.	Návratnosť investície.....	38
1.13.	Zhrnutie	39
2	ANALÝZA SÚČASNÉHO STAVU.....	40
2.1.	Základné informácie o spoločnosti	40
2.2.	Analýza medzier.....	41
2.2.1.	Identifikované medzery v rámci oddelení	42
2.3.	Maturity Assessment.....	45
2.3.1.	Výsledky DMBOK maturity assessmentu.....	46
2.3.2.	Výsledok podľa PPTG.....	47
2.4.	Vyhodnotenie súčasného stavu	48
2.4.1.	Data Governance Framework	50
2.4.2.	Metadáta.....	52
2.4.3.	Data Quality	54
2.5.	Zhrnutie	57
3	VLASTNÉ NÁVRHY RIEŠENIA.....	58
3.1.	Popis navrhovanej zmeny	58
3.2.	Lewinov model.....	59

3.2.1.	Fáza rozmrazenia	59
3.2.2.	Fáza prechodu a aplikácie zmeny	61
3.2.3.	Fáza zmrazenia	62
3.2.4.	Verifikácia dosiahnutých výsledkov.....	62
3.3.	Vypracovanie projektu	63
3.3.1.	Vymenovanie sponzora DG programu	63
3.3.2.	Menovanie vedúceho tímu DG	63
3.3.3.	Pridelenie rozpočtu na DG.....	63
3.3.4.	Definícia dátovej stratégie	63
3.3.5.	Prijatie špecialistov na Data Governance	63
3.3.6.	Zriadenie výboru pre Data Governance (DGSC)	64
3.3.7.	Definovanie základnej politiky	64
3.3.8.	Vymedzenie politiky vlastníctva dát.....	65
3.3.9.	Definovanie vlastníctva na úrovni domén	65
3.3.10.	Identifikácia DO, DS a TDS	65
3.3.11.	Vytvorenie politiky a postupov v oblasti metadát.....	65
3.3.12.	Zabezpečenie platformy pre business glossary	66
3.3.13.	Zabezpečenie dátového katalógu	66
3.3.14.	Zabezpečenie platformy pre data lineage.....	66
3.3.15.	Zabezpečenie Confluence ako centrálného CMS	67
3.3.16.	Zabezpečenie report katalógu	67
3.3.17.	Počiatkové zhromažďovanie metadát.....	67
3.3.18.	Zhromažďovanie pojmov do business glossary	69
3.3.19.	Zhromažďovanie reportov do report katalógu	70
3.4.	Definovanie dátových rolí	71
3.4.1.	Vlastník dátovej domény	71

3.4.2.	Data Owners (DO)	72
3.4.3.	Business Data Stewards (DS)	73
3.4.4.	Technical Data Stewards (TDS)	74
3.4.5.	Model dátového vlastníctva	75
3.4.6.	Časová náročnosť pre jednotlivé role	76
3.4.7.	Organizačné zavedenie	77
3.5.	Časová analýza	79
3.6.	Analýza rizík	82
3.6.1.	Identifikácia a ohodnotenie rizík	82
3.6.2.	Mapa rizík	84
3.6.3.	Návrh opatrení	85
3.6.4.	Riziká po opatreniach	86
3.7.	Ekonomické zhodnotenie	87
3.7.1.	Náklady	88
3.7.2.	Benefity	90
3.7.3.	Zhodnotenie z hľadiska bezpečnosti	92
3.8.	Zhrnutie	95
	ZÁVER	96
	ZOZNAM POUŽITÝCH SKRATIEK	97
	REGISTER OBECNE POUŽÍVANÝCH POJMOV	98
	ZOZNAM POUŽITÝCH ZDROJOV	103
	ZOZNAM OBRÁZKOV	106
	ZOZNAM TABULIEK	107
	ZOZNAM GRAFOV	108
	ZOZNAM PRÍLOH	109

ÚVOD A CIEĽ PRÁCE

Pre svoju diplomovú prácu som si zvolil tému zavedenia systému riadenia dát v rámci spoločnosti, v ktorej už dva roky pracujem. Počas týchto dvoch rokov som pracoval v rôznych dátovo orientovaných tímoch a všimol som si, že niektoré oddelenia majú podstatne väčší prístup k dátam a pracujú s nimi efektívnejšie, než iné. Zároveň bolo možné pozorovať negatívne dopady nejednotného prístupu k dátam v podobe nejednotného názvoslovía, reportingu či samotnej kvality dát.

Tieto skúsenosti vyvolali otázky: Ako by bolo možné zlepšiť správu dát? Ako k problematike pristupujú iné organizácie? Existuje univerzálny rámec, ktorý by mohol byť prínosom aj pre našu organizáciu?

Cieľom tejto diplomovej práce je zhodnotiť súčasný stav správy dát vo vybranej spoločnosti a využitím poznatkov zo zavedených postupov následne zaviesť systém správy dát, ktorý povedie k zlepšeniu súčasnej situácie.

Aby bolo možné cieľ dosiahnuť, boli stanovené štyri čiastkové ciele. Prvým čiastkovým cieľom je rešeršou odbornej literatúry vypracovať teoretické a odborné zázemie a vymedziť pojmy pre analytickú a návrhovú časť.

Druhým čiastkovým cieľom je vhodným spôsobom analyzovať súčasný stav správy dát. Na základe osobných skúseností je predpoklad, že v rámci oddelení budú prostredníctvom rozhovorov so zamestnancami a analýzou relevantnej dokumentácie zistené rozličné nedostatky, ktoré bude potrebné kategorizovať a prioritizovať.

Tretím čiastkovým cieľom je využiť poznatky získané z teoretickej a analytickej časti pre zavedenie systému správy dát, ktorý bude reflektovať špecifické požiadavky spoločnosti a zároveň bude v súlade s uznávanými rámcami a najlepšimi postupmi. V priamej nadväznosti je vytýčený štvrtý čiastkový cieľ, ktorým je ekonomické zhodnotenie navrhnutého systému, zmien a opatrení.

Diplomová práca predstavuje spojenie akademických poznatkov a praktických skúseností získaných v korporátnom prostredí. Za možnosť jej realizácie som rád, pretože má potenciál nielen zlepšiť operácie v spoločnosti, ale aj poskytnúť cenný príspevok do teórie správy dát.

1 TEORETICKÉ VÝCHODISKÁ PRÁCE

Táto časť diplomovej práce predstavuje rešerš pojmov, ktoré tvoria teoretický základ pre tému tejto práce. Jej cieľom je vytvoriť komplexný základ pre nasledujúcu analytickú a návrhovú časť.

V úvode kapitoly budú vymedzené základné pojmy a názvoslovie, ktoré budú použité vo zvyšku práce. Ďalej bude súčasný stav riešenej problematiky popísaný pomocou definícií kľúčových pojmov Data Management, Data Governance, ich porovnaníu a zároveň predstaveníu rámca DAMA, ktorý je kľúčový pre analytickú časť. Samostatné kapitoly budú venované aj ďalším dátovým oblastiam ako Metadata Management a Data Security. Po dátových oblastiach bude nasledovať kapitola popisujúca nástroje, ktoré podporujú fungovanie týchto oblastí. Pre zhodnotenie výsledkov analytickej časti bude popísaný teoretický základ pre Maturity Assessment.

Zvyšné kapitoly budú slúžiť najmä pre oboznámenie sa s postupmi, ktoré budú použité v návrhovej časti. Ide o manažment zmien pomocou Lewinovho modelu, časové plánovanie projektu pomocou PERT, analýzu rizík a postup pre ekonomické zhodnotenie investície pomocou ROI a ROSI.

1.1. Základné pojmy

Táto kapitola je zameraná na vysvetlenie kľúčových pojmov, ktoré sú základom pre pochopenie správy dát v kontexte informačných technológií. Kapitola sa postupne zaoberá definíciami a charakteristikami dát, ich premenou na informácie, až po širšie pochopenie dát ako cenného podnikového aktíva. Tiež je priblížené, ako kvalita, riziko a hodnota dát ovplyvňujú ich využitie a správu. Tieto základné pojmy slúžia ako fundament pre ďalšie kapitoly diplomovej práce.

1.1.1. Dáta

Dlhodobo zaužívané definície dát zdôrazňujú ich úlohu pri reprezentácii faktov o svete. V súvislosti s informačnými technológiami sa pod dátami rozumejú aj informácie, ktoré sú uložené v digitálnej forme (DAMA, 2017, str. 18).

„Väčšina ľudí usúdi, že keďže dáta reprezentujú fakty, sú formou pravdy o svete a že fakty do seba zapadajú. Ale "fakty" nie sú vždy jednoduché alebo priamočiare. Dáta sú prostriedkom reprezentácie. Zastupujú iné veci než seba samé“ (Chisholm, 2010). Dáta sú interpretáciou objektov, ktoré reprezentujú, a zároveň sú objektom, ktorý musí byť interpretovaný (Sebastian-Coleman, 2013). Je to ďalší spôsob, ako povedať, že na to, aby dáta dávali zmysel, potrebujeme kontext. Kontext možno považovať za reprezentačný systém dát; takýto systém zahŕňa spoločný slovník a súbor vzťahov medzi komponentmi. Ak poznáme konvencie takéhoto systému, môžeme údaje v ňom interpretovať. Tieto konvencie sú často zdokumentované v špecifickom druhu dát, ktorý sa označuje ako metadáta (DAMA, 2017, str. 19).

1.1.2. Dáta a informácie

Dáta sú často označované ako "surové dáta" a informácie ako "dáta v kontexte". Často sa na opísanie vzťahu medzi dátami (v základni), informáciami, znalosťami a múdrosťou (na samom vrchole) používa vrstvená pyramída. Toto znázornenie predstavuje pre správu dát niekoľko výziev (DAMA, 2017, str. 20).

- Vychádza z predpokladu, že dáta jednoducho existujú. Ale dáta jednoducho neexistujú. Musia sa vytvoriť.
- Tým, že opisuje lineárnu postupnosť od dát po múdrosť, neberie na vedomie, že na vytvorenie dát sú v prvom rade potrebné znalosti.
- Naznačuje, že dáta a informácie sú oddelené veci, hoci v skutočnosti sú tieto dva pojmy navzájom prepojené a závislé jeden od druhého. Dáta sú formou informácií a informácie sú formou dát (DAMA, 2017, str. 20).

Ako príklad je možné uviesť Report o tržbách za posledný štvrťrok [informácia]. Je založená na dátach z dátového skladu [dáta]. V nasledujúcom štvrťroku sa tieto výsledky [dáta] použijú na vytvorenie našich štvrťročných výkonnostných ukazovateľov [informácie]. Uvedomenie si, že dáta a informácie je potrebné pripraviť na rôzne účely, potvrdzuje hlavnú zásadu správy dát - dáta aj informácie je potrebné riadiť (DAMA, 2017, str. 20). Z tohto dôvodu sa budú v tejto diplomovej práci tieto pojmy používať zameniteľne.

1.1.3. Dáta ako aktívum

Aktívum je všetok hmotný a nehmotný majetok (Sedlák, 2023, str. 15). Je to ekonomický zdroj, ktorý možno vlastníť alebo kontrolovať a ktorý má alebo vytvára hodnotu. Aktíva sa dajú premeniť na peniaze. Dáta sú všeobecne uznávané ako podnikové aktívum, hoci chápanie toho, čo znamená spravovať dáta ako aktívum, sa stále vyvíja. V súčasnosti sa "hodnota goodwillu" bežne uvádza ako položka vo výkaze ziskov a strát. Podobne, hoci nie je všeobecne prijatá, je čoraz bežnejšia aj monetizácia dát (DAMA, 2017, str. 20).

1.1.4. Dáta a hodnota

Hodnota je rozdiel medzi cenou veci a úžitkom, ktorý z nej plynie. Pri niektorých aktívach, ako sú napríklad akcie, je výpočet hodnoty jednoduchý. V prípade dát sú však tieto výpočty zložitejšie, pretože náklady ani úžitky z dát nie sú štandardizované. Keďže údaje každej organizácie sú pre ňu jedinečné, prístup k oceňovaniu údajov sa musí začať formulovaním všeobecných kategórií nákladov a prínosov, ktoré možno v rámci organizácie dôsledne uplatňovať. Medzi vzorové kategórie patria:

- Náklady na získanie a uchovávanie dát.
- Náklady na nahradenie dát v prípade ich straty.
- Dopad na organizáciu, ak by dáta chýbali.
- Náklady na zmiernenie rizík a potenciálne náklady na riziká spojené s dátami.
- Náklady na zlepšenie dát.
- Výhody vyššej kvality dát.
- Koľko by za tieto dáta zaplatila konkurencia.
- Za koľko by bolo možné dáta prediť.
- Očakávané príjmy z inovatívneho využitia dát (DAMA, 2017, str. 25).

Hlavnou výzvou pri oceňovaní dátových aktív je skutočnosť, že hodnota dát je závislá od kontextu (to, čo má hodnotu pre jednu organizáciu, nemusí mať hodnotu pre inú) a často aj od času (to, čo malo hodnotu včera, nemusí mať hodnotu dnes). Napriek tomu je pravdepodobné, že v rámci organizácie budú určité typy dát v priebehu času konzistentne hodnotné. Napríklad dôveryhodné informácie o zákazníkoch. Informácie o zákazníkoch môžu časom dokonca získať na hodnote, pretože sa akumuluje viac dát súvisiacich s aktivitami zákazníkov (DAMA, 2017, str. 25).

1.1.5. Dáta a kvalita

Pojem pre kvalitu dát, Data Quality, sa vzťahuje na charakteristiky spojené s vysokou kvalitou dát, ako aj na procesy používané na meranie alebo zlepšovanie kvality dát. (DAMA, 2017, str. 26, 452) Spoločnosti, ktoré chcú zo svojich dát získať hodnotu, si uvedomujú, že vysokokvalitné dáta sú cennejšie ako dáta nízkej kvality. Dáta s nízkou kvalitou sú rizikové. Môžu poškodiť povest' organizácie, čo môže mať za následok pokuty, stratu príjmov, stratu zákazníkov a negatívnu publicitu v médiách. Regulačné požiadavky často vyžadujú vysokú kvalitu dát. Okrem toho je s nekvalitnými dátami spojených mnoho priamych nákladov (DAMA, 2017, str. 26, 452).

Kvalitné dáta nielen zmierňujú riziko a znižujú náklady, ale aj zvyšujú efektívnosť. Zamestnanci dokážu rýchlejšie a dôslednejšie odpovedať na otázky, ak pracujú so spoľahlivými dátami. Trávia menej času zisťovaním, či sú dáta správne, a viac času na to, aby ich využívali na získavanie informácií, prijímanie rozhodnutí a poskytovanie služieb zákazníkom (DAMA, 2017, str. 26, 452).

1.1.6. Dáta a riziko

Dáta nepredstavujú len hodnotu, ale aj riziko. Dáta nízkej kvality (nepresné, neúplné alebo neaktuálne) samozrejme predstavujú riziko, pretože ich obsah nie je správny, ako bolo spomenuté vyššie. Dáta však predstavujú riziko aj preto, že môžu byť nesprávne pochopené a zneužitú (DAMA, 2017, str. 30). Organizácie získajú najväčšiu hodnotu z dát najvyššej kvality - dostupných, relevantných, úplných, presných, konzistentných, včasných, použiteľných, zmysluplných a pochopených. Napriek tomu pri mnohých dôležitých rozhodnutiach máme informačné medzery - rozdiel medzi tým, čo vieme, a tým, čo potrebujeme vedieť na prijatie efektívneho rozhodnutia. Informačné medzery predstavujú záväzky podniku s potenciálne hlbokým vplyvom na prevádzkovú efektívnosť a ziskovosť. Organizácie, ktoré si uvedomujú hodnotu vysokokvalitných dát, môžu podniknúť konkrétne, proaktívne kroky na zlepšenie kvality a použiteľnosti dát a informácií v rámci regulačných a etických kultúrnych rámcov. Podobne, keďže spotrebitelia si čoraz viac uvedomujú, ako sa ich dáta používajú, očakávajú nielen plynulejšie a efektívnejšie fungovanie procesov, ale aj ochranu svojich údajov a ochranu súkromia (DAMA, 2017, str. 30).

1.2. Data Management

Mnohé organizácie si uvedomujú, že ich dáta sú dôležitým podnikovým aktívom. Dáta a informácie im poskytujú prehľad o ich zákazníkoch, produktoch a službách. Takisto im pomáhajú pri inováciách a dosahovaní strategických cieľov. Napriek tomuto poznaniu len málo organizácií aktívne spravuje údaje ako aktívum, z ktorého môžu získavať trvalú hodnotu (Evans a Price, 2012).

Získavanie pridanej hodnoty z dát sa nedeje vo vzduchoprázdne alebo náhodou. Vyžaduje si zámer, plánovanie, koordináciu a rozhodnutie. Vyžaduje si správu a vedenie (DAMA, 2017, str. 17).

Data Management je vývoj, realizácia a dohľad nad plánmi, politikami, programami a postupmi, ktoré poskytujú, kontrolujú, chránia a zvyšujú pridanú hodnotu dát a informačných aktív počas celého ich životného cyklu (DAMA, 2017, str. 17). V určitom bode všetky organizácie zavedú manažment dát. Ak sa zhromažďujú a ukladajú dáta, technicky ide o manažment týchto dát. Pri manažmente dát je dôležitý stupeň vyspelosti, ktorý sa uplatňuje pri manažovaní hodnoty a kvality dátových setov (Reichental, 2022, str. 53).

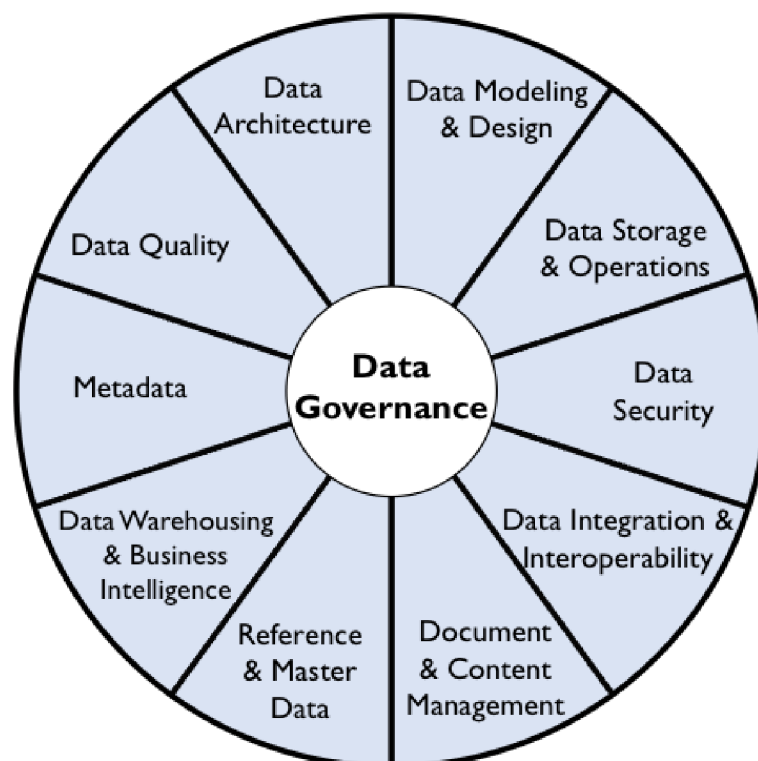
Činnosti Data Managementu majú široký záber. Zahŕňajú všetko od schopnosti prijímať konzistentné rozhodnutia o tom, ako z dát získať strategickú hodnotu, až po technické nasadenie a výkonnosť databáz. Preto si Data Management vyžaduje technické aj netechnické (t. j. "biznis") zručnosti. Zodpovednosť za správu dát sa musí rozdeliť medzi obchodné a IT role a ľudia z oboch oblastí musia byť schopní spolupracovať, aby zabezpečili, že organizácia bude mať vysoko kvalitné dáta, ktoré budú spĺňať jej strategické potreby (DAMA, 2017, str. 17). Kvalitný manažment dát poskytuje príležitosť na výrazné zvýšenie výkonnosti organizácie (Reichental, 2022, str. 53).

S cieľom podporiť odborníkov vykonávajúcich prácu v oblasti správy dát pripravila The Data Management Association (DAMA) International knihu DAMA DMBOK (The DAMA Guide to the Data Management Body of Knowledge - DAMA DMBOK2). V nasledujúcej kapitole je priblížený rámec DAMA Data Management Framework, ktorý poskytuje kontext pre prácu vykonávanú odborníkmi na správu dát v rámci rôznych Data Management znalostných oblastí (DAMA, 2017, str. 17).

1.3. DAMA Framework

DAMA International's Guide to the Data Management Body of Knowledge (DAMA DMBOK Guide) je súbor procesov a oblastí znalostí, ktoré sú všeobecne považované za najlepšie postupy v oblasti správy dát. Správa dát je zastrešujúci pojem, ktorý opisuje procesy používané na plánovanie, špecifikáciu, umožnenie, vytváranie, získavanie, údržbu, používanie, archiváciu, vyhľadávanie, kontrolu a čistenie dát. Tieto procesy sa navzájom prekrývajú a ovplyvňujú oblasti znalostí správy dát (Cupoli et al., 2014). Prvé vydanie vzniklo v roku 2009.

DAMA Wheel, uvedený na obrázku č. 1, definuje znalostné oblasti riadenia dát. Do centra činností Data Managementu kladie Data Governance, pretože správa je potrebná na zabezpečenie konzistentnosti v rámci jednotlivých funkcií a rovnováhy medzi nimi. Ostatné oblasti znalostí (data security, data quality atď.) sú vyvážené okolo kolesa. Všetky sú nevyhnutnou súčasťou vyspelej správy dát, ale môžu byť implementované v rôznom čase v závislosti od požiadaviek organizácie (DAMA, 2017, str. 36).



Obrázok č. 1: DAMA Wheel
(Zdroj: DAMA, 2017, str. 36)

Ciele publikácie sú zamerané na štandardizáciu a rozvoj prístupov k správe dát prostredníctvom jednotných definícií, osvedčených postupov a zásad. Zahrňujú vytvorenie konsenzuálneho pohľadu na oblasti znalostí, stanovenie hlavných zásad, vyjasnenie rozsahu a hraníc správy dát, identifikáciu organizačných problémov a poskytnutie zdrojov pre ďalšie vzdelávanie v tejto oblasti (Cupoli et al., 2014, str. 7).

Framework je vhodný pre rôzne skupiny, ktoré môže zaujímať správa dát, vrátane certifikovaných odborníkov, IT profesionálov, manažérov či vedúcich pracovníkov (Cupoli et al., 2014, str. 7, 8). Zároveň poskytuje rôzne možnosti využitia, ktoré zahŕňajú vzdelávanie širokého spektra odborníkov o správe dát, budovanie konsenzu v tejto oblasti, usmerňovanie správcov a vlastníkov dát v ich rolách, hodnotenie efektívnosti správy dát, podporu implementácie a zlepšenia v správe dát, vývoj vzdelávacích programov a podporu pri príprave na profesijné certifikácie v oblasti správy dát. (Cupoli et al., 2014, str. 8).

1.3.1. Znalostné oblasti

Každá oblasť znalostí v rámci DAMA frameworku je definovaná kontextovým diagramom, ktorý určuje a vymedzuje rozsah danej oblasti. Tento diagram je špecificky prispôsobený na popis procesov vrátane ich vstupov, výstupov, podnikových podnetov, použitých nástrojov a technického vybavenia. Každý diagram obsahuje:

- **Definície:** Stručný popis oblasti znalostí.
- **Ciele:** Požadované výsledky oblasti znalostí v rámci tohto tématu.
- **Proces:** Zoznam samostatných činností ktoré majú byť vykonané.
- **Vstupy:** Aké vstupy sú priamo nevyhnutné pre začiatok/pokračovanie procesu?
- **Role dodávateľa:** Role a/nebo tímy, ktoré dodávajú vstupy do procesu.
- **Zodpovedné role:** Role a/nebo tímy, ktoré proces vykonávajú.
- **Role zúčastnených strán:** Role alebo tímy, ktoré sú informované/konzultované.
- **Nástroje:** Typy technológií, ktoré proces využíva k vykonávaniu funkcie.
- **Výstupy:** Čo je priamo vytvorené procesmi?
- **Spotrebiteľské role:** Role a/nebo tímy, ktoré očakávajú a prijímajú výstupy.
- **Metriky:** Merania, ktoré kvantifikujú úspešnosť procesov na základe cieľov (Cupoli et al., 2014, str. 8-10).

Oblasť je celkom 11 pričom Data Governance je centrálna, ako bolo vyššie spomenuté. Popis jednotlivých oblastí, ktoré je možné vidieť na obrázku č. 1:

- **Data Governance** – plánovanie, dohľad a kontrola správy dát a využívania dát a zdrojov súvisiacich s dátami. Hoci chápeme, že správa zahŕňa "procesy", nie "veci", bežný termín pre správu dát je Data Governance (nie dáta manažment), a preto bude v práci používaný tento termín.
- **Data Architecture** – celková štruktúra dát a zdrojov súvisiacich s dátami ako neoddeliteľná súčasť podnikovej architektúry.
- **Data Modeling & Design** – analýza, návrh, konštrukcia, testovanie a údržba.
- **Data Storage & Operations** – nasadenie a správa štruktúrovaných fyzických dátových úložísk.
- **Data Security** – zabezpečenie súkromia, dôveryhodnosti a vhodného prístupu.
- **Data Integration & Interoperability** – akvizícia, extrakcia, transformácia, presun, doručenie, replikácia, federácia, virtualizácia a prevádzková podpora.
- **Documents & Content** – ukladanie, ochrana, indexácia a umožnenie prístupu k dátam z neštruktúrovaných zdrojov (elektronické súbory a fyzické záznamy), sprístupnenie týchto dát pre integráciu a interoperabilitu so štruktúrovanými (databázovými) dátami.
- **Reference & Master Data** – správa zdieľaných dát s cieľom znížiť nadbytočnosť a zabezpečiť lepšiu kvalitu dát prostredníctvom štandardizovanej definície a používania dátových hodnôt.
- **Data Warehousing & Business Intelligence** – riadenie analytického spracovania dát a umožnenie prístupu k dátam na podporu rozhodovania pre účely výkazníctva a analýzy.
- **Metadata Management** – zhromažďovanie, kategorizácia, údržba, integrácia, kontrola, správa a poskytovanie metadát.
- **Data Quality** – definovanie, monitorovanie, udržiavanie integrity dát a zlepšovanie ich kvality (Cupoli et al., 2014, str. 10).

Väčšina podnikov nevykonáva všetky činnosti opísané v jednotlivých oblastiach znalostí. Spôsob, akým konkrétna organizácia spravuje svoje dáta, závisí od jej cieľov, veľkosti, zdrojov a zložitosti, ako aj od jej vnímania toho, ako dáta podporujú jej celkovú stratégiu (DAMA, 2017, str. 46).

1.4. Data Governance

Pre Plotkina (2021, str. 1) predstavuje Data Governance vykonávanie rozhodovacích procesov v otázkach súvisiacich s dátami. Plotkin zároveň žartuje, že keď položíte miestnosti plnej špecialistov otázku, čo Data Governance znamená, dostanete toľko definícií, koľko tam je ľudí. DAMA definuje Data Governance ako vykonávanie právomocí, riadenia a spoločného rozhodovania (plánovanie, monitorovanie a vynucovanie) v oblasti správy dátových aktív (DAMA, 2017, str. 69). Eryureka a kol. (2021, str. 10) definujú Data Governance tak, že to je predovšetkým funkcia riadenia dát, ktorá zabezpečuje kvalitu, integritu, bezpečnosť a použiteľnosť dát, ktoré organizácia zbiera. Podľa Hopper (2023, str. 22) je Data Governance organizačný framework, ktorý stanovuje stratégiu, ciele a politiky pre spoločné dátové aktíva. Všetci spomenutí autori sa zhodujú, že Data Governance musí byť implementovaná od okamihu, kedy sa dáta zbierajú, až po okamih, kedy dochádza k ich zničeniu.

Oblasť Data Governance usmerňuje všetky ostatné funkcie dáta manažmentu. Účelom Data Governance je zabezpečiť, aby boli dáta riadne spravované v súlade so smernicami a osvedčenými postupmi (Ladley, 2012). Zatiaľ čo hybnou silou dáta manažmentu celkovo je zabezpečiť, aby organizácia získala hodnotu zo svojich dát, Data Governance sa zameriava na to, ako sa prijímajú rozhodnutia o dátach a ako sa majú ľudia a procesy správať vo vzťahu k dátam (DAMA, 2017, str. 67).

Data Governance musí tiež zabezpečiť, že zainteresované strany získavajú vysoko kvalitný integrovaný pohľad na všetky dáta v rámci spoločnosti. Existuje viacero aspektov vysoko kvalitných dát – dáta musia byť správne, aktuálne a konzistentné s ostatnými dátami v spoločnosti (Eryureka a kol., 2021, str. 11).

Ďalej musí byť Data Governance implementovaná aj pre zabezpečenie bezpečnosti dát, čo znamená, že:

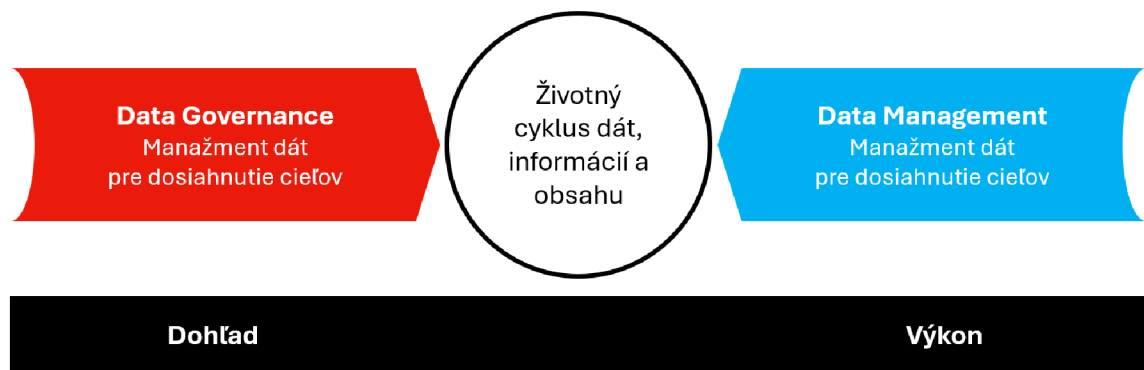
- k dátam majú prístup iba oprávnení používatelia povolenými spôsobmi,
- dáta sú auditovateľné, čo znamená, že všetky prístupy, vrátane zmien, sú zaznamenané,
- sú v súlade s predpismi (Eryureka a kol., 2021, str., 11).

Cieľom Data Governance je zvýšiť dôveru v dáta. Dôveryhodné dáta sú nevyhnutným predpokladom pre umožnenie používateľom využívať podnikové dáta na podporu

rozhodovania, hodnotenia rizík a riadenie pomocou KPI (Eryureka a kol., 2021, str. 11). Vo väčšine organizácií si prijatie formálnej Data Governance vyžaduje podporu vo forme change manažmentu, ako aj sponzorstvo zo strany riadiaceho predstaviteľa na úrovni C, ako je napríklad Chief Financial Officer alebo Chief Data Officer (DAMA, 2017, str. 68).

1.4.1. Data Management a Data Governance

Data Governance sa používa zameniteľne s pojmami Data Quality alebo Data Management vzhľadom na to, že spolu úzko súvisia (Mahanti, 2021, str. 53). Častou analógiou je prirovnanie data governance k auditu a účtovníctvu. Audítori a kontrolóri stanovujú pravidlá správy finančných aktív. Odborníci na data governance stanovujú pravidlá pre správu dátových aktív. Ostatné oblasti tieto pravidlá vykonávajú (DAMA, 2017, str. 73).



Obrázok č. 2: Data Governance a Data Management
(Zdroj: Vlastné spracovanie podľa DAMA, 2017, str. 73)

Tak ako audítor kontroluje finančné procesy, ale v skutočnosti nevykonáva finančné riadenie, data governance zabezpečuje správnu správu dát bez toho, aby priamo vykonával manažment. Data Governance predstavuje prirodzené oddelenie povinností medzi dohľadom a výkonom (DAMA, 2017, str. 73).

1.4.2. Data Stewardship

Data Stewardship je najčastejšie používané spojenie na označenie povinnosti a zodpovednosti za dáta a procesy, ktoré zabezpečujú efektívnu kontrolu a používanie dátových aktív. Správcovstvo môže byť formalizované prostredníctvom pracovných názvov a popisov, alebo môže ísť o menej formálnu funkciu riadenú ľuďmi, ktorí sa snažia pomôcť organizácii získať hodnotu z jej dát. Pojmy ako kustód alebo steward

(správca) sú často synonymami pre tých, ktorí vykonávajú funkcie podobné správcovským (DAMA, 2017, str. 75).

Steward je človek, ktorého prácou je spravovať majetok inej osoby. Data Stewards spravuje dátové aktíva v mene ostatných a v najlepšom záujme organizácie (McGilvray, 2008). Data Steward reprezentuje záujmy všetkých stakeholderov a musí zohľadňovať perspektívu podniku, aby bola zabezpečená vysoká kvalita podnikových dát a ich efektívne využitie. Efektívny Data Stewardi sú zodpovední za data governance aktivity a majú časť svojho pracovného času vyhradenú priamo pre tieto aktivity (DAMA, 2017, str. 76). Zvyčajne sa ich aktivity zameriavajú na:

- **vytváranie a manažovanie metadát** – definovanie biznisovej terminológie, validovanie dátových hodnôt a pod.,
- **dokumentovanie pravidiel a štandardov** – definovanie biznisových pravidiel, dátových štandardov a pravidiel pre dátovú kvalitu,
- **spravovanie dátovej kvality** – identifikácia a riešenie problémov s dátami,
- **spoluúčasť pri data governance aktivitách.**

V závislosti na komplexnosti organizácie a cieľov jej Data Governance programu, formálne zvolený Data Stewardi môžu byť rozlíšený podľa ich role v organizácii, zameraní ich práce, alebo oboch. Ide napríklad o:

- **Vlastníci dátových domén** dohliadajú na dátové domény naprieč biznisom.
- **Data Owner** je Data Steward, ktorý má schvaľovaciu právomoc pre rozhodnutia týkajúce sa dát v doméne, ktorú spravuje.
- **Business** (biznisový) **Data Stewardi** sú špecialisti z obchodných oddelení zodpovední za podskupinu dát. Častokrát sú považovaní za expertov vo svojej oblasti. Spolupracujú so stakeholdermi a spoločne definujú a riadia dáta.
- **Technical** (technický) **Data Stewardi** sú IT profesionáli so špecifickými znalosťami, ktorí pôsobiaci v oblasti integrácie dát, BI, dátovej kvality, alebo ako správcovia databázy (DAMA, 2017, str. 97).

Data Stewardship je prevádzkovým aspektom celkového programu Data Governance. Predstavuje oblasť, kde sa vykonáva praktická každodenná práca na správe podnikových dát. Je možné povedať, že bez Data Stewardship je Data Governance len rámcom dobrých úmyslov, ktoré sa nikdy nerealizujú (Plotkin, 2021, str. 10).

1.5. Metadata Management

Spracovanie dát a elektronická výmena dát silne spolieha na presnosť, kontrolovateľnosť a overiteľnosť dát zaznamenaných v databázach. Nevyhnutným predpokladom správnosti, použitia a interpretácie dát je, že vlastníci dát a ich používatelia majú rovnaké pochopenie ich významu. Pre uľahčenie spoločného pochopenia sú definované číselné charakteristiky alebo atribúty týchto dát. Tieto charakteristiky sú známe ako metadáta. Metadáta sa najčastejšie definujú ako „dáta, ktoré popisujú dáta“ (Hopper, 2023, str. 83). Data Governance využíva Metadata management pre tvorbu politík súvisiacich s definíciou dát, používaním dát, bezpečnosťou dát a data lineage. Avšak veľká časť manuálnej správy metadát môže byť časovo veľmi náročná. Preto by systémy mali generovať metadáta spoločne s dátami a Data Governance by malo využiť tieto metadáta spoločne s automatizáciou metadát, čím potreba sledovať metadáta manuálne zanikne. V tomto ohľade využíva Data Governance Metadata Management na vynútenie správcovskej disciplíny v oblasti zberu, objavovania a riadenia dát, ako aj znižovaní rizika spojeného s nimi (Mahanti, 2021, str. 104). Metadata Management a Data Governance pracujú ruka v ruke pri implementácii správnych kontrolných mechanizmov na podnikových dátach. Metadáta sú typom dát a teda rovnako potrebujú byť spravované, ako aj iné dátové typy (Mahanti, 2021, str. 105).

Nástroje na sledovanie metadát rovnako pomáhajú organizáciám sledovať pohyb citlivých údajov. Identifikácia citlivých údajov s použitím Metadát poskytuje najlepší spôsob, ako zabezpečiť, aby boli údaje riadne chránené (DAMA, 2017, str. 257).

Na druhú stranu tieto nástroje vytvárajú riziko s ich zneužitím a teda, že externí aktéri môžu zistiť interné informácie z metadát spojených s dokumentmi. Avšak keďže najväčší počet prípadov úniku dát je dôsledkom nedostatočnej ochrany citlivých údajov z dôvodu neznalosti ich citlivosti, dokumentácia metadát úplne zatieni akékoľvek hypotetické riziko, ktoré by mohlo nastať, ak by sa metadáta nejakým spôsobom dostali do rúk útočníka. Toto riziko je o to zanedbateľnejšie, že pre skúseného útočníka je triviálne nájsť nechránené citlivé údaje v sieti. Ľudia, ktorí si s najväčšou pravdepodobnosťou neuvedomujú potrebu ochrany citlivých údajov, sú zväčša zamestnanci a manažéri (DAMA, 2017, str. 258).

1.6. Data Security

Hlavnými hnacími silami činností v oblasti informačnej bezpečnosti sú znižovanie rizík a obchodný rast. Zabezpečenie dát v organizácii znižuje riziko a zvyšuje konkurenčnú výhodu. Samotná bezpečnosť je cenným aktívom (DAMA, 2017, str. 220).

Riziká spojené s bezpečnosťou údajov súvisia s dodržiavaním právnych predpisov, fidiárnou zodpovednosťou za podnik a akcionárov, reputáciou a právnou a morálnou zodpovednosťou chrániť súkromné a citlivé informácie zamestnancov, obchodných partnerov a zákazníkov. Za nedodržanie predpisov a zmluvných záväzkov hrozia organizáciám pokuty. Únik údajov môže spôsobiť stratu dobrého mena a dôvery zákazníkov (DAMA, 2017, str. 220).

Rast podniku zahŕňa dosahovanie a udržiavanie prevádzkových obchodných cieľov. Problémy s bezpečnosťou údajov, narušenia a neoprávnené obmedzenia prístupu zamestnancov k dátam môžu priamo ovplyvniť prevádzkový úspech (DAMA, 2017, str. 220).

Informačná bezpečnosť sa začína klasifikáciou údajov organizácie s cieľom určiť, ktoré údaje si vyžadujú ochranu. Celkový proces zahŕňa tieto kroky:

1. Zaistenie súhlasu a podpory vedenia spoločnosti.
2. **Identifikácia** a klasifikácia citlivých dátových **aktív**, ich **ocenenie** a vypracovanie celkovej **analýzy rizík**: V závislosti od odvetvia a organizácie môže existovať málo alebo veľa aktív a celý rad citlivých údajov (vrátane osobných identifikačných, lekárskejších, finančných a ďalších) (Sedlák, 2023, str. 81; DAMA, 2017). Bezpečnostné požiadavky sa môžu líšiť v závislosti od toho, kde sú údaje uložené. Značné množstvo citlivých údajov na jednom mieste predstavuje vysoké riziko vzhľadom na možné škody spôsobené jediným narušením (DAMA, 2017).
3. Určenie **spôsobu**, **akým** je potrebné **chrániť jednotlivé aktíva**: Tento krok je kľúčový a priamo nadväzuje na analýzu rizík. Na základe zistených bezpečnostných potrieb je nutné vybrať vhodné opatrenia (Sedlák, 2023, str. 81; DAMA, 2017).

4. Určenie, **ako tieto informácie spolupracujú s podnikovými procesmi**: Na určenie toho, aký prístup je povolený a za akých podmienok, je potrebná analýza obchodných procesov (DAMA, 2017).
5. **Certifikácia**: Tento krok je z hľadiska ISMS nepovinný. Je postačujúce implementovať bez certifikácie, nakoľko systém je funkčný aj bez nej (Sedlák, 2023, str. 81).

Okrem klasifikácie samotných údajov je potrebné posúdiť externé a interné hrozby (DAMA, 2017). Značná väčšina údajov je stratená alebo odhalená v dôsledku neznalosti zamestnancov, ktorí si neuvedomili, že ide o veľmi citlivé informácie, alebo ktorí obišli bezpečnostné politiky (DAMA, 2023). Používatelia predstavujú jeden z najväčších zdrojov bezpečnostných rizík IS (Sedlák, 2023, str. 384). Podľa štúdie Verizon (2023) sa ľudské faktory, ako sú chyby, zneužitie oprávnení a sociálne inžinierstvo, podieľali na približne 74 % všetkých analyzovaných narušení.

1.6.1. Big Data Security

Podľa Sedláka (2021, str. 160) si minimálne 40% dát v digitálnom svete vyžaduje nejakú úroveň ochrany. Zároveň sa objem dát na svete každé dva roky viac než zdvojnásobuje. Big Data (v preklade Veľké Dáta) je pojem pre obrovské objemy dát, ktorých spracovanie v rozumnom čase tradičnými databázovými nástrojmi je náročné. Ich bezpečnosť má dva rozmery. Prvým je samotná bezpečnosť dát a druhým je bezpečnosť technológií pre prácu s nimi. Základné technológie, ktoré sú odporúčané pre ochranu veľkých dát, sú:

- anonymizovanie,
- šifrovanie,
- bezpečnostné kontroly,
- transparentnosť a prístup k veľkým dátam,
- súhlas, vlastníctvo, kontrola (Sedlák, 2021, str. 162, 168, 173).

1.7. Nástroje

Správa dát je v prvom rade o správaní organizácie. Tento problém sa nedá vyriešiť pomocou technológie. Existujú však nástroje, ktoré celkový proces podporujú. Opis týchto nástroj je predmetom nasledujúcich podkapitol.

1.7.1. Business Glossary

Business Glossary je jedným zo základných nástrojov Data governance. Nachádzajú sa v ňom definície biznisových pojmov, na ktorých sa zhodli kľúčový používatelia, a ich nadväznosti na ďalšie dátové objekty (DAMA, 2017, str. 93)

Za tvorbu obsahu Business Glossary sú zvyčajne zodpovední Data Stewardi. Glosár je potrebný, pretože ľudia používajú rozličné slová na rozličné veci. To kladie dôraz na to mať jasné a presné definície pre dáta. Zároveň si väčšina organizácií časom vytvorí svoj vlastný, interný slovník. Glosár je spôsob, ako tento interný slovník zdieľať v rámci organizácie (DAMA, 2017, str. 90).

Vytvorením a zdokumentovaním bežných dátových definícií sa znižuje nejednoznačnosť a zlepšuje sa komunikácia. Definície musia byť zreteľné, jasne formulované a musia vysvetliť výnimky, synonymá či varianty. Schvaľovatelia terminológie by mali zahŕňať zástupcov hlavných skupín používateľov. Dátoví architekti môžu poskytnúť prvotný draft definícií a typové rozdelenie z nimi spravovaných oblastí (DAMA, 2017, str. 90).

Business Glossary by mal mať nasledujúce ciele:

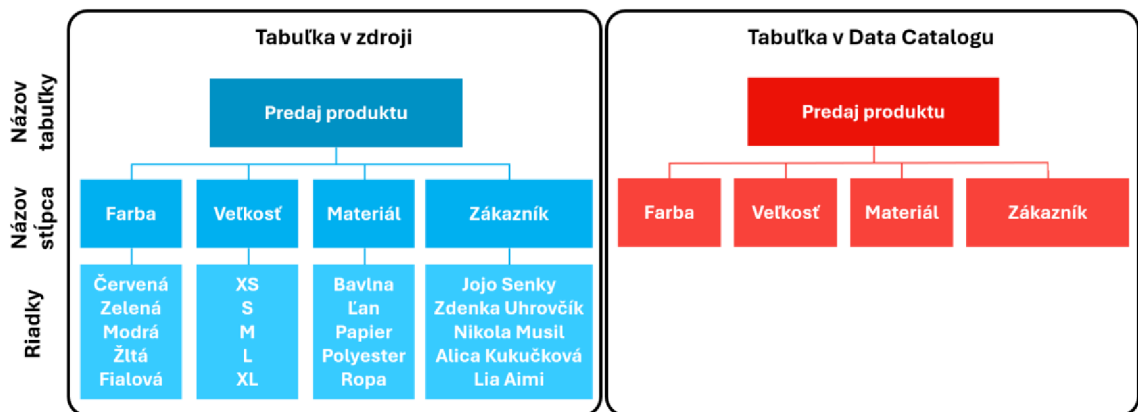
- sprístupniť bežné porozumenie základných podnikových pojmov a terminológii,
- znížiť riziko, že dáta budú nesprávne použité v dôsledku nekonzistentnému porozumeniu podnikových pojmov,
- vylepšiť zladenie medzi technickými aktívami (a ich technickým názvoslovím) a podnikovou organizáciou,
- maximalizovať schopnosť vyhľadávania a umožniť prístup do dokumentovej znalostnej báze (DAMA, 2017, str. 90).

Business Glossary by nemal byť iba zoznam pojmov a ich definícií. Každý pojem by mal byť asociovaný s ďalšími dôležitými metadátami – synonymami, metrikami, biznisovými pravidlami, dátovou líniou, stewardom zodpovedným za pojem atď. (DAMA, 2017, str. 90).

Na začiatok je v tomto bode dobré sa odraziť už z existujúcich dáta artefaktov, vrátane už vybudovaných dátových modelov a databáz. Aj keď môžu byť niektoré dáta neaktuálne, niektoré ich časti môžu byť stále pre tento účel použiteľné. Samozrejme je vhodné si ich presnosť a aktuálnosť overiť u kľúčových užívateľov (DAMA, 2017, str. 155).

1.7.2. Data Catalog

Podstatou dáta katalógu je organizovaný inventár dát v spoločnosti. Poskytuje prehľad len na úrovni metadát, čo znamená, že žiadne skutočné hodnoty dát nie sú vystavené, a teda je možné ukázať každému všetko bez obáv z odhalenia dôverných alebo citlivých údajov (Ole Olesen-Bagneux, 2023, str. 4). Obsahuje údaje o tom, kde sa dáta nachádzajú a aké technické informácie sú s nimi spojené (myslite na: schéma tabuľky, názov tabuľky, názov stĺpca, popis stĺpca), ale tiež umožňuje pripojiť ďalšie "obchodné" metadáta, ako napríklad kto v organizácii vlastní dáta, či sú dáta lokálne vytvorené alebo externe zakúpené, či sa týkajú produkčných prípadov použitia alebo testovania atď. (Eryureka a kol., 2021, str. 70).



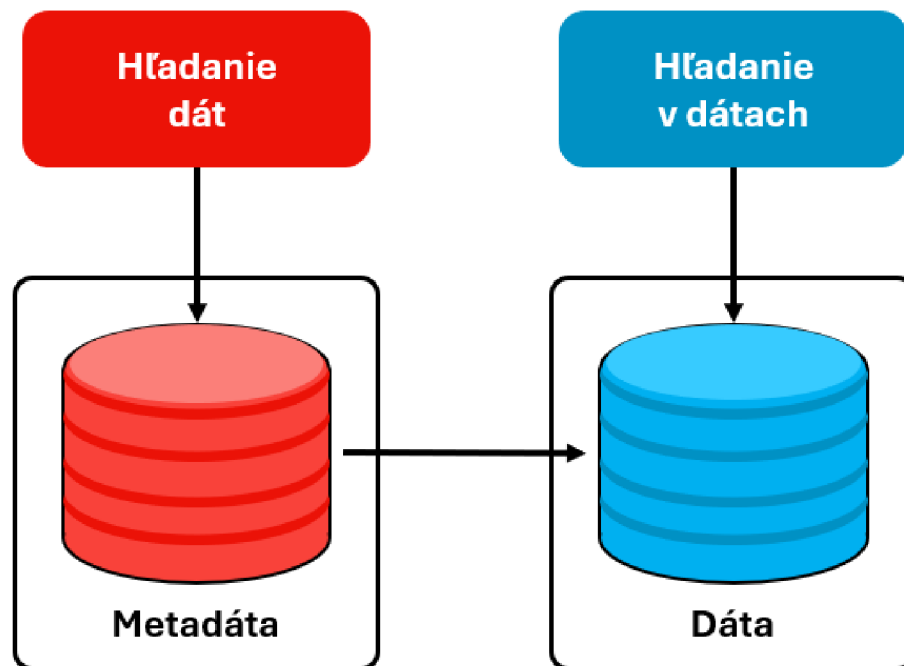
Obrázok č. 3: Tabuľka v zdroji a jej viditeľnosť v data catalogu
(Zdroj: Vlastné spracovanie podľa Olesen-Bagneux, 2023)

Ako je možné vidieť na pravej strane obrázka, žiadne hodnoty nie sú zahrnuté v aktíve v dátovom katalógu. V tomto prípade, citlivé údaje – meno zákazníka – nie sú v dátovom katalógu viditeľné tak, ako tomu je v dátovom zdroji. V katalógu je zobrazený len názov stĺpca, vďaka čomu môže každý vidieť všetko (Olesen-Bagneux, 2023, str. 7).

Dátový katalóg je v podstate databáza s metadátami, ktoré boli „pushed“ alebo „pulled“ z dátových zdrojov v IT prostredí danej spoločnosti. Súčasťou dátového katalógu je aj vyhľadávač, ktorý umožňuje prehľadávať metadáta zozbierané z dátových zdrojov. Olesen-Bagneux (2023, str. 4) zároveň tvrdí, že táto možnosť vyhľadávania je najdôležitejšou vlastnosťou dátového katalógu.

1.7.3. Data Discovery

Dátový katalóg umožňuje všetkým zamestnancom vyhľadávať všetky dáta v ich spoločnosti. Vyhľadávanie a skutočné objavenie dát sa nazýva „data discovery.“ Napriek tomu sa za „data discovery“ zvyčajne nepovažuje hľadanie dát (searching for data), ale len samotné hľadanie v dátach (searching in data), v databázach, za účelom získania nových náhľadov na zákazníkov, produkty a pod., čo nie je správne, nakoľko oba pohľady sú v tomto kontexte relevantné. Zjednodušene povedané, „data discovery“ začína zistením, že určité dáta vôbec existujú, nie toho, čo sa v nich nachádza. Zavedenie dáta katalógu exponenciálne urýchľuje hľadanie v dátach, pretože predchádzajúce objavenie existencie dát je omnoho efektívnejšie s dátovým katalógom, než bez neho (Olesen-Bagneux, 2023, str. 13).



Obrázok č. 4: Rozdiel medzi hľadaním dát a hľadaním v dátach
(Zdroj: Vlastné spracovanie Olesen-Bagneux, 2023)

Data discovery v dátovom katalógu slúži dvom účelom:

- data analytike,
- data governance.

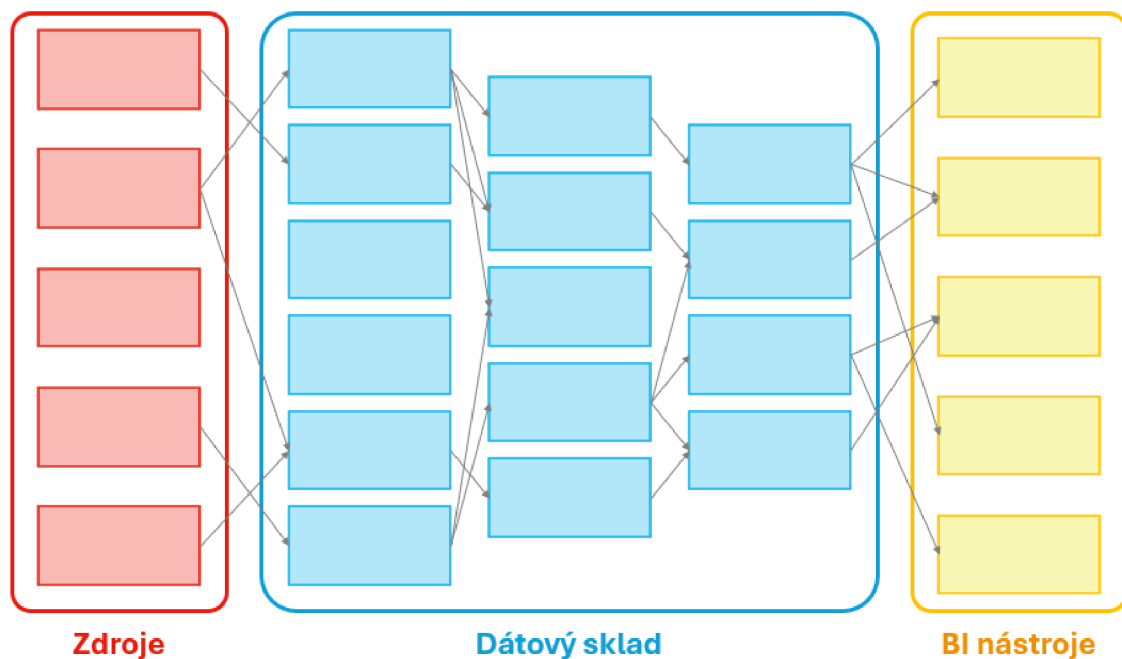
Dátová analytika podporovaná dátovým katalógom je pomerne jednoduchá. Dátový vedci (data scientists) – analytici a podobný zamestnanci – všetci potrebujú dáta. Bez

jednotného, globálneho prehľadu pracujú iba s dátami, ktoré náhodou poznajú – v ich dátových silách – a nie s údajmi, ktoré sa najlepšie hodia pre to, čo chcú robiť. Dátový katalóg pomáha túto prekážku preklenúť vytvorením kompletného prehľadu všetkých dát v spoločnosti (Olesen-Bagneux, 2023, str. 13).

Data governance podporovaná dátovým katalógom má mnoho výhod, pričom najdôležitejšou z nich je možnosť klasifikovať všetky údaje v IT prostredí z hľadiska citlivosti a dôvernosti. Táto funkcia je extrémne nápomocná najmä pre data protection officera (DPO) a chief information security officera (CISO) (Olesen-Bagneux, 2023, str. 14).

1.7.4. Data Lineage

Data Lineage identifikuje, mapuje a popisuje zdroj a cieľ dát, vrátane ich pôvodu a jednotlivých zastávok na ceste. (Reichental, 2023, str. 4) Je to záznam napríklad o tom, kde údaje boli, kam smerujú, ako boli použité, kto ich použil a či boli transformované. Data Lineage pomáha s integritou údajov, znižovaním počtu chýb, dodržiavaním zásad, bezpečnosťou a ďalšími aspektmi (Reichental, 2023, str. 220).



Obrázok č. 5: Príklad E2E data lineage systému od zdrojových dát do BI nástrojov
(Zdroj: Vlastné spracovanie Boisredon, 2023)

1.7.5. CMS

Systém správy obsahu (Content Management System) sa používa na zhromažďovanie, organizovanie, indexovanie a vyhľadávanie obsahu, ktorý sa ukladá buď ako časti, alebo ako celé dokumenty, pričom sa zachovávajú prepojenia medzi jednotlivými prvkami. CMS môže poskytovať aj ovládacie prvky na revíziu obsahu v rámci dokumentov. CMS je v podstate nezávislý od toho, kde a ako sú dokumenty uložené (DAMA, 2017 str. 332).

Dokument je pre content (obsah) to isté, čo vedro pre vodu: kontajner. Content (obsah) sa vzťahuje na dáta a informácie v súbore, dokumente alebo na webovej stránke. Content (obsah) sa často spravuje na základe konceptov, ktoré dokumenty predstavujú, ako aj na základe typu alebo stavu dokumentov. Tiež má svoj životný cyklus (DAMA, 2017, str. 307). Content management (správa obsahu) zahŕňa procesy, techniky a technológie na organizovanie, kategorizáciu a štruktúrovanie informačných zdrojov tak, aby ich bolo možné ukladať, publikovať a opakovane používať viacerými spôsobmi (DAMA, 2017, str. 307).

1.8. Maturity Assessment

Maturity Assessment, v preklade hodnotenie vyspelosti, v oblasti manažmentu dát je systematický prístup, ktorý umožňuje organizáciám hodnotiť úroveň svojich procesov a praxí v riadení dát vzhľadom na preddefinované štandardy a modely zrelosti. Tento proces je aplikovateľný v rôznych rámcoch a metodikách, a zahŕňa nasledujúce kroky:

1. Výber rámca: Organizácia si zvolí model zrelosti, ktorý najlepšie vyhovuje jej potrebám a zameraniu. Tento model poskytuje kritéria pre meranie úrovne zrelosti v rôznych oblastiach manažmentu dát.
2. Diagnostika súčasného stavu: Zhromaždenie a analýza informácií o existujúcich dátových procesoch, ktoré zahŕňa preskúmanie interných politík, rozhovory so zainteresovanými stranami a technickú inventarizáciu.
3. Porovnanie a hodnotenie: Aktuálne praktiky sú porovnané s normami definovanými v zvolenom rámci, čo umožňuje identifikovať oblasti silných stránok a oblasti, ktoré vyžadujú zlepšenie.

4. Plánovanie a implementácia zlepšení: Na základe porovnania a hodnotenia sa vypracuje plán zlepšení zahŕňajúci strategické ciele, iniciatívy na zlepšenie a priradenie zdrojov na dosiahnutie požadovaných zmien (DAMA, 2017).

Tento prístup umožňuje organizáciám nielen identifikovať slabé stránky v ich procesoch riadenia dát, ale tiež poskytuje metodiku pre trvalé zlepšovanie a adaptáciu na meniace sa technologické a obchodné požiadavky (DAMA, 2017).

1.8.1. Maturity Assessment založený na DMBOK2

Model vyspelosti DAMA DMBOK2 je definovaný v zmysle postupu cez úrovne, ktoré opisujú charakteristiky procesov. Keď organizácia pochopí charakteristiky procesov, môže vyhodnotiť svoju úroveň vyspelosti a zaviesť plán na zlepšenie svojich schopností. S každou novou úrovňou sa vykonávanie procesov stáva konzistentnejšie, predvídateľnejšie a spoľahlivejšie. Procesy sa zlepšujú, keď nadobúdajú charakteristiky jednotlivých úrovní. Pokrok sa deje v stanovenom poradí. Žiadnu úroveň nemožno preskočiť. Úrovne DAMA (2017, str. 531) zahŕňajú:

- Úroveň 0: Absencia.
- Úroveň 1: Počiatočná alebo ad hoc - úspech závisí od spôsobilosti jednotlivcov.
- Úroveň 2: Opakovateľná - Je zavedená minimálna procesná disciplína.
- Úroveň 3: Definovaná - Štandardy sú zavedené a používané.
- Úroveň 4: Riadená - Procesy sú kvantifikované a riadené.
- Úroveň 5: Optimalizovaná - Ciele zlepšovania procesov sú kvantifikované.

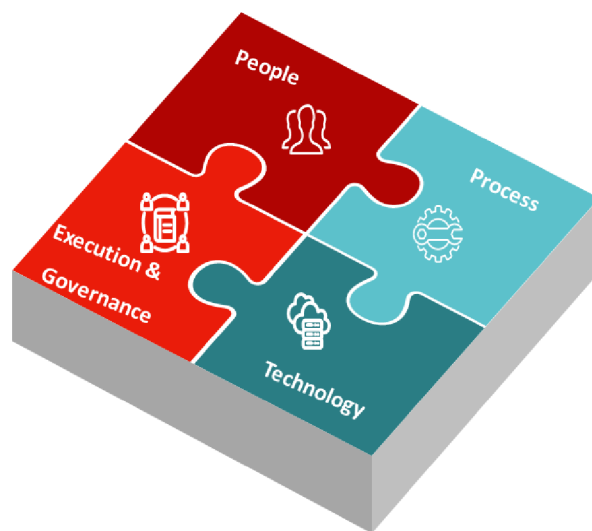


Obrázok č. 6: Úrovne vyspelosti
(Zdroj: Vlastné spracovanie podľa DAMA, 2017)

Vo všeobecnosti je prvou správne fungujúcou úrovňou úroveň 3 - Definovaná. Nižšie úrovne znamenajú nedostatočnú vyspelosť a na zabezpečenie správneho Data Managementu sú potrebné ďalšie zlepšenia Data Governance. Úroveň 5 - Optimalizovaná je takmer nemožné dosiahnuť bez rozsiahlych investícií a intenzívnej podpory vrcholového manažmentu a podnikovej stratégie (DAMA, 2017, str. 536).

1.8.2. Maturity Assessment založený na PPTG

Tradičné rámce riadenia dát, ako je napríklad DAMA, často obsahujú príliš rozsiahle a komplexné štruktúry, ktoré môžu byť ťažko aplikovateľné a nie sú vždy navzájom kompatibilné. V rámci skúmania interných dokumentov bolo zistené, že materská spoločnosť so sídlom v Nemecku rovnako otvorila tému správy a riadenia dát, pričom pre zhodnotenie vyspelosti použili model PPTG. Model vyspelosti PPTG bol vyvinutý interným tímom materskej spoločnosti a sústreďuje iba na štyri hlavné oblasti: ľudí (People), procesy (Process), technológie (Technology) a správu (Governance). Tento prístup umožňuje jasnejšie a efektívnejšie zhodnotenie vyspelosti (Interný dokument, 2023). Z obrázku je možné vidieť, že tento prístup pre zhodnotenie vychádza z PPT rámca pre digitálnu transformáciu a manažmentu zmien, ktorý zas vychádza Leavittovho diamantu.

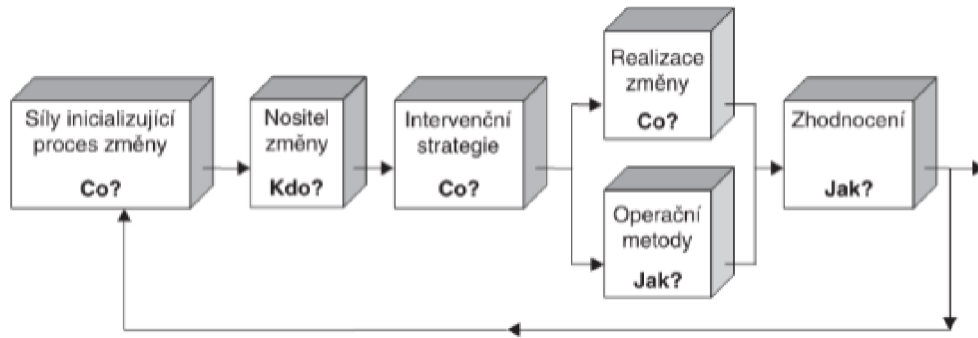


Obrázok č. 7: PPTG model
(Zdroj: Interný dokument, 2023)

Model hodnotí jednotlivé oblasti na stupnici pomocou šiestich úrovní rovnako, ako DMBOK2, no hodnotí štyri oblasti namiesto jedenástich (Interný dokument, 2023).

1.9. Lewin

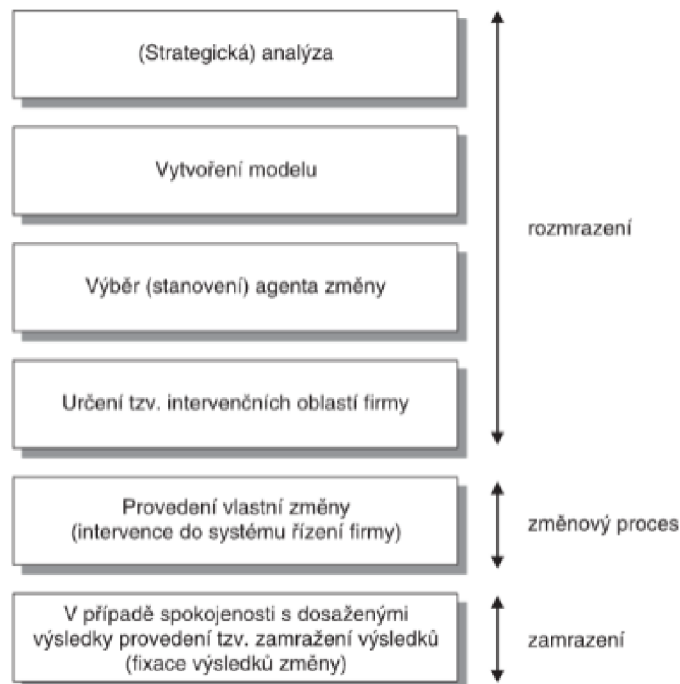
Lewin zdôrazňoval demokratický a prístup k riadenej zmene a jeho model je jedným z najznámejších modelov riadenej zmeny. Lewinov model zachytáva správnu postupnosť krokov, ktoré musia nastať pri riadení zmien (Smejkal, 2013, str. 64).



Obrázok č. 8: Lewinov model riadenej zmeny

(Zdroj: Smejkal, 2013, str. 67)

Model je zložený z troch fáz - fáza rozmrazenia, fáza vlastnej zmeny (zmenový proces) a fáza zmrazenia.



Obrázok č. 9: Základné kroky procesu modelovania riadenej zmeny

(Zdroj: Smejkal, 2013, str. 68)

Prvou fázou je vytrhnúť ľudí zo súčasného stavu, komunikovať s nimi a presvedčiť ich o potrebe zmeny. Na uskutočnenie prvej zmeny musíme vykonať analýzu silového poľa,

určiť tri kľúčové úlohy pre úspešnú zmenu a identifikovať a opísať intervenčné oblasti (Smejkal, 2013, str. 71).

Druhá fáza vlastnej zmeny zahŕňa realizáciu zmeny a definovanie projektu, ktorý pozostáva zo súboru vzájomne súvisiacich činností. Tieto činnosti sa analyzujú pomocou metódy kritickej cesty alebo metódy PERT (Smejkal, 73).

Poslednou **treťou fázou** je fáza zmrazenia, ktorá stabilizuje systém a umožňuje dosiahnuť požadovaný výkon a výsledky. V poslednom kroku je potrebné vyhodnotiť dosiahnuté výsledky (Smejkal, 2013, str. 74).

1.10. PERT

Časová analýza PERT (Project Evaluation and Review Technique) je jednou zo sieťových metód používaných na časové plánovanie projektu, určenie trvania jednotlivých činností a určenie kritickej cesty. Metóda PERT bola vyvinutá na odstránenie nedostatkov Ganttových diagramov (Smejkal, 2013).

Na rozdiel od metódy kritickej cesty (CPM) sa pri metóde PERT nepredpokladá presné trvanie činností, ale pracuje sa s pravdepodobnostnými odhadmi trvania jednotlivých činností. Zároveň pomáha odhaliť riziká. Metóda PERT pracuje s očakávaným trvaním činnosti y_{ij} a štatistickými ukazovateľmi σ_{ij} a σ_{ij}^2 (Smejkal, 2013).

Očakávané trvanie činnosti y_{ij} sa vypočíta podľa vzorca $y_{ij} = \frac{a_{ij} + 4m_{ij} + b_{ij}}{6}$, kde

- a - optimistický čas trvania činnosti,
- m - odhad najpravdepodobnejšej doby trvania činnosti,
- b - pesimistický čas trvania činnosti.

Štatistické ukazovatele:

- Smerodajná odchýlka σ_{ij} podľa $\sigma_{ij} = \frac{b_{ij} - a_{ij}}{6}$.
- Rozptyl ako σ_{ij}^2 .

1.11. Analýza rizík

Analýza rizík je proces definovania hrozieb, pravdepodobnosti ich výskytu a vplyvu na aktíva alebo stanovovanie rizík a ich závažnosti. **Riziko** vyjadruje mieru ohrozenia aktíva, mieru nebezpečenstva, že sa realizuje hrozba a dôjde k nežiaducemu výsledku,

t. j. ku škode. **Hrozba** je sila, udalosť alebo osoba, ktorá má škodlivý vplyv na bezpečnosť alebo môže spôsobiť škodu (Smejkal, 2013).

1.11.1. Identifikácia rizík

Pre identifikáciu rizík je dôležité zisťovať a štruktúrovane evidovať významné **potenciálne riziká**, na ktorých sa potom vykoná ich klasifikácia do určených klasifikačných skupín. Riziká sú zisťované napríklad pomocou brainstormingu (Smejkal, 2013).

1.11.2. Vyhodnotenie rizík

Vyhodnotenie rizika je proces porovnávania odhadnutého rizika s danými kritériami pre určenie jeho významu (Sedlák, 2023, str. 16). Vyhodnotenie dôležitosti identifikovaných rizík sa obvykle deje skupinou odborníkov, napríklad vo forme riadenej skupinovej diskusie s respondentami. Vyhodnotenie môže mať formu kvalitatívnej bodovej stupnice 1-5, kde sa **pravdepodobnosti** a **dopadu** rizika pridávajú hodnoty na tejto stupnici (Smejkal, 2013).

1.11.3. Mapa rizík

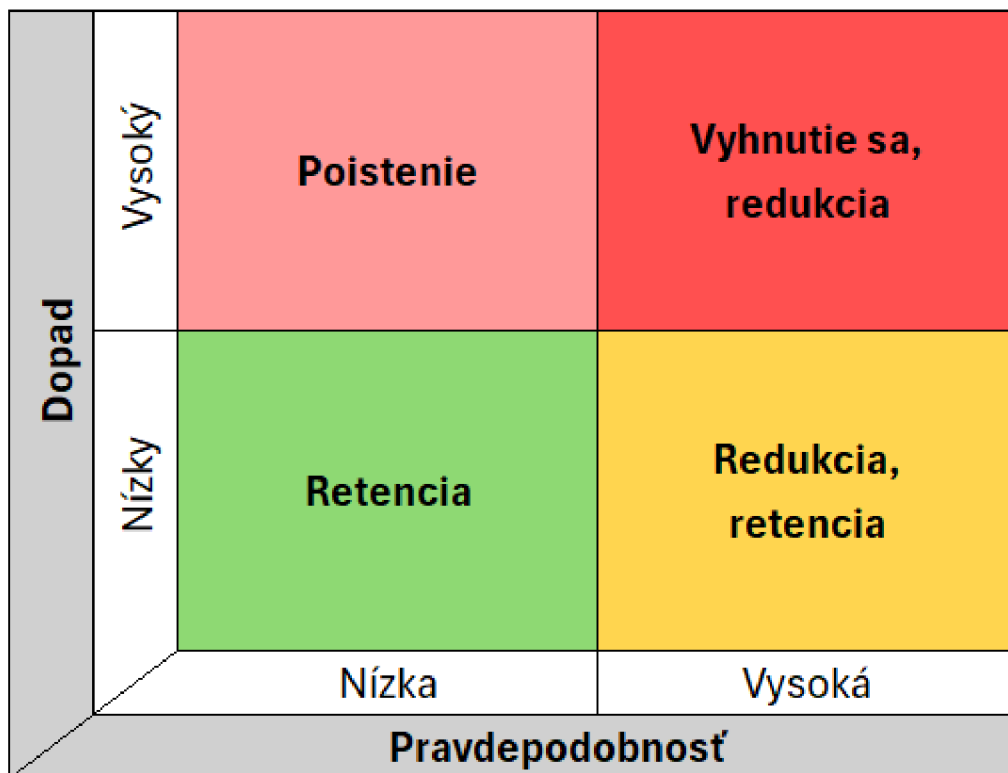
Riziká, ktoré boli identifikované a vyhodnotené ako hlavné, sú predmetom ďalšieho podrobnejšieho skúmania pomocou mapy rizík. Mapa rizík je grafickým znázornením relatívneho postavenia a významnosti hlavných rizík. Ide o dvojrozmerný graf polôh rizík v súradniciach pravdepodobnosti výskytu rizika P a dopadu rizika D. Významnosť rizika R sa potom vypočíta ako $R = D \times P$ (Smejkal, 2013).

1.11.4. Protiopatrenia a metódy znižovania rizika

Pod pojmom protiopatrenie sa rozumie proces, postup alebo čokoľvek, čo je špeciálne navrhnuté na zmiernenie alebo odstránenie účinku hrozby. Protiopatrenia sú navrhované tak, aby zabránili vzniku škody alebo uľahčili prekonanie následkov škody (Smejkal, 2013).

Po identifikácii rizík je kľúčové navrhnúť vhodný postup ich riešenia. Riziká môžeme efektívne adresovať v závislosti od pravdepodobnosti ich vzniku a možného dopadu. Existujú rôzne stratégie napríklad:

- **Retencia** rizika sa odvoláva na analýzu, ktorá naznačuje, že pravdepodobnosť naplnenia hrozby je zanedbateľná alebo že jej potenciálny dopad je prijateľný.
- **Redukcia** rizika sa sústreďuje na elimináciu príčin, ktoré môžu riziko spôsobiť, a zároveň na minimalizáciu jeho nepriaznivých dôsledkov.
- **Poistenie** rizika je koncepcia vychádzajúca z teórie rizika zmeny, kde malá, ale istá strata je akceptovaná výmenou za ochranu pred veľkými stratami.
- Možnosť **vyhnúť sa** riziku úplne, čo je vtedy, keď je situácia hodnotená ako príliš riziková, a preto sa rozhodneme danú zmenu nepreviesť (Smejkal, 2013).



Obrázok č. 10: Metódy znižovania rizika
(Zdroj: Vlastné spracovanie podľa Smejkal, 2013)

1.12. Návratnosť investície

ROI

Návratnosť investície (ROI - Return on investment) je ukazovateľ výkonnosti, ktorý sa používa na hodnotenie efektivity alebo ziskovosti investície, alebo na porovnanie efektivity rôznych investícií. ROI sa snaží priamo merať množstvo návratnosti danej investície v porovnaní s nákladmi na investíciu (Investopedia, 2024).

Na výpočet ROI sa zisk/benefit/návratnosť investície delí nákladmi na investíciu. Výsledok sa vyjadruje ako percento alebo pomer (Investopedia, 2024). Vzorec je nasledovný:

$$ROI = \frac{\text{Hodnota investície} - \text{Máklady na investíciu}}{\text{Náklady na investíciu}}$$

ROSI

Návratnosť investície do zabezpečenia (ROSI – Return on Security Investment) je výpočet navrhnutý ENISA a vychádza z obecné rozšírenej metodiky ROI. Ide o porovnanie peňažnej hodnoty investície s peňažnou hodnotou zníženia rizika (Sedlák, 2023). Počíta sa ako:

$$ROSI = \frac{\text{Monetary loss reduction (MLR)} - \text{Cost of the solution}}{\text{Cost of the solution}}$$

kde:

MLR je finančné vyjadrená redukcia straty. Počíta sa ako $ALE * MR$.

MR je percentuálne vyjadrená čiastka toho, o aký relatívny pomer dokáže opatrene znížiť nebezpečie dopadu rizika.

ALE je ročná očakávaná strata a počíta sa ako $SLE * ARO$.

ARO je miera pravdepodobnosti, že sa riziko zrealizuje v období jedného roku.

SLE je očakávaný finančný objem straty, pokiaľ sa zrealizuje bezpečnostné riziko (Sedlák, 2023).

Vzorec by teda bolo možné zapísať aj ako:

$$ROSI = \frac{((SLE * ARO) * MR) - \text{Cost of the solution}}{\text{Cost of the solution}}$$

1.13. Zhrnutie

Cieľom tejto časti diplomovej práce bolo pomocou rešerše odbornej literatúry vytvoriť komplexný teoretický základ pre nasledujúcu analytickú a návrhovú časť. V úvode kapitoly boli vymedzené základné pojmy a názvoslovie pre vymedzenie problematiky tejto práce a pre jednotné porozumenie nadväzujúcich pojmov a definícií.

Riešená problematika bola ďalej rozvinutá pomocou definícií kľúčových pojmov, ktorými sú Data Management, Data Governance, ich porovnanie a zároveň predstaveniu rámca DAMA, ktorý je obsahuje najlepšie postupy v oblasti správy dát a je kritický pre nasledujúcu analytickú časť. V samostatných kapitolách boli priblížené dátové oblasti Metadata Management a Data Security, na ktoré naviazala kapitola nástroje, ktorej obsahom bol opis nástrojov podporujúcich tieto dátové oblasti. Pre zhodnotenie výsledkov analytickej časti bol zároveň popísaný teoretický základ pre Maturity Assessment.

Zvyšné kapitoly boli určené pre oboznámenie sa s postupmi, ktoré budú použité v návrhovej časti, a to manažment zmien pomocou Lewinovho modelu, časové plánovanie projektu pomocou PERT, analýza rizík a postup pre ekonomické zhodnotenie investície pomocou ROI a ROSI.

2 ANALÝZA SÚČASNÉHO STAVU

Cieľom analytickej časti je analyzovať a zhodnotiť súčasný stav správy dát vo vybranej spoločnosti. Analytická časť vychádza najmä z interných zdrojov informácií. Interné informácie o zamestnávateľovi poskytnú zamestnanci vo forme rozhovorov, e-mailovej komunikácie a interných dokumentov.

V úvode tejto časti budú poskytnuté základné informácie o spoločnosti, ktoré je predmetom ďalších analýz a následne navrhnutého riešenia. Prvou aktivitou pre vykonanie analytickej časti bude analýza existujúcich dokumentov súvisiacich so správou dát vo vybranej spoločnosti. Po analýze dokumentov budú zorganizované stretnutia s oddeleniami, ktoré pracujú s dátami, za účelom získania hlbšieho porozumenia praktických aspektov správy dát. Druhou aktivitou bude vyhotovenie formulárov a ich vyplnenie vrcholovým manažmentom a kľúčovými používateľmi dát v spoločnosti.

Z uvedených aktivít vyplynú dva hlavné výstupy. Prvým bude prehľad identifikovaných medzier v správe dát, ktoré budú systematizované pomocou DAMA a PPTG. Tieto prehľady budú mať za cieľ pomôcť prioritizovať oblasti, ktoré vyžadujú zlepšenie. Prehľady budú doplnené o konkrétne identifikované medzery v rámci oddelení.

Druhým výstupom bude Maturity Assessment, ktorý poskytne vizualizáciu súčasnej úrovne vyspelosti v spoločnosti v kontexte správy dát. Zhodnotenie bude rovnako vykonané pomocou DAMA a PPTG.

Nakoniec budú tieto dva výstupy skombinované a na základe ich výsledkov budú identifikované oblasti s najväčším potenciálom pre zlepšenie. Oblasti s najväčším potenciálom pre zlepšenie budú detailne popísané v zhrňujúcej kapitole a zistenia v nich formulované budú slúžiť ako základ pre návrhovú časť práce. Zároveň je potrebné podotknúť, že z dôvodu citlivosti údajov budú niektoré dáta anonymizované.

2.1. Základné informácie o spoločnosti

Spoločnosť XYZ, pôsobí na českom energetickom trhu ako dôležitý dodávateľ elektriny a plynu. Svoje služby a produkty orientuje na domácnosti, podniky a veľké priemyselné spoločnosti, pričom sa zameriava na inovatívne energetické riešenia a zvýšenie efektívnosti energie. Súčasťou jej stratégie je rozvoj obnoviteľných zdrojov, kde sa snaží

expandovať vo využívaní solárnej a veternej energie. XYZ je tiež aktívna v oblasti energetických služieb, kde ponúka poradenstvo v oblasti úspor energie, energetické audity a riešenia pre smart metering, čo sú inteligentné meracie systémy, ktoré umožňujú zákazníkovi lepšie monitorovať a riadiť ich spotrebu energie.

V snahe o dosiahnutie uhlíkovej neutrality, XYZ plánuje postupnú transformáciu svojich energetických zdrojov a infraštruktúry smerom k väčšej ekologickosti a udržateľnosti. To zahŕňa nielen investície do zelených technológií, ale aj zlepšovanie operatívnej efektívnosti a redukciu emisií v rámci svojich operácií.

Spoločnosť XYZ si tak kladie za cieľ nielen poskytovať spoľahlivé a bezpečné dodávky energie, ale aj aktívne prispievať k ochrane životného prostredia a podpore udržateľného rozvoja v Českej republike.

2.2. Analýza medzier

V rámci analýzy medzier (GAP analýzy) bol po definovaní kľúčových procesov v rámci organizácie vytvorený zoznam mítingov so zamestnancami, ktoré bolo potrebné uskutočniť s cieľom získať podrobné informácie o uplatňovaní procesov správy dát a ich vplyve na každodennú činnosť spoločnosti. Stretnutia boli rozdelené na tie, ktoré sa týkali IT, a tie, ktoré sa týkali obchodných oddelení.

Celkovo bolo uskutočnených 56 stretnutí a identifikovaných 109 medzier, ktoré majú taktický alebo operatívny charakter. Všetky stretnutia prebehli v priebehu 60 minút a ich cieľom bolo zistiť hlavné kompetencie, každodenné interakcie s dátami spoločnosti a problémy súvisiace s dátami.

Zistené medzery boli overené členmi centrálného tímu spoločnosti a boli priebežne prioritizované. Medzery boli zoskupené do skupín podľa konkrétnych oblastí, ktorých sa týkajú a ich počet a zaradenie je možné vidieť v tabuľkách uvedených nižšie. Prehľad vybraných medzier je predmetom nasledujúcej kapitoly tejto práce.

Tabuľka č. 1: Zaradenie medzier v rámci oblastí PPTG
(Zdroj: Vlastné spracovanie)

PPTG oblasť	Počet medzier	% z celku
Ľudia (People)	12	11%
Procesy (Process)	32	29%
Technológia (Technology)	34	31%
Správa (Governance)	31	28%

V tabuľke č. 1 je možné vidieť medzery a ich zaradenie v rámci PPTG oblastí. V tabuľke uvedenej nižšie sú tieto medzery zaradené do znalostných oblastí podľa DAMA.

Tabuľka č. 2: Zaradenie medzier v rámci znalostných oblastí
(Zdroj: vlastné spracovanie)

ID	Znalostná oblasť	Počet medzier	% z celku
0	Data Governance Core	7	6%
1	Data Architecture	3	3%
2	Data Modelling & Design	8	7%
3	Data Storage & Operations	6	6%
4	Data Security	3	3%
5	Data Integration & Interoperability	22	20%
6	Document & Content Management	7	6%
7	Reference & Master Data	3	3%
8	Data Warehousing & BI	28	26%
9	Metadata	13	12%
10	Data Quality	9	8%

2.2.1. Identifikované medzery v rámci oddelení

V rámci analýzy medzier bolo zistené, že existujú mnohé priame dôsledky chýbajúcich dátových procedúr na obchodnú činnosť spoločnosti. Niektoré z hlavných obchodných vertikáli a oddelení majú nedostatočný prístup a nevhodné nástroje, preto si vytvárajú vlastné "zdroje pravdy" dát v Exceli. Hlavné medzery, ktoré boli zistené, sú:

1. nedostatočné vlastníctvo dát,
2. množstvo manuálnej práce s dátami,
3. chýbajúce metadáta,
4. chýbajúca dokumentácia v súvislosti dát a dátových procesov,
5. nedostatočná dokumentácia v súvislosti s manažmentom zmien.

Detailnejší opis uvedených päť medzier je predmetom zvyšku tejto kapitoly.

1. Nedostatočné vlastníctvo dát

Mnohé problémy vznikajú z toho, že v spoločnosti nie sú pridelené role vlastníkov dát. Absencia týchto rolí a implementovaného modelu vlastníctva vedie k tomu, že neexistuje náležité alebo jednotné chápanie kľúčových ukazovateľov, primeranej dátovej kvality

a štandardu, zodpovednosti za dôveryhodnosť dát, definovanie potreby uchovávanía dát atď. Medzi niektoré príklady patria napr.:

- Žiadni vlastníci dát v dátovom sklade.
- Obchodní zástupcovia a tretie strany (agentúry) môžu predložiť zmluvu bez toho, aby boli vyplnené všetky potrebné údaje. V dôsledku toho v systéme chýbajú niektoré kľúčové údaje, napr. číslo zmluvy.
- Neexistuje definícia vlastníkov dát vo všetkých dátach z hľadiska správy dát.
- Proces tvorby reportov vo všeobecnosti nie je formalizovaný a vynucovaný, čo v niektorých prípadoch vedie k tomu, že oddelenia musia krížovo kontrolovať výstupy reportov v Power BI so zdrojovými systémami.

2. Manuálna práca

Pravdepodobne najvýraznejší problém týkajúci sa práce obchodných oddelení s dátami v spoločnosti je v spojitosti s ich manuálnou prácou. Existuje mnoho oddelení, ktoré denne manuálne spracovávajú dátové sady. Medzi niekoľko takýchto príkladov patrí napr.:

- Finančné oddelenie A nemôže pre svoje výkazy používať jediný dátový zdroj a na zhromažďovanie a úpravu potrebných dát musí používať program Microsoft Access - to isté platí pre finančné oddelenie B a C.
- Zdrojové dáta pre interné výkazy v rámci B2C obchodného oddelenia sú uložené v Excelových tabuľkách, ktoré si vyžadujú každodennú aktualizáciu.
- Reporty Power BI pre zákaznicke oddelenie sú závislé od manuálnej aktualizácie zdrojových Excelov a prístup k nim má len ich vlastník.
- Neexistuje žiadna integrácia dát medzi online chatovou aplikáciou na webovej stránke a CRM systémom. Zamestnanci musia pridávať komunikačné udalosti ručne.

Väčšina týchto medzier je dôsledkom chýbajúcej integrácie dát s DWH. Oddelenia sú potom nútené zbierať a upravovať dáta manuálne pomocou najflexibilnejšieho dostupného nástroja - Microsoft Excel alebo Microsoft Access. Tieto riešenia sú alternatívou technicky spoľahlivejšieho riešenia, ktoré by si vyžadovalo úplnú integráciu do DWH.

3. Chýbajúce metadáta

Spoločnosť nespravuje metadáta systematickým spôsobom. V súčasnosti chýbajú prvky metadát, ako je dátový katalóg, biznisový glosár a dátová línia. To spôsobuje problémy v oblasti efektívneho vyhľadávania dát, odstraňovania problémov s dátovou kvalitou a konzistentnosťou dát. Medzi príklady chýbajúceho manažmentu metadát patria napr.:

- Neexistuje aktuálny katalóg pre všetky KPI a žiadna synchronizácia so skupinou.
- Analytický tím nemá zoznam aktuálne nasadených reportov ani centrálnu úložisko pre ich skripty.
- Neexistuje centrálnu úložisko týkajúce sa používania CRM pre zákaznícke obchodné oddelenia a prípadne ďalšie oddelenia (napr. definícia leadov nie je jednotná v rámci obchodných vertikál).

4. Chýbajúca dokumentácia

Dokumentácia procesov je nedostatočná alebo neúplná a neexistuje jednotný postup dokumentovania a mapovania činností alebo aktualizácie dokumentácie. V oddeleniach, kde sa dokumentácia vedie, sa na jej uchovávanie používajú rôzne metódy a nástroje (napr. Confluence, Teams, SharePoint). Roztrieštená obchodná dokumentácia znamená, že každé oddelenie má určitú dokumentáciu obchodných procesov, ale je udržiavaná neštruktúrovaným a nejednotným spôsobom. Medzi príklady tohto stavu patrí napr.:

- Dokumentácia k reportom Power BI nie je riadne udržiavaná. K dispozícii je len pôvodná dokumentácia, ktorá však neodráža úpravy v priebehu času.
- Oddelenie auditu sa nemôže spoliehať na mapy procesov, ktoré boli vytvorené v prechádzajúcich rokoch, pretože sú neúplné a zastarané. Aby porozumeli jednotlivým procesom, musia sa stretávať s vlastníkmi procesov a učiť sa od nich. Početné obchodné vertikály a oddelenia neudržiavajú alebo nemajú aktualizované mapy procesov a dokumentáciu.

5. Manažment zmien

Zmeny v dátach vrátane pridania, odstránenia a úpravy dátových modelov a tabuliek sa uskutočňujú bez komplexnej dopadovej analýzy. V súčasnosti nie je v rámci podniku zavedený jednotný proces riadenia zmien. Hoci tímy, ktoré vykonávajú drobný vývoj, postupujú podľa procesu v JIRA, chýbajú ľudia, ktorí by vykonávali analýzu dopadu zmeny na dáta, DWH a reporting.

2.3. Maturity Assessment

Hodnotenie vyspelosti (Maturity Assessment) je proces, ktorý zisťuje aktuálny stav vyspelosti správy dát v spoločnosti. Výstup je založený na výpočte úrovne vyspelosti na vopred definovanej stupnici. Stupnice pre Assessment sú predmetom prílohy č. 1 a prílohy č. 2.

Spoločnosť bola hodnotená v dvoch rôznych rámcov pomocou dvoch hodnotiacich modelov. Použité rámce:

- hodnotenie na základe znalostných oblastí definovaných DAMA,
- hodnotenie na základe oblastí PPTG.

Tabuľka č. 3: Počet tém použitých v modeloch podľa znalostnej oblasti
(Zdroj: Vlastné spracovanie)

ID	DMBOK Znalostná oblasť	Všetky témy	Témy v KH	Témy v DH
0	Data Governance Core	32	3	17
1	Data Architecture	30	2	18
2	Data Modeling & Design	29	8	20
3	Data Storage & Operations	39	13	25
4	Data Security	26	6	20
5	Data Integration & Interoperability	55	11	30
6	Document & Content Management	38	4	9
7	Reference & Master Data	34	1	11
8	Data Warehousing & BI	45	8	26
9	Metadata	27	5	15
10	Data Quality	23	5	10
	Celkom	358	66	201

Použité modely:

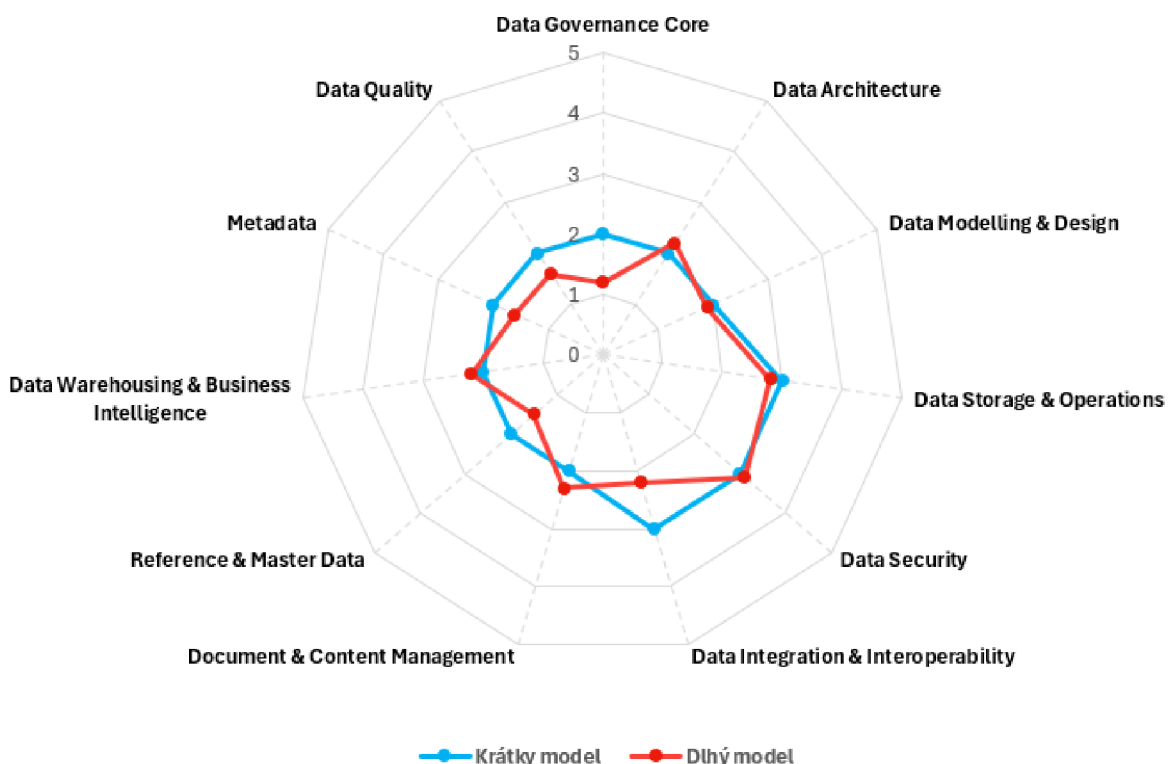
- krátky hodnotiaci model (KH) s pevne stanoveným súborom 66 tém, bez podpory modelu PPTG - cieľovou skupinou krátkeho hodnotenia je vrcholový manažment,
- dlhý hodnotiaci model (DH) s 358 témami, ktoré možno rozšíriť a prevážiť, s podporou modelu PPTG - dlhé hodnotenie je lepšie pre dátových odborníkov.

Krátke hodnotenie bolo rýchle a časovo nenáročné riešenie. Model dlhého hodnotenia bol hlbší/rozsiahlejší a časovo náročnejší. Krátke hodnotenie bolo podmnožinou modelu dlhého hodnotenia. Každá znalostná oblasť DAMA v krátkom/dlhom hodnotení bola priradená k PPTG oblasti. Normalizované výstupy umožňujú porovnávať výsledky v čase. Bez hodnotenia vyspelosti neexistuje praktický spôsob, ako efektívne zlepšiť Data Governance alebo definovať merateľné ciele.

2.3.1. Výsledky DMBOK maturity assessmentu

Na grafe č. 1 je možné vidieť výsledky krátkeho a dlhého assessmentu DAMA. Krátke hodnotenie, ako bolo vyššie spomenuté, zahŕňalo obmedzený súbor 66 tém. Konečný výsledok je 2,3 (priemer všetkých oblastí znalostí DMBOK2), čo znamená **úroveň 2 - opakovateľná** a teda je zavedená minimálna procesná disciplína. Dlhé hodnotenie poskytuje detailnejší pohľad. Výsledok je 2,1, čo opäť znamená **úroveň 2 - opakovateľná**. Dlhý výsledok je o niečo nižší než v krátkej verzii.

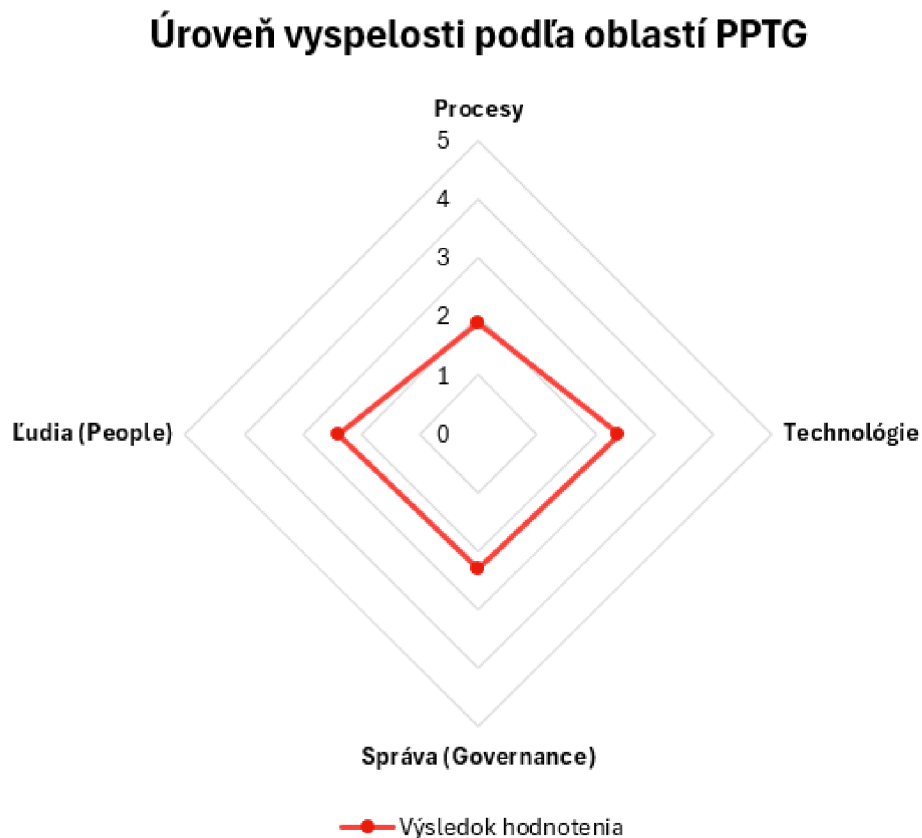
Súčasná úroveň vyspelosti podľa znalostných oblastí DAMA



Graf č. 1: Úroveň vyspelosti podľa DMOK
(Zdroj: Vlastné spracovanie podľa DAMA, 2017)

2.3.2. Výsledok podľa PPTG

V kapitole 1.8.2. bolo spomenuté, že tradičné rámce riadenia dát, ako je DAMA, často obsahujú príliš komplexné štruktúry, ktoré môžu byť ťažko aplikovateľné a preto bol rámec PPTG rovnako aplikovaný na súčasný ekosystém spoločnosti súbežne s hodnotením DMBOK. Model hodnotí jednotlivé oblasti na stupnici pomocou šiestich úrovní rovnako, ako DMBOK2. Výsledky je možné vidieť na grafe uvedenom nižšie.



Graf č. 2: Úroveň vyspelosti podľa PPTG
(Zdroj: Vlastné spracovanie)

Model vyspelosti PPTG bol vyvinutý interným tímom materskej spoločnosti a na rozdiel od DAMA, ktorá sa sústreďuje na 11 oblastí, sa PPTG sústreďuje iba na štyri hlavné oblasti: ľudí (People), procesy (Process), technológie (Technology) a správu (Governance).

2.4. Vyhodnotenie súčasného stavu

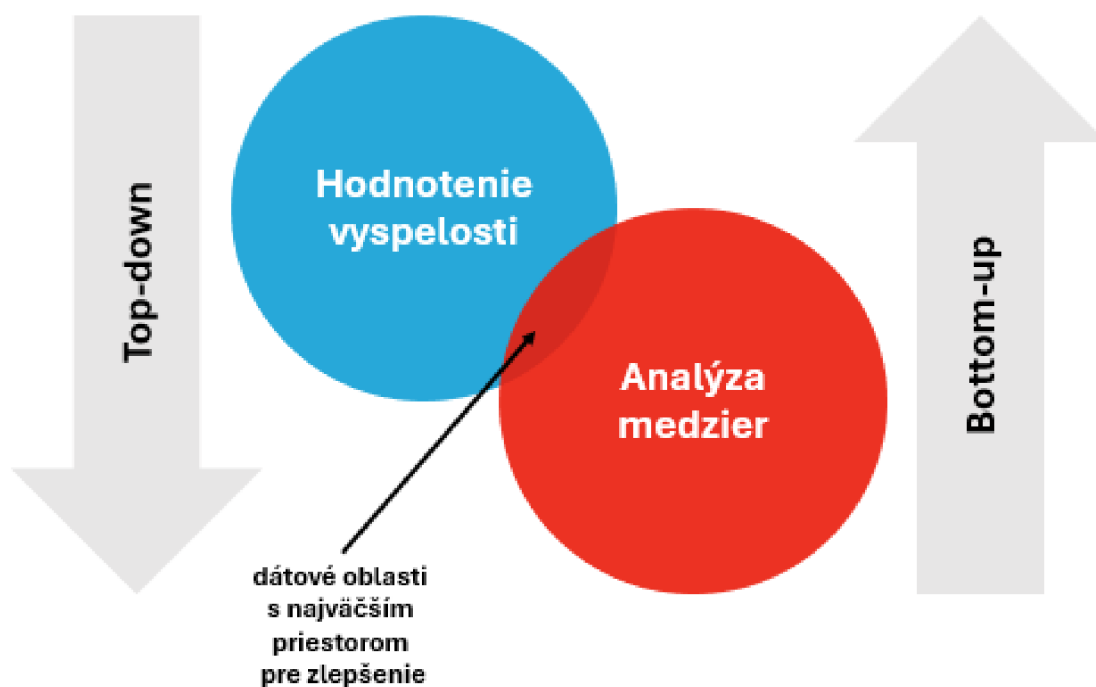
V predchádzajúcich kapitolách bol analyzovaný súčasný stav spoločnosti v dvoch dimenziách:

- úroveň vyspelosti Data Governance (strategická a taktická úroveň),
- Data Governance medzery a medzery súvisiace s dátami (operatívna úroveň).

Úroveň vyspelosti popisuje, aké dátové kapacity spoločnosť má a ako ich riadi, kontroluje a používa. Proces úrovne vyspelosti využíva prístup „top-down“ na preskúmanie všetkých dátových schopností súčasného stavu.

Na druhej strane sú tu medzery v dátach. Dátové medzery sú predmetom každodenných problémov na nižších úrovniach, ktoré narúšajú manažment dát, spôsobujú dodatočné náklady, zhoršujú zákaznícku/užívateľskú skúsenosť a spomaľujú firemné operácie. Proces zisťovania medzier je prístupom „bottom-up“ súčasného stavu.

Analýza medzier (GAP analýza) bola vypracovaná na základe rozhovorov so zainteresovanými stranami z oblasti IT/dát a z oblasti biznisu, ktoré denne pracujú s dátami. Informácie získané počas stretnutí spolu s podrobným prehľadom dostupných materiálov pomohli vykresliť obraz o postupoch správy dát v spoločnosti.



Obrázok č. 11: Oblasti s najväčším priestorom pre zlepšenie
(Zdroj: Vlastné spracovanie)

Aby bolo možné plne pochopiť súčasný stav spoločnosti, sú v tejto kapitole spojené výstupy oboch procesov s cieľom identifikovať najhoršie fungujúce časti manažmentu a správy dát. Výsledkom sú 3 kľúčové oblasti v oblasti dát, ktoré boli definované ako najhoršie fungujúce:

- Data Governance Framework,
- Metadata,
- Data Quality.

Vyspelosť všetkých týchto oblastí je veľmi nízka a v podstate chýbajú základné stavebné prvky vo všetkých oblastiach. To má za následok slabé porozumenie dátam vo všetkých biznis oddeleniach a množstvo operácií súvisiacich s dátami, ktoré nevytvárajú hodnotu pre spoločnosť. Vyššie uvedené 3 kriticky nízke dátové oblasti, ako aj ďalšie oblasti, boli predmetom komplexného hodnotenia vyspelosti, ktoré kombinovalo metodiku DAMA a rámca PPTG.

Tabuľka č. 4: Výsledné hodnotenie vykonaných analýz

(Zdroj: Vlastné spracovanie)

#	Znalostná oblasť	Výsledné hodnotenie
0	Data Governance Core	1,2
9	Metadata	1,5
10	Data Quality	1,6
7	Reference & Master Data	1,7
2	Data Modelling & Design	1,9
1	Data Architecture	2,2
5	Data Integration & Interoperability	2,2
8	Data Warehousing & Business Intelligence	2,2
6	Document & Content Management	2,3
3	Data Storage & Operations	2,8
4	Data Security	3,1

Výsledkom bolo určenie úrovne **vyspelosti analyzovanej spoločnosti na úroveň 2,1 - opakovateľná**. To znamená, že analyzovaná spoločnosť vykonáva reaktívne procesy správy údajov, ktoré sú podporované nedostatočne alebo v podstate neexistujúcimi normami alebo politikami bez vhodných nástrojov, ktoré by umožnili automatickú správu údajov. Oblasti sú detailne opísané v nasledujúcich podkapitolách.

2.4.1. Data Governance Framework

Všeobecne platí, že rámec je skutočná alebo koncepcná štruktúra, ktorá slúži ako opora alebo sprievodca pri budovaní niečoho, čo rozširuje štruktúru na niečo hodnotné a spracovateľné. Data Governance Framework by mal byť hlavnou kostrou každej iniciatívy Data Governance, ktorá spája politiky, štandardy a postupy do celého ekosystému podporujúceho dátovú stratégiu.

Súčasný stav

Z hľadiska **ľudí** neexistuje na lokálnej úrovni tím pre DG. Úlohou centrálného DG tímu je poskytovať poradenstvo neexistujúcemu tímu na lokálnej úrovni. Vlastníci dát nie sú zadaní a vlastníci aplikácií ich napodobňujú.

Z hľadiska **procesov** centrálny tím poskytol rámec pre zhodnotenie Data Governance vybranej spoločnosti. Súčasný procesy súvisiace s dátami sú obmedzené na operatívne úkony. Iné procesy zas napodobňujú nástroje Data Governance rôznymi obchádzkami (napr. vlastníctvo aplikácie napodobňuje vlastníctvo dát atď.)

Súčasná **technológia** pre DG má len dve prevádzkové časti a to Confluence pre spoluprácu a SAP PowerDesigner na dátové modelovanie/dopadovú analýzu v obmedzenom rozsahu.

Z hľadiska **správy** spoločnosť nezaviedla lokálny rámec Data Governance. Existuje len niekoľko globálnych politík s obmedzeným rozsahom, dopadom a detailmi. Rámec Data Governance musí definovať nasledovné:

- ciele stratégie,
- sponzora Data Governance,
- zainteresované strany v oblasti údajov,
- výbor pre Data Governance (DGSC),
- Data Governance ako takú,
- KPI pre manažment týkajúci sa dát,
- základné politiky a najkritickejšie postupy a spoločné šablóny,
- relevantné nástroje.

Najväčšie medzery

Na základe predchádzajúcej kapitoly boli ako najväčšie medzery vyšpecifikované tieto oblasti.

1. Chýbajúca dátová stratégia

Spoločnosť nedisponuje žiadnym štruktúrovaným dokumentom o dátovej stratégii ani zoznamom dátových priorít z biznisového hľadiska. Technická časť dátovej stratégie je opísaná ako rozšírenie súčasných možností pre replikácie dát, Big Data a Sandboxing bez akejkoľvek závislosti od priorít biznis oddelení.

2. Slabé koncepcie a role v oblasti Data Governance

Súčasný koncept Data Governance na lokálnej úrovni je založený na centrálnej vízii Data Governance bez náležitého základu na lokálnej úrovni. Samotný koncept je nevypracovaný a nezrelý. Téma Data Governance sa však na lokálnej úrovni vyvíja a hľadá sa jej najefektívnejšia forma. Neexistuje jasný, dohodnutý dlhodobý plán rozvoja IT, ktorý by obsahoval Data Governance. Je prítomný rozpor medzi požiadavkami IT a biznisu.

3. Limitovaný počet globálnych noriem

Existuje limitovaný počet globálnych noriem bez dlhodobého publikačného plánu. Stratégia skupiny je založená na federatívnom prístupe. Tiež je nejasná minimálna požiadavka na Data Governance pre lokálne obchodné jednotky. Tento chýbajúci základ predstavuje riziko nekompatibility a duplicity so skupinovými riešeniami v budúcnosti. Lokálne jednotky musia byť aktívnym členom globálneho programu Data Governance, aby boli v súlade s víziou a programami skupiny.

4. Žiadne dátové vlastníctvo

Neexistujú žiadni dedikovaní vlastníci dát. Chýbajúce vlastníctvo dát je v mnohých prípadoch napodobované vlastníctvom aplikácie, ale tento prístup nie je vhodný pre dátovú analýzu, ktorá integruje dáta do sekundárnych dátových kópií. Hranice odvodeného vlastníctva dát sa strácajú a neexistuje spôsob, ako dáta manažment. Takisto nie je možné efektívne riadiť dátovú kvalitu bez správnych vlastníkov.

2.4.2. Metadáta

Pri vytváraní funkčnej stratégie metadát a jej implementácii musí spoločnosť zväziť svoje zdroje metadát, metódy spracúvania metadát a požadované prístupové práva.

Súčasný stav

Z hľadiska **ľudí** na správu metadát nie je vyčlenený žiadny zamestnanec.

Mnohé **procesy** súvisiace so správu metadát nie sú v spoločnosti formálne organizované, ako napr.:

- analýzy metadát, ktorý hodnotí metadáta v spoločnosti, ich potenciál na získanie a ich úlohu v celkovej stratégii metadát,
- získavania metadát, ktorý zabezpečuje získavanie a pravidelnú aktualizáciu relevantných metadát,
- integrácie metadát, ktorý pomáha integrovať dáta vo všetkých potrebných systémoch.

Neexistencia týchto procesov a nedostatok ich dokumentácie zvyšuje náročnosť ďalších súvisiacich procesov, ako je proces dopadovej analýzy a proces riadenia problémov, čo sa prejavuje napríklad tým, že zamestnanci nemôžu nájsť príslušné informácie v centrálnom úložisku, ale musia sa pýtať kompetentných zamestnancov a sami metadáta vyhľadávať.

Z hľadiska **technológie** v súčasnosti neexistuje v spoločnosti koordinovaný postup správy metadát, preto ak sa metadáta vytvárajú z nutnosti alebo z osobnej iniciatívy jednotlivých zamestnancov, vytvárajú sa bez centrálne stanovených usmernení a ukladajú sa na osobné alebo tímovo zdieľané disky. V organizácii neexistuje žiadny špecializovaný nástroj na správu metadát, ako napríklad biznisový glosár alebo dátový katalóg. Dátová línia je vytvorená v programe SAP PowerDesigner, ale nie je úplná a nepodporuje nástroj na správu dát erwin DIS, ktorý bude v spoločnosti pravdepodobne implementovaný.

Z hľadiska **správy** sa v spoločnosti udržiavajú čiastkové metadáta, ktoré sú však slabo zdokumentované. Oddelenia často nemajú zoznamy reportov, ktoré používajú, a majú obmedzenú dokumentáciu zdrojových údajov. Okrem toho spoločnosť nie je synchronizovaná v oblasti obchodných pojmov a KPI s materskou spoločnosťou.

Najväčšie medzery

Na základe predchádzajúcej kapitoly boli ako najväčšie medzery v oblasti metadát vyšpecifikované nasledovne:

1. Chýbajúci Dáta katalóg

Dátové katalógy sa v spoločnosti nevedú a preto existuje obmedzený prehľad o obsahu databáz a tabuliek.

2. Chýbajúci Business Glossary

Žiadne oddelenie nemá formálne vedený biznis glosár, ktorý by definoval a popisoval jednotlivé biznis pojmy, ktoré sa na ich oddelení používajú, a to či už v zdieľaných dokumentoch alebo v špecializovanom biznis glosárskom nástroji. Jedinou výnimkou je novo vzniknuté oddelenie D, kde sa zavádzajú nové procesy a počíta sa s obchodným slovníkom pre onboardovanie nových zamestnancov.

3. Chýbajúci centralizovaný katalóg reportov

Centralizovaný katalóg reportov obsahuje zoznam všetkých reportov, ktoré sa v spoločnosti aktuálne vytvárajú, spolu s popisom ich obsahu, zdrojových údajov a zdrojových skriptov. Tento katalóg slúži ako referenčný bod v prípade, že niekto potrebuje vidieť určitý report alebo chce vytvoriť nový report a overiť, či podobný report už neexistuje.

4. Nekompletná Data Lineage

Súčasná dátová línia je k dispozícii len pre malú časť dátového prostredia. Tento obmedzený rozsah sa v aplikácii SAP PowerDesigner udržiava manuálne a nemôže zaručiť prehľad o pohybe údajov a závislostiach v dátovom prostredí.

5. Aktualizácia a udržiavanie katalógu služieb

Súčasný katalóg služieb v systéme LeanIX nie je aktualizovaný o všetky aplikácie, ktoré spoločnosť používa. Tento katalóg je jedným zo základných úložísk metadát, ktoré slúži na lepšiu orientáciu v technologickom prostredí spoločnosti.

6. Chýbajúce centralizované úložisko metadát pre biznis aj IT

IT oddelenia používajú Confluence a niektoré pracovné priestory sú dobre udržiavané, ale toto úložisko by sa malo zaviesť pre všetky oddelenia spolu s pravidlami jeho používania a údržby (pravidlá pre indexovanie, katalogizáciu, šablónovanie).

2.4.3. Data Quality

Všetky disciplíny riadenia dát prispievajú ku kvalite dát, a preto možno riadenie kvality dát v istom zmysle vnímať ako vrcholovú činnosť, ktorá ťaží alebo trpí vyspelosťou všetkých ostatných disciplín riadenia dát. Formálne by sa manažment kvality dát mal riadiť počas celého životného cyklu stanovením štandardov pre procesy, ktoré vytvárajú, transformujú, uchovávajú a vizualizujú dáta.

Správny program riadenia kvality dát by sa mal riadiť jasne definovanými hlavnými firemnými cieľmi. Mal by zaviesť kritériá kvality dát, pravidlá, kontrolné mechanizmy a procesy riešenia problémov s cieľom zabezpečiť čo najvyššiu kvalitu dát počas celého životného cyklu dát.

Súčasný stav

Z hľadiska **ľudí** je riadenie dátovej kvality naráža predovšetkým na obmedzené kapacity súčasných zamestnancov, ktorí nemajú dostatok času na ďalší rozvoj jej rozsahu a implementáciu širšej škály pravidiel a kontrol. V ideálnom prípade by existoval špecializovaný tím, ktorý by riadenie dátovej kvality prebral ako svoju hlavnú agendu. Napriek tomu sa však k výraznému zlepšeniu postupov v riadení dátovej kvality dosiahne až vtedy, keď spoločnosť zavedie správne DG role v rámci celej svojej spoločnosti, pretože role, ako sú Data Stewards a Data Custodians, sú kľúčové pri udržiavaní kvality údajov v ich špecializovaných dátových doménach. Tieto úlohy v súčasnosti v spoločnosti chýbajú.

Z hľadiska **procesov** súčasný rámec dátovej kvality zaviedol niekoľko procesov, ktoré hodnotia kvalitu dát, ale jeho rozsah je pomerne obmedzený a dokáže riešiť len malú časť celkového dátového ekosystému v spoločnosti.

Z hľadiska **technológií** je v súčasnosti riadenie dátovej kvality reprezentované najmä rámcom dátovej kvality, čo je súbor pravidiel, kontrol a zásad, ktoré sú prevedené do skriptov. Rámec kvality údajov v súčasnosti čiastočne pokrýva len technické a biznis kontroly dát zo systémov SAP CRM a ISU. Existuje interne vyvinutý nástroj na monitorovanie a riadenie dátovej kvality, ktorý však nepokrýva všetky aspekty dátovej kvality. Spoločnosť nemá formalizovaný a štruktúrovaný proaktívny procesný prístup k riadeniu dátovej kvality v zmysle definovania, porovnávania, bodovania,

monitorovania, opravy a vykazovania dátovej kvality. Prístup k riadeniu dátovej kvality je väčšinou reaktívny, pričom existujú niektoré nástroje, ktoré uľahčujú zlepšovanie dátovej kvality, ale tieto nie sú kontextualizované v žiadnej dohodnutej stratégii a sú implementované len v niekoľkých kritických systémoch. Súčasný postupy nezahŕňajú najmä tieto prvky:

- Kontroly manuálneho zadávania dát - keď systém vyžaduje manuálne vloženie dát, často neexistujú žiadne kontroly pre samotné vloženie týchto dát. Spoločnosť sa pri zadávaní správnych dát do systémov spolieha na skúsenosti svojich zamestnancov.
- Profilovanie dát - dáta vstupujúce do DWH nie sú nijako posudzované a sú vkladané priamo z primárnych systémov.
- Bodové hodnotenie a vyhodnocovanie údajov - spoločnosť nestanovila, ktoré údaje spĺňajú požiadavky na kvalitu a ktoré nie, a následne nevyhodnotila kvalitu odrážajúcu tieto požiadavky.
- Overovanie dát - údaje v DWH sa overujú v obmedzenom počte prípadov (napr. mesačne na finančnom kontrolingu, faktúry vykonávané manuálne).
- Riešenie problémov s dátami - v súčasnosti existuje proces, ale jeho výkonnosť často trpí nedostatkom kapacít.

Najvýznamnejšou politikou v oblasti **správy**, ktorá je implementovaná len v obmedzenej miere, je stratégia dátovej kvality a s ňou súvisiace postupy. Stratégia často obsahuje definíciu business motivácie, ktorá vedie k vytvoreniu konkrétneho opatrenia, a konkrétne business problémy, ktoré má program riadenia dátovej kvality vyriešiť. Okrem toho sa však riadenie dátovej kvality reprezentované rámcom dátovej kvality uplatňuje len v SAP ISU a SAP CRM a nepokrýva údaje v celom ekosystéme.

Najväčšie medzery

Ako bolo uvedené v predchádzajúcich podkapitolách, aj v tej podkapitole bolo na základe predchádzajúcej kapitoly definované najväčšie medzery:

1. Chýbajúce vlastníctvo

V spoločnosti chýba riadne a formálne nastavenie rolí DG, čo je jedným z hlavných faktorov, ktoré prispievajú k problémom s dátovou kvalitou v spoločnosti, pretože dáta

nemajú určených vlastníkov, stevardov a kustódov. V súčasnosti sa spoločnosť spolieha na "kľúčových používateľov" jednotlivých systémov, ktorí v prípade akýchkoľvek problémov musia prejaviť osobnú iniciatívu a nezrovnalosti odstrániť.

2. Definovaná stratégia dátovej kvality nie je v požadovanom detaile a nie je ani plne implementovaná

V súčasnosti neexistuje žiadna zastrešujúca politika alebo dokument, ktorý by jasne špecifikoval podnikové ciele, rozsah, role, konkrétne iniciatívy a činnosti, ktoré by malo riadenie kvality údajov v spoločnosti dosiahnuť. Tento dokument by poskytol základ pre budúce činnosti a určil by ich priority.

Ktoré dáta sú pre spoločnosť kritické? Ktoré dáta musia byť vždy bezchybné a ktoré dáta nemusia byť v takom dokonalom stave? Kto sú ľudia zodpovední za udržiavanie dátovej kvality? Základom riadenia dátovej kvality by mala byť komplexná stratégia dátovej kvality.

3. Dáta v DWH občas nezodpovedajú zdrojom

V spoločnosti chýba riadne a formálne nastavenie rolí DG, čo je jedným z hlavných faktorov, ktoré prispievajú k problémom s dátovou kvalitou v spoločnosti, pretože dáta nemajú určených vlastníkov, stevardov a kustódov. V súčasnosti sa spoločnosť spolieha na "kľúčových používateľov" jednotlivých systémov, ktorí v prípade akýchkoľvek problémov musia prejaviť osobnú iniciatívu a nezrovnalosti odstrániť.

4. Rozsah súčasného rámca DQ je obmedzený na SAP ISU a SAP CRM

Existujúci rámec dátovej kvality je skvelým príkladom osobnej proaktivity a iniciatívy, ale vzhľadom na obmedzené kapacity nie je možné rámec ďalej rozvíjať. Preto pokrýva len malú časť ekosystému firemných dát a dokáže odhaliť len určitú časť problémov spojených s dátovou kvalitou.

5. Kontakty neplatných klientov nie sú v DWH po neúspešnom doručení e-mailu označené ako "Neplatné".

Existuje len obmedzený počet mechanizmov, ktoré by overovali kontakty v systéme CRM na účely GDPR. Niektoré kontakty chýbajú alebo sú neaktuálne. Ak sa po e-mailovej kampani zistí, že kontakt je neplatný, nie je v DWH takto označený.

2.5. Zhrnutie

Cieľom tejto časti diplomovej práce bolo vhodným spôsobom analyzovať súčasný stav správy dát vo vybranej spoločnosti. Zároveň bol v úvode práce na základe osobných skúseností vyslovený predpoklad, že v rámci analýzy budú zistené rozlične nedostatky, ktoré bude potrebné vhodným spôsobom kategorizovať a prioritizovať. Z toho dôvodu bol súčasný stav spoločnosti skúmaný na operatívnej úrovni pomocou analýzy medzier a na taktickej a strategickej úrovni pomocou Maturity Assessmentu.

V rámci analýzy medzier prebehli stretnutia s používateľmi dát naprieč organizáciou, vďaka ktorým bolo identifikovaných 109 medzier, ktoré súviseli s dátami, prácou s nimi a ich správou. Medzery boli zoskupené do skupín podľa konkrétnych oblastí, ktorých sa týkali a to v rámci znalostných oblastí DAMA a oblastí PPTG. Dopady medzier na činnosť jednotlivých oddelení boli predmetom samostatnej podkapitoly.

Úlohou Maturity Assessmentu bolo zhodnotiť úroveň vyspelosti z pohľadu Data Governance. Za týmto účelom dva modely, jeden krátky a jeden dlhý v podobe hodnotiacich dotazníkov. Cieľovou skupinou krátkeho modelu bol vrcholový manažment. Na dlhý hodnotiaci model odpovedali kľúčový zamestnanci a odborníci na dátové oblasti.

Výsledky týchto skúmaní boli skombinované v kapitole 2.3, kde bol poskytnutý prehľad celkových výsledkov. Zároveň bol v tejto kapitole vyhodnotený súčasný stav v spoločnosti z hľadiska správy dát a to identifikovaním oblastí s najväčším potenciálom pre zlepšenie. Oblasti s najväčším potenciálom pre zlepšenie boli detailne popísané v jednotlivých podkapitolách a zistenia v nich formulované slúžia ako základ pre nasledujúcu návrhovú časť práce.

3 VLASTNÉ NÁVRHY RIEŠENIA

Na základe analýz vykonaných v predchádzajúcej časti diplomovej práce bola úroveň vyspelosti z pohľadu Data Governance vybranej spoločnosti vyhodnotená na úrovni dva z päť. Zároveň boli identifikované oblasti najväčším potenciálom pre zlepšenie z pohľadu správy dát.

Činnosti uvedené v tejto časti diplomovej práce predstavujú nevyhnutné kroky na prechod spoločnosti z úrovne vyspelosti 2 na úroveň 3 v krátkodobom horizonte a na úroveň 4 v dlhodobom horizonte. Je zrejmé, že zmena nemôže nastať zo dňa na deň; namiesto toho si vyžaduje veľmi komplexný plán týchto krokov, zoradených do postupného poradia, zohľadňujúc závislosti medzi nimi, postavený na konkrétnych problémoch, s ktorými sa potykajú jednotlivé oddelenia. K navrhovanej zmene je teda pristupované technokraticky, teda projektovo, čo znamená, že súčasťou návrhu bude sieťová analýza PERT. Pre zahájenie procesu riadenej zmeny je využitý Lewinov model. Zároveň, nakoľko je modelovaná plánovaná zmena pre potreby vrcholového vedenia spoločnosti, budú uvedené kritické miesta návrhu a bude upozornené na riziká, ktoré súvisia s vykonaním zmeny. Rovnako budú navrhnuté metódy, ako tieto rizikové miesta prekonať, prípadne ako znížiť riziko na únosnú mieru.

3.1. Popis navrhovanej zmeny

Na základe analytických údajov vytýčených v predchádzajúcej časti je možné vidieť, že spoločnosť pracuje s obrovským množstvom rozličných dát, ktorých správa spadá pod rôzne individuálne projekty a ad-hoc analýzy. Na údržbu metadát nie je vyčlenený žiadny personál. V organizácii neexistuje žiadny špecializovaný nástroj na správu metadát, ako napríklad business glosár alebo dáta katalóg. V spoločnosti sa udržiavajú čiastkové metadáta, ktoré sú však veľmi zriedka zdokumentované. Oddelenia často nemajú zoznamy reportov, ktoré používajú, a majú obmedzenú dokumentáciu zdrojových dát.

Z týchto dôvodov spoločnosť potrebuje v krátkodobom horizonte nastaviť úlohy a zodpovednosti v oblasti riadenia dát a v dlhodobom horizonte centralizovať činnosti manažmentu dát (ideálne pod vedením CDO alebo ekvivalentnej role). Ide teda o zmenu dátovej stratégie, ktorá bude vykonaná pomocou implementácie Data Governance Framework-u. Na základe vyššie uvedeného vyplýva potreba prijať 3 nových

pracovníkov s plným úväzkom (FTE) na podporu programu Data Governance Framework (vedúci DG, špecialisti DG) pod aktuálnym CDO. Zároveň bude spoločnosť musieť implementovať centrálnu DG platformu erwin DIS, ktorá je schválená na globálnej úrovni v rámci skupiny.

3.2. Lewinov model

V tejto kapitole bude navrhnutý model zmeny, ktorý pomôže spoločnosti prispôbiť sa zmene v správe dát a metadát a následne túto zmenu implementovať. Proces pozostáva z troch fáz: rozmrazenie, skutočnej zmeny a zmrazenie.

3.2.1. Fáza rozmrazenia

V prvej fáze Lewinovho modelu boli identifikované a kvantifikované sily, ktoré podporujú alebo bránia plánovanej zmene. Sily, ktoré iniciujú zmenu, alebo hnacie sily, boli hodnotené na stupnici od 1 do 5. Sily, ktoré iniciujú proti zmene, známe ako brzdiace sily, boli hodnotené na stupnici od -1 do -5. V oboch prípadoch ± 1 predstavuje minimálne hodnoty a ± 5 predstavuje maximálne hodnoty. Súčet hodnôt predstavuje konečné rozhodnutie. Pre identifikáciu síl bola použitá metóda silového poľa.

Tabuľka č. 5: Analýza silového poľa
(Zdroj: Vlastné spracovanie)

Sily podporujúce zmenu	Skóre	Sily brániace zmene	Skóre
Manuálna práca s dátami	+5	Návyk zamestnancov na existujúci proces	-5
Podpora vrcholového manažmentu	+5	Nevyužívanie novej technológie	-4
Podpora rýchlejšieho rozhodovania	+4	Neochota kľúčových používateľov	-3
Zber dát a ich malé využitie	+3	Technické obmedzenia pre implementáciu	-2
Integrované sledovanie všetkých metadát na jednom mieste	+2	Integrácia veľkého množstva zdrojov dát	-2
Celkové skóre	+ 19	Celkové skóre	-16

Hodnotenie poukazuje na to, že pozitívne sily, ktoré sú v prospech zmeny, majú väčšiu váhu (+19) v porovnaní s negatívnymi silami (-16). Zmenu je teda možné vykonať.

Hlavné sily podporujúce zmenu vyplývajú hlavne z negatívnych skúseností zamestnancov v súvislosti s prácou s dátami, najmä v oblasti zberu dát a následnej manuálnej práce s nimi. Dôležitými silami je silná motivácia vrcholového manažmentu pre vykonanie tejto zmeny. Na druhej strane hlavnú silu brániacu zmenu tvoria opäť zamestnanci a to najmä tí, ktorí dáta používajú na nepravidelnej báze a ktorí si zvykli na súčasný stav a nemajú iniciatívu svoje návyky meniť. V tejto súvislosti je dôležité dbať aj na názory kľúčových používateľov, ktorí rovnako majú zabehnuté procesy, ktoré nie vždy radi menia za nové. Celý návrh zmeny sa samozrejme opiera o predpoklad, že zavedený nástroj a politiky budú zamestnanci aktívne používať.

Identifikácia agenta zmeny

Agentmi pre zmenu budú CDO a vedúci oddelenia DG, ktorí budú zmenu realizovať.

Sponzorom bude CFO, ktorá má právomoc zaviesť zmenu a podporuje agentov zdrojmi.

Obhajcom zmeny je vedúci oddelenia BI Delivery, ktorý zmenu podporuje.

Intervenčné oblasti

Ludské zdroje: Implantácia Data Governance Frameworku pomerne zdĺhavý proces, ktorý sa dotkne najmä zamestnancov analytických a projektových tímov a v širšom rozsahu každého zamestnanca, ktorý je používateľom dát. Zmena bude mať dopad aj na tých zamestnancov, ktorí budú pasívnymi používateľmi rámca Data Governance, čoho následkom bude, že sa im zlepšia ich skúsenosti s dátami a s dátovými výstupmi. Chod samotných oddelení táto zmena nijak nenarušuje.

Organizačná štruktúra: Organizačná štruktúra spoločnosti sa nemení, avšak budú identifikovaný vlastníci jednotlivých dátových domén ako aj vlastníci, stevardi a správcovia dátových aktív a objektov. Tieto dopady sú bližšie objasnené v kapitolách 3.4.5. a 3.4.7.

Technológia firmy: Implementácia Data Governance nástroja erwin DIS umožní spoločnosti pravidelne mapovať dátové prostredie, identifikovať dátové aktíva, evidovať ich dôležitosť a skenovať metadáta v nich a ukladať ich v jednom nástroji, kde budú dostupné všetkým zamestnancom, bez rizika zobrazenia skutočných dát. Zároveň bude možné v tomto nástroji možné pracovať aj s ďalšími modulmi, akými je katalóg reportov alebo katalóg KPI.

Komunikačné a organizačné toky: Implementácia nového Data Governance nástroja erwin DIS povedie k zefektívneniu a automatizácii procesov súvisiacich so zberom dát a metadát a tvorby reportov, ktoré v súčasnom riešení zvyčajne vykonáva analytik/používateľ dát manuálne. Plánuje sa tiež prepojenie tohto systému s inými BU v rámci skupiny, aby sa získané metadáta zdieľali.

3.2.2. Fáza prechodu a aplikácie zmeny

Pre úspešné vykonanie druhej fázy zmeny bol vytvorený súbor na seba naväzujúcich činností, ktoré sa týkajú implementácie tejto zmeny. Cieľom týchto činností je pokryť zistené medzery a zvýšiť úroveň zrelosti kritických oblastí, ktoré boli identifikované v analytickej časti tejto diplomovej práce. Samotné činnosti sú rozpísané v tabuľke nižšie a ich obsah bude v detaile popísaný v kapitole 3.3. Vypracovanie projektu.

Tabuľka č. 6: Jednotlivé činnosti pre zmenu
(Zdroje: Vlastné spracovanie)

# č.	Činnosť	# č.	Činnosť
A	Vymenovanie sponzora DG programu	K	Vytvorenie politiky a postupov v oblasti metadát
B	Menovanie vedúceho tímu DG	L	Zabezpečenie platformy pre business glossary
C	Pridelenie rozpočtu na DG	M	Zabezpečenie dátového katalógu
D	Definícia dátovej stratégie	N	Zabezpečenie platformy pre data lineage
E	Prijatie špecialistu DG	O	Zabezpečenie Confluence ako centrálného CMS
F	Zriadenie výboru pre DG (DGSC)	P	Zabezpečenie report katalógu
G	Definovanie základnej politiky	R	Počiatočné zhromažďovanie metadát
H	Vymedzenie politiky vlastníctva dát	S	Zhromažďovanie pojmov do business glossary
I	Definovanie vlastníctva na úrovni domén	T	Zhromažďovanie reportov do report katalógu
J	Identifikácia DO, DS a TDS		

3.2.3. Fáza zmrazenia

V poslednej fáze Lewinovho modelu budú všetky základy pre Data Governance implementované do bežného fungovania spoločnosti.

3.2.4. Verifikácia dosiahnutých výsledkov

V Lewinovom modeli sa hodnotenie nevykonáva iba na konci procesu zmeny, ale vykonáva sa priebežne. Pokiaľ čiastkové výsledky zmenového procesu nie sú priebežne sledované, hrozí, že výsledná situácia vo firme bude rovnaká alebo horšia.

Odborná literatúra sa zhoduje v tom, že aby sa dosiahlo úspechu pri zavedení systému riadenia dát, musí sa program postupne vyvíjať. To znamená začať s menšími, dosiahnuteľnými krokmi, ktoré je možné priebežne merať a vyhodnocovať. Za týmto účelom je vhodné merať pokrok implementácie, súlad s požiadavkami a hodnotu, ktorú projekt spoločnosti prináša. Možné metriky zahŕňajú:

Tabuľka č. 7: Metriky pre meranie dosiahnutých výsledkov
(Zdroj: Vlastné spracovanie)

KPI	Cieľová hodnota
Vytvorený tím DG	Tím má 3/3 členov.
Vytvorené strategické dokumenty a politiky	4
Zmapované procesy pre aktivity DG	Procesy sú definované (BPMN2)
Účasť členov na DGSC	80%
Zadefinované dátové domény	6 až 12
Domény a majú svojich vlastníkov	100%
Identifikovaný DO/DS/DC	Min. 1/1/1 na doménu
Systémy sú evidované a automaticky skenovaných	100%
Automatizované reporty sú evidované a automaticky skenované	100% reportov v PBI a BOBJ
Vytvorený katalóg business pojmov spolu so všetkými potrebnými údajmi	100 pojmov (najviac používaných)
Vytvorený katalóg KPI spolu so všetkými potrebnými údajmi	Top 5 KPI pre každé oddelenie
Vzájomné prepojenie aktív v Data Lineage nástroji	5 až 10% (najviac kritických)

3.3. Vypracovanie projektu

Táto kapitola je zameraná na jednotlivé činnosti, ktoré boli stanovené v predchádzajúcej kapitole 3.2.2. a na ich vypracovanie.

3.3.1. Vymenovanie sponzora DG programu

Pre program Data Governance je potrebné vymenovať *sponzora*, ktorým vo vybranej spoločnosti bola CFO.

3.3.2. Menovanie vedúceho tímu DG

Vymenovanie kľúčovej osoby, *agenta zmeny*, ktorá bude riadiť činnosti Data Governance.

3.3.3. Pridelenie rozpočtu na DG

Prideliť a odsúhlasiť rozpočet. Rozpočet by mal byť dostačujúci aspoň do opätovného posúdenia časového plánu. Aktuálny časový plán by mal byť kľúčovým ukazovateľom rozpočtu. Zodpovednosť za aktivitu má sponzor zmeny.

3.3.4. Definícia dátovej stratégie

Vytvorenie zastrešujúceho dokumentu, ktorý musí celej spoločnosti povedať, prečo sa Dátová stratégia zavádza a aké benefity to prinesie. Dátová stratégia je dlhodobý plán, ktorý definuje technológie, procesy, ľudí a pravidlá potrebné na podporu získavania, organizácie, analýzy a poskytovania údajov na podporu cieľov spoločnosti. Jeho súčasťou by mala byť aj stratégia Data Governance. Zodpovednosť za aktivitu má agent zmeny.

3.3.5. Prijatie špecialistov na Data Governance

Prijatie špecialistov na Data Governance, ktorí budú pomáhať vedúcemu tímu DG pri realizácii iniciatív programu. Spolu s vedúcim tímu DG vytvorí tím DG. Zodpovednosť za aktivitu má agent zmeny.

Tím DG je podriadený CDO. Tím DG by mal pozostávať z týchto členov:

- tímoví vedúci pre DG,

- špecialista pre DG so zameraním pre informačný manažment,
- špecialista pre DG s technickým zameraním (zdrojové systémy, dátová kvalita,...)

Tímoví vedúci bude viesť program DG na každodennej báze. Špecialisti DG budú zodpovední za zavádzanie kľúčových procesov v rámci programu súvisiacich so získavaním, správou, vývojom a využívaním dátových zdrojov. Budú úzko spolupracovať s vedúcim tímu DG, CDO a vlastníkami údajov a môžu v niektorých prípadoch plniť úlohu správcov údajov.

3.3.6. Zriadenie výboru pre Data Governance (DGSC)

Vytvorenie výboru pre Data Governance (DGSC), ktorý bude hlavným riadiacim orgánom plánu DG. Zodpovednosť za aktivitu má sponzor zmeny.

Zvyčajne DGSC pozostáva z niekoľkých zamestnancov na úrovni B alebo B-1 vrátane vedúceho pracovníka pre Data Governance a teda CDO alebo vedúceho oddelenia pre DG. Toto zloženie je vhodným východiskom pre rozvoj programu DG vo vybranej spoločnosti. Výbor je kľúčovým orgánom pri rozhodovaní a riadení projektov v rámci programu. Jeho úlohou je schvaľovať jednotlivé činnosti a projekty a zabezpečovať komunikáciu a akceptáciu zo strany biznisových útvarov. Členovia by tiež mali pomáhať pri definovaní obchodných pojmov, pôsobiť ako orgán na riešenie konfliktov a nominovať jednotlivých vlastníkov dát. Výbor je zodpovedný za spoluprácu interného tímu s prípadnými zapojenými externými stranami.

DGSC je orgán, v ktorom sa môžu stretnúť zástupcovia všetkých biznisových oddelení a zosúladiť svoje priority. Nominácia jednotlivých členov závisí od interného rozhodnutia vybranej spoločnosti.

3.3.7. Definovanie základnej politiky

Definovanie základnej politiky správy údajov, tzv. Data Governance Policy, ktorá umožní vytvárať ďalšie politiky potrebné pre činnosť K a na ňu nadväzujúce činnosti. Táto politika vytvára prepojenie medzi organizačnou štruktúrou a funkciami DG. Tiež musí definovať hlavné prvky a metódy DG v spoločnosti. Rovnako predstavuje lokálnu aplikáciu globálnych Data Governance štandardov s tým, že rešpektuje všetky aspekty,

potreby a limity súčasného prostredia vybranej spoločnosti. Zároveň podporuje zásady a ciele Dátovej stratégie. Zodpovednosť za aktivitu má agent zmeny.

3.3.8. Vymedzenie politiky vlastníctva dát

Definujte politiku vlastníctva dát, ktorá umožní vytvárať ďalšie politiky. Zodpovednosť za aktivitu má agent zmeny. Činnosť je bližšie popísaná v samostatnej kapitole 3.4.

3.3.9. Definovanie vlastníctva na úrovni domén

Zodpovedný za udržiavanie a oznamovanie aktuálnej definície dátových domén a pridelovanie zodpovednosti vlastníkom dátových domén je CDO. Činnosť je bližšie popísaná v samostatnej kapitole 3.4.

3.3.10. Identifikácia DO, DS a TDS

DO, DS a DC sú základné role pre Data Governance. Činnosť je bližšie popísaná v samostatnej kapitole 3.4.

3.3.11. Vytvorenie politiky a postupov v oblasti metadát

Metadátová politika a jej základné postupy/normy poskytnú usmernenie pre implementáciu potrebných metadátových nástrojov, ako je napríklad biznisový glosár, dátový katalóg a dátová línia. Zodpovednosť za aktivitu má agent zmeny.

Zároveň, nakoľko nadväzujúce činnosti spoliehajú na jasné definície dát, odporúčam v tomto kroku prijať aj pravidlá, ktoré definujú formuláciu definícií dát. Politika, ktorá bude pravidlá obsahovať, by mala stanoviť, že definícií dát musia byť:

- a) **v jednotnom čísle** – definície sa píše výhradne v jednotnom čísle, s výnimkou pojmov, ktoré sú samé o sebe v množnom čísle,
- b) **definovať čo pojem je, nie čo pojem nie je** – nie je možné pojem definovať výhradne stanovením, čo pojem nie je, napríklad „sú to náklady, ktoré nezahŕňajú,...“,
- c) **byť v opisnej forme** - postačujúcou definíciou nie je uvedenie jedného alebo viacerých synonym, ani zmena poradia slov v názve prvku; pokiaľ je potrebné viac než jedno opisné ustálené vyjadrenie, použijú sa celé, gramaticky správne vety,

- d) **mať vysvetlené skratky** – v definícii musia byť použité celé slová, nie skratky z dôvodu, že v rôznych prostrediach môže byť skratka nesprávne alebo zmätočne interpretovaná,
- e) **byť bez vonkajších definícií** - v definícii by sa nemala objaviť definícia ďalšieho prvku alebo súvisiaceho pojmu.

Pravidlá uvedené vyššie by mali byť dostačujúce pre písanie definícií do platforiem uvedených v nasledujúcich bodoch.

3.3.12. Zabezpečenie platformy pre business glossary

Zavedenie modulu Business Glossary v Erwin pre vybranú spoločnosť po diskusii s centrálnym tímom. Zahŕňa nastavenie prostredia, licenčnú zmluvu, role a používateľský účet. Zodpovednosť za aktivitu má agent zmeny.

Ide o centralizované úložisko všetkých obchodných termínov dostupných vo vybranej spoločnosti s informáciami o vlastníctve, ich organizačnom zaradení a vzťahu k iným termínom v business glossary alebo aktívam v dátovom katalógu.

3.3.13. Zabezpečenie dátového katalógu

Zavedenie Dátového katalógu v Erwin pre vybranú spoločnosť po diskusii s centrálnym tímom. Zahŕňa nastavenie prostredia, licenčnú zmluvu, roly a používateľský účet. Zodpovednosť za aktivitu má agent zmeny.

Úložisko metadát s informáciami o všetkých existujúcich dátových zdrojoch. Malo by to byť jediné miesto, kde sa používajú všetky informácie o dátach a infraštruktúre používanej na správu dát.

3.3.14. Zabezpečenie platformy pre data lineage

Zavedenie modulu pre Data Lineage v Erwin pre vybranú spoločnosť po diskusii s centrálnym tímom. Zahŕňa nastavenie prostredia, licenčnú zmluvu, roly a používateľský účet. Zodpovednosť za aktivitu má agent zmeny.

Data lineage predstavuje podrobnú mapu všetkých priamych a nepriamych závislostí medzi dátovými entitami v spoločnosti. Poskytuje východiskový bod pre akúkoľvek dopadovú analýzu a umožňuje pochopiť dátový ekosystém.

3.3.15. Zabezpečenie Confluence ako centrálného CMS

V súčasnosti spoločnosť používa Confluence ako CMS (v preklade systém na správu obsahu), v ktorom sa ukladajú dokumenty a iné digitálne aktíva. Tento systém je potrebné zlepšiť z hľadiska jeho správy (štandardy indexácie, šablóny, zálohovanie) a mali by mať k nemu prístup všetky obchodné a IT oddelenia. V ideálnom prípade by mala byť nastavená štruktúra pre každú podstránku so správnym indexovaním a označením verzie všetkých dokumentov, ktoré by mali mať aj svoje preddefinované šablóny.

Confluence v spoločnosti technicky spravuje oddelenie IT a jednotlivé pracovné priestory by mali spravovať vedúci oddelení/projektov, ktorí by mali dodržiavať štandardy pre tvorbu pracovných priestorov a súvisiacej dokumentácie.

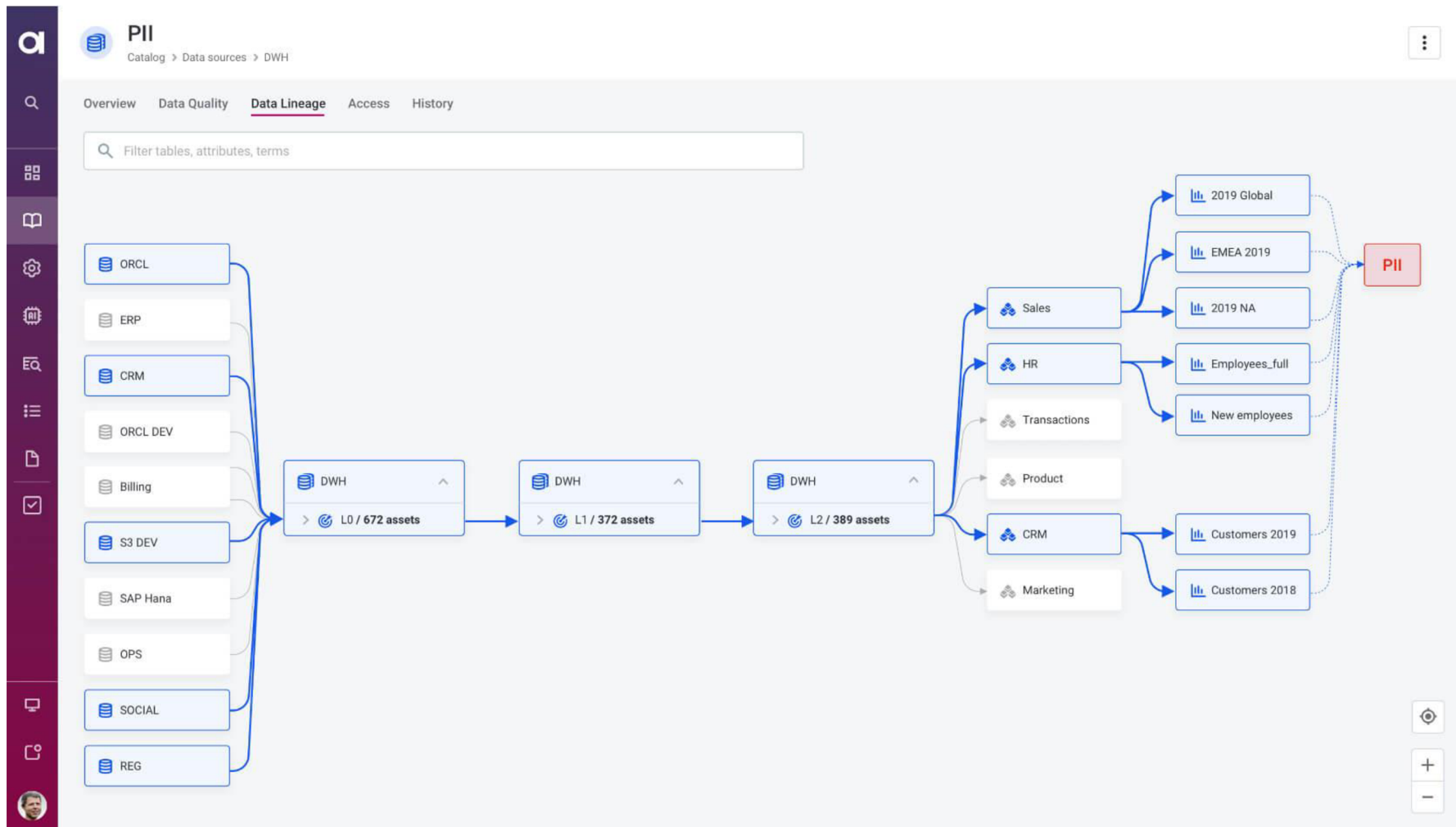
3.3.16. Zabezpečenie report katalógu

Zavedenie platformy pre report katalóg, v ktorom budú uvedené všetky reporty spolu s ich zdrojovými skriptami, popismi KPI, verziami atď.

3.3.17. Počiatočné zhromažďovanie metadát

Počiatočný proces zberu zahŕňajúci metadát z rôznych zdrojov v spoločnosti. Niektoré metadáta je možné zbierať automaticky pomocou nástrojov DG, niektoré metadáta bude potrebné zozbierať a aktualizovať ručne. Nakoľko ide o počiatočné štádium, odporúčam zamerať sa na biznisovo najvýznamnejšie systémy, z ktorých je zároveň možné zber metadát zautomatizovať pomocou connectora. Connector by mal byť schopný skenovať systém a všetky tabuľky a stĺpce v ňom. Zároveň je možné v rámci tohto procesu začať proces počiatočného zberu metadát pre data lineage, napr. zber metadát pre PII.

Na obrázku č. 12 je možné vidieť, že na ľavej strane sú skenované jednotlivé zdrojové systémy dát s príslušnými tabuľkami a stĺpcami. Dáta z týchto systémov prechádzajú do DWH, kde sa spracovávajú v troch vrstvách a následne vstupujú do dátových modelov až konkrétnych reportoch. Zároveň je v pravom hornom rohu možné vidieť červenou farbou označené pole PII, čo značí, že dátová línia, ktorá je zobrazená, sa týka práve dát kategorizovaných ako PII.



Obrázok č. 12: Zobrazenie PII v dátových aktívach pomocou Data Lineage
 (Zdroj: Ataccama, 2023)

3.3.18. Zhromažďovanie pojmov do business glossary

Proces napĺňania business glossary obchodnými termínmi. Ako bolo vyššie spomenuté, ide o globálne definované termíny, ktoré predstavujú používanie obchodnej terminológie. Slúžia ako spoločná a zdieľaná definícia kľúčových obchodných pojmov a dátových prvkov, čím pomáhajú zabezpečiť konzistentnosť, zrozumiteľnosť a zosúladenie komunikácie a správy dát v rámci podniku. Zohrávajú základnú úlohu pri uľahčovaní efektívnej správy v rámci DG a podporujú lepšiemu porozumeniu dát.

Tabuľka č. 8: Polia pre business terms
(Zdroj: Vlastné spracovanie)

Pole	Popis	Príklad
Názov	Pole špecifikuje názov pojmu.	Odberné miesto
Definícia	Pole špecifikuje definíciu pojmu.	Miesto, na ktorom zákazník odoberá elektrinu alebo plyn z distribučnej siete.
Popis	Toto pole dokumentuje širší kontext pojmu, ktorý inak nie je vhodné zahrnúť do definície.	Miesto, na ktorom zákazník odoberá elektrinu alebo plyn z distribučnej siete, pričom je presne geograficky lokalizované pre účely merania spotreby.
Poznámky	Slúži pre referencie, poznámky, ak nejaké sú.	Jeden zákazník môže mať viacero odberných miest, napríklad dom a chata. Spotreba elektriny je obvykle uvádzaná v kilowattoch (kWh), spotreba plynu v kubických metroch (m ³).
Synonymá	Zoznam alternatívnych pojmov alebo slov, ktorými je možné pojem nahradiť.	Miesto odberu
Skratky	Pole pre skratky a akronymy s ich vysvetlením.	OM (Odberné Miesto), MO (Miesto Odberu)
Klasifikácia bezpečnosti	Označenie úrovne citlivosti dát. <ul style="list-style-type: none"> • Public (Low) • Internal (Medium) • Confidential (High) • Strictly Confidential (Very High) 	Internal

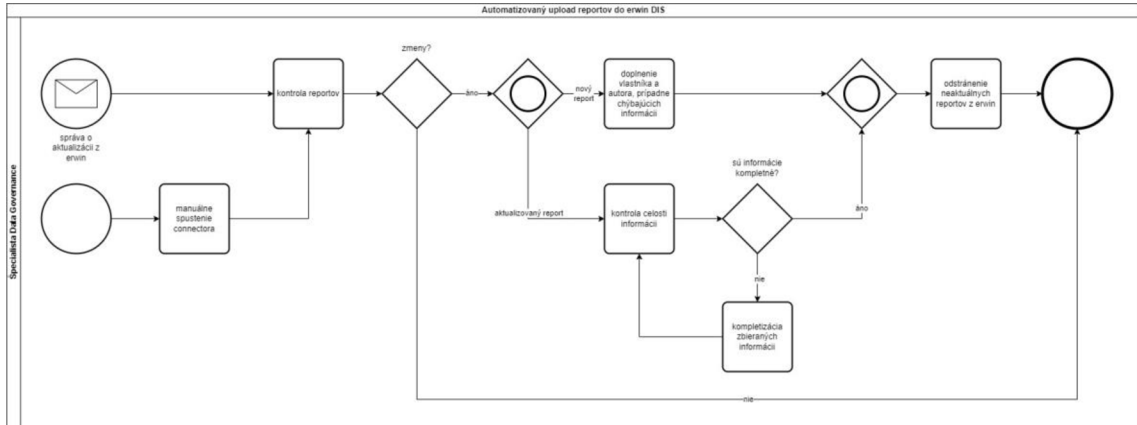
3.3.19. Zhromažďovanie reportov do report katalógu

Vo vybranej spoločnosti sa pre reportovanie používajú reporty vytvorené v Power BI, SAP Business Objects (BOBJ), Microsoft Excel a Microsoft Power Point. Pre počiatkové zhromažďovanie reportov odporúčam zamerať sa na reporty v Power BI a BOBJ, nakoľko ich zber by malo byť možné automatizovať prostredníctvom príslušných API. Pokiaľ ide o reporty manuálne, je potrebné zvážiť dôležitosť jednotlivých reportov a ich užitočnosť pre ostatných používateľov dát a teda zbierať a evidovať v katalógu iba také reporty, ktoré sú významné pre biznis. V nasledujúcej tabuľke je možné vidieť navrhované polia pre report katalóg.

Tabuľka č. 9: Polia pre report katalóg
(Zdroj: Vlastné spracovanie)

Pole	Popis	Príklad
Názov	Pole špecifikuje názov reportu.	Dashboard kampaní
Definícia	Pole špecifikuje definíciu reportu.	Report obsahuje počet zákazníkov, ktorí boli oslovení v priamych komunikačných kampaniach.
Popis	Toto pole obsahuje podrobný popis reportu a obsahuje viac kontextu.	Report obsahuje počet zákazníkov, ktorí boli oslovení v priamych komunikačných kampaniach podľa kanála, prostredníctvom ktorého bol zákazník oslovený s ponukou.
Vytvorený	Program, v ktorom bol report vytvorený	Power BI
Úložisko	Miesto, na ktorom je report uložený.	Worskpace „02 CRM“
URL	Miesto pre link na konkrétny report.	app.powerbi
Verzia	Aktuálna verzia reportu	1.0
Frekvencia aktualizácie	Ako často je report aktualizovaný/ako staré sú dáta.	Týždenne
Zdroj dát	Špecifikácia, či dáta pre daný report sú automatizované (DWH), manuálne alebo semi-automatizované.	Automatizovaný
Vlastník	Meno osoby, pre ktorú bol report vytvorený.	Anna Lytics
Autor	Meno osoby, ktorá report pripravila.	Terry Byte

Prípadný proces automatizovaného zberu reportov z Power BI a BOBJ je možné vidieť na obrázku uvedenom nižšie. Pre reporty tvorené manuálne, typicky vo formáte .xlsx, by bolo potrebné vytvoriť samostatný proces, ktorý môže zdieľať určité časti nižšie uvedeného, avšak jeho trvanie bude určite dlhšie.



Obrázok č. 13: Proces automatizovaného zberu reportov
(Zdroj: Vlastné spracovanie)

3.4. Definovanie dátových rolí

V tejto kapitole bude bližšie priblížená problematika definovania dátových rolí, ich identifikácia v organizácii a organizačné začlenenie.

3.4.1. Vlastník dátovej domény

Model dátovej domény je špecifikácia logických oblastí dát v spoločnosti. Príkladom domény dát môže byť "Zákazník" a všetky údaje týkajúce sa zákazníkov budú v tejto doméne a jej sub-doménach. Tento model bude základom pri definovaní vlastníctva dát. Vlastník domény dát by mal byť vedúci pracovník zodpovedný za celú doménu údajov a za menovanie dátových vlastníkov. Je to konečný subjekt vlastníci dáta. Jeden zamestnanec môže byť vlastníkom viacerých dátových domén. Ich hlavnými úlohami by mali byť:

- definícia hraníc domény, ktorú spravujú,
- definícia obchodných pojmov,
- riešenie konfliktov,
- menovanie vlastníkov dát.

Zároveň, nakoľko vo vybranej spoločnosti dátové domény definované nie sú, navrhujem zdefinovať osem hlavných dátových domén, ako je možné vidieť na obrázku nižšie. Uvedený spôsob prístupu k tvorbe domén je tzv. „grid“, alebo v preklade mriežkový, čo znamená, že domény sú v grid-e a sú si navzájom ekvivalentné. Tento prístup je veľmi flexibilný pre rozdeľovanie a zlúčenie dátových domén, avšak stanovenie hraníc je vecou dohody medzi vlastníkami. Okrem hlavných dátových domén je vhodné zdefinovať aj sub-domény, ktoré by obsahovali opis udalostí a činností definovaných v hlavných dátových doménach.

Z uvedeného vyplýva, že je v rámci spoločnosti potrebné identifikovať a menovať osem vlastníkov dátových domén. Nemusi ísť však o osem rôznych zamestnancov, nakoľko sa môže stať, že jeden zamestnanec bude vlastníkom napríklad dvoch dátových domén, ako bolo spomenuté vyššie. Zároveň je tento počet možné zväčšiť alebo zmenšiť podľa potrieb spoločnosti. Ak by sa granulácia ukázala ako príliš rozsiahla, je možné spojiť doménu Sales a doménu Market do jednej s názvom Sales & Market. Naopak je možné doménu Corporate rozdeliť na viacero menších celkov, ako napríklad doménu Business Entity a doménu Enterprise.



Obrázok č. 14: Mriežkový prístup k hlavným dátovým doménam
(Zdroj: Vlastné spracovanie)

3.4.2. Data Owners (DO)

V rámci programu DG je potrebné určiť data ownerov v jednotlivých dátových doménach. Data owneri budú mať zodpovednosť za dôveryhodnosť údajov, ako

aj za podporu bezpečnosti a ochrany dát v rámci ich dátovej domény. Data owneri budú niesť konečnú zodpovednosť za tieto úlohy týkajúce sa dát v rámci jeho dátovej domény:

- **vymenovanie biznisových dáta stewardov**, ktorí riadia každodenné činnosti,
- **dôveryhodnosť dát** – identifikácia kritických dátových elementov, dohodnutie sa na definícii biznisových pojmov, definovanie pravidiel, riešenie dátových problémov a zlepšovanie dátovej kvality,
- **bezpečnosť a ochrana** – špecifikácia toho, čo sa považuje za citlivé údaje, schvaľovanie prístupu k dátam a pravidelná správa opätovného overovania prístupov k dátam.
- **monitorovanie aktivít** – monitorujte priebežnú pokrok programu Data Governance, zúčastňujte sa na stretnutiach DGSC a zodpovedá za interné kontroly v prípade zistenia akýchkoľvek nedostatkov vo vyššie uvedených činnostiach.

Role DO by mali byť pokryté existujúcimi pozíciami v spoločnosti, konkrétne zamestnancami na úrovni B-2 a teda nie je potrebné zvažovať žiadny externý nábor.

3.4.3. Business Data Stewards (DS)

Business Data Stewards, tiež známy ako Data Stewards, budú zamestnanci, ktorí sú odborníkmi na jednu alebo viacero dátových domén a všetky ich elementy. Sú praktickými správcami dát, ktorí sa budú starať o to, aby dáta boli objavené na všetkých miestach, správne opísané, klasifikované a dostupné tam, kde sú potrebné, vo formáte požadovanom pravidlami politiky. Z uvedeného vyplýva, že súčasťou ich zodpovednosti bude aj dátová kvalita a priame aplikovanie náprav všetkých problémov zistených v rámci ich oblasti zodpovednosti. Mali by to byť členovia biznisových tímov na úrovni B-3. Zodpovednosti správcu údajov sú tieto:

- identifikovať kritické dátové elementy (CDE),
- definovať pravidlá na základe CDE,
- riešenie problémov s dátami odhalených počas práce s nimi,
- definovať hierarchie a zoskupenia,
- definovať, čo sa považuje za citlivé údaje,
- definuje hodnoty kódov a reviduje tabuľky s kódmi,
- mapujte hodnoty kódov medzi tabuľkami s kódmi v rôznych aplikáciách,

- overovať prístup používateľov ku kľúčovým systémom,
- poskytovať prístup k dátam s ohľadom na potreby biznisu, regulácie, politiky a reputáciu spoločnosti.

Je vhodné, aby úlohy DS boli pokryté existujúcimi pozíciami v spoločnosti. Konečné rozhodnutie však závisí od zmapovania očakávaných a existujúcich zručností budúcich DS v spoločnosti. Nie je neobvyklé, že správcovia údajov sú najímaní externe na plný úväzok v prípade, že v spoločnosti nie sú zamestnanci s požadovanými zručnosťami.

3.4.4. Technical Data Stewards (TDS)

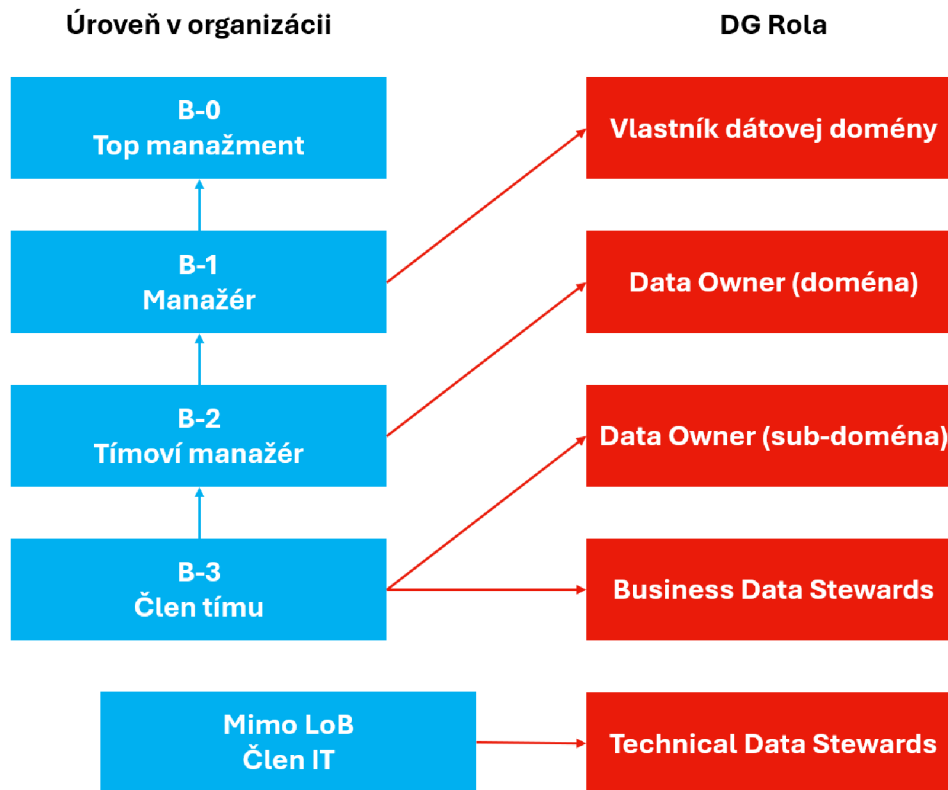
Technical Data Stewards, tiež pod pojmom Data Custodian, bude spadať pod oddelenie IT a bude mať technické znalosti o aplikáciách, platformách, dátasetoch atď. Bude musieť podporovať činnosti DS. Povinnosti tejto role budú nasledujúce:

- vytvárať SQL príkazy na základe pravidiel, ktoré definovali DS,
- zabezpečovať príslušné prepojenie so zdrojovými systémami na profilovanie dát,
- vytvárať postupy, aby sa (dátové) výnimky presmerovali na riešenie príslušný DS,
- pridávať, upravovať a odstraňovať kategórie pre obchodné pojmy, obchodné pravidlá a iné dátové artefakty.
- pracovať s technickými metadátami pre relačné databázy, nástrojmi na integráciu dát, zostavami, analytickými nástrojmi a dátovými modelmi,
- implementovať pravidlá párovania na základe vstupov od DS,
- pridávať/odstraňovať používateľov a pridávať/upravovať/odstraňovať oprávnenia pre používateľov,
- oznamovať zmeny v kóde,
- pracovať s nástrojom pre data discovery s cieľom odhaliť skryté citlivé údaje,
- monitorovať obsah databázy.

Opäť je vhodné, aby úlohy DC boli pokryté existujúcimi pozíciami v spoločnosti. Konečné rozhodnutie však znova závisí od zmapovania existujúcich zručností budúcich DC v spoločnosti a v prípade zistených nedostatkov je možné na túto rolu prijať externistov.

3.4.5. Model dátového vlastníctva

Aby bolo možné si role vlastníkov v spoločnosti lepšie predstaviť, je na obrázku nižšie uvedený model, ktorý poskytuje prehľad o tom, ako je vlastníctvo dát distribuované v rámci organizácie



Obrázok č. 15: Model dátového vlastníctva v spoločnosti
(Zdroj: Vlastné spracovanie)

V spoločnosti sú štyri úrovne hierarchie, začínajúc od B-0 úrovne, ktorou sa označuje top manažment až po B-3 úroveň, na ktorej sú jednotliví členovia tímu (referenti, špecialisti a pod.).

V modeli je okrem týchto štyroch úrovní oblasť s označením „Mimo LoB,“ čo znamená, že o podporný tím, ktorý pracuje mimo hlavnú líniu podnikania. Ako bolo spomenuté v predchádzajúcej kapitole, ide o členov IT tímu.

Klady tohto modelu spočívajú v praktickom a hlbokom porozumení dát, ktorým zamestnanci disponujú vďaka špecializácii na konkrétne oblasti. Model aplikuje princíp "rozdeľ a panuj", čo umožňuje efektívne rozdelenie zodpovedností a zvyšuje granularitu dohľadu nad dátami. Tento prístup zároveň zaisťuje, že práca s dátami je začlenená

do každodennej práce zamestnancov, čo by v ideálnom prípade malo viesť k zvyšovaniu ich zručností a kvality práce s dátami.

Na druhej strane, model má niekoľko záporov, medzi ktoré patrí obmedzená kapacita bez možnosti rozšírenia tímu novými zamestnancami. Zároveň, nakoľko zamestnanci pracujú s dátami, ktoré spravujú, na dennej báze, môžu mať tzv. „bias“ voči svojim dátam a teda náklonnosť nevedomky prehliadnuť nesúladu, ktoré sa v dátach nachádzajú. Ďalej môže nastať situácia, že kritické aktíva budú v rukách "slabších" vlastníkov, ktorí nemusia mať potrebné skúsenosti alebo zdroje na ich správu. Okrem toho vyžaduje model komplexnú spoluprácu medzi rôznymi úrovňami a oddeleniami, čo môže byť náročné na koordináciu a môže viesť k pomalšiemu rozhodovaciemu procesu.

Tento model môže byť upravený na viac úrovňový model tým, že B-1 manažér by bol vlastníkom domény a zároveň dátovým vlastníkom v danej doméne. Ďalej B-2 tímový manažér by bol dátovým vlastníkom v doméne a zároveň dátovým vlastníkom v sub-doméne. Tento model by vyriešil nedostatok „slabších“ vlastníkov kritických aktív. Zároveň by boli znalosti a správa o doménach optimálnejšie rozložené na viacerých zamestnancov. Model však so sebou prináša nové nevýhody, medzi ktorými sú najvýraznejšie nedostatok kapacít na DG činnosti na B-1 a B-2 úrovniach a nejednoznačnosť povinností. Z tohto dôvodu odporúčam sa držať modelu tak, ako je zobrazený na obrázku vyššie.

3.4.6. Časová náročnosť pre jednotlivé role

V nasledujúcej tabuľke sa je možné vidieť FTE alokáciu na vyššie uvedené role, ktoré je potrebné zaviesť v rámci programu DG. Je potrebné poznamenať, že ide o odhady a konečný počet FTE bude do veľkej miery závisieť od veľkosti definovaných domén.

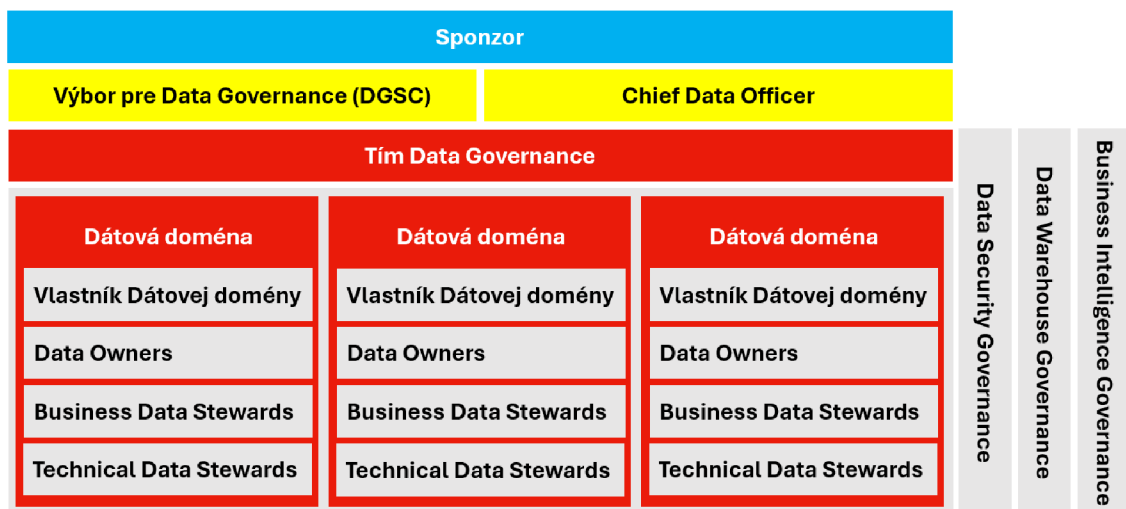
Tabuľka č. 10: FTE alokácia na role definované v rámci DG
(Zdroj: Vlastné spracovanie)

Rola	FTE na doménu
Vlastník dátovej domény	0,05 – 0,1
Data Owner	0,1 – 0,3
Business Data Stewards	0,3 – 0,5
Technical Data Stewards	0,3 – 0,8

Data Governance role by nemali byť vytvorené tzv. ad-hoc a dočasne vymenované pre dosiahnutie cieľa projektu. Mali by byť súčasťou organizačnej štruktúry a bez zmien v stratégii programu by sa nikdy nemali stať neobsadenými alebo nepridelenými. Väčšina rolí nie je plnohodnotnými rolami v zmysle pracovnej pozície, ale skôr doplnkovými povinnosťami ľudí na súčasných pracovných pozíciách. Program DG a povinnosti z neho vyplývajúce by však pre nich nemali mať nižšiu prioritu. Ak dôjde k nižšej alebo menej prioritnej úlohe, malo by sa to považovať za porušenie povinností danej osoby a malo by sa podľa toho postupovať.

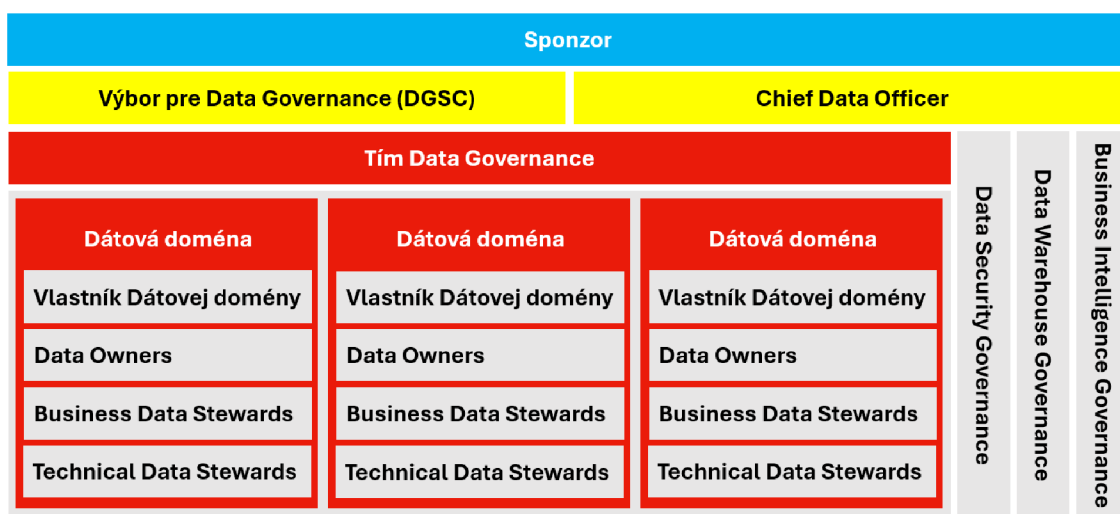
3.4.7. Organizačné zavedenie

Organizačné zavedenie Data Governance je navrhnuté v dvoch variantoch, z ktorých by mal byť jeden zvolený. **Prvý variant** má obmedzený vplyv na dátovú kultúru a jeho primárne zameranie sa týka vlastníctva dát a formálnej kontroly Data Managementu. V rámci tejto štruktúry by tím DG bol priamo podriadený CDO a vlastníci dátových domén by boli podriadení CDO po bodkovanej línii, ako je spomenuté v kapitole vyššie. Tímy podporujúce činnosti DG (security, DWH, BI) by nepodliehali CDO, ale spolupracovali by s tímom DG v rámci pravidelných stretnutí. Tento rámec je ľahko použiteľný ako jednoduchý doplnok k súčasnej organizačnej štruktúre a jeho grafické znázornenie je možné vidieť nižšie.



Obrázok č. 16: Organizačné zaradenie Data Governance – prvý variant
(Zdroj: Vlastné spracovanie)

Druhý variant podporuje kultúru založenú na dátach a pokrýva celý životný cyklus dát vrátane poskytovania informácií. Tento variant si vyžaduje úlohu CDO ako osoby zodpovednej za celkové činnosti data management-u vo vybranej spoločnosti. V rámci tejto štruktúry by tím DG bol priamo podriadený CDO a vlastníci dátových domén by boli podriadení CDO po bodkovanvej línii rovnako, ako je tomu v prechádzajúcom variante. Ostatné tímy, ktoré podporujú DG v spoločnosti, by tiež podliehali priamo CDO. Oproti predchádzajúcej variante si tento rámec vyžaduje hlbokú integráciu do súčasnej organizačnej štruktúry.



Obrázok č. 17: Organizačne zaradenie Data Governance – druhý variant
(Zdroj: Vlastné spracovanie)

Ako je možné vidieť, zavedenie druhého variantu si vyžaduje omnoho hlbšie integrovanie do organizačnej štruktúry, pričom zároveň poskytuje možnosť riadiť dáta end-to-end. Avšak nakoľko ide o počiatočne zavedenie Data Governance vo vybranej spoločnosti, odporúčam zaviesť prvý variant s tým, že v dlhom období je možné postupne prejsť na druhý variant a dosiahnuť požadovaného end-to-end účinku.

3.5. Časová analýza

Celý proces implementácie zmeny má 19 činností. Na určenie celkového trvania bola použitá metóda PERT. Každéj činnosti bol priradený optimistický, najpravdepodobnejší a pesimistický odhad trvania. Z týchto troch odhadov bolo určené priemerné trvanie, ktoré bolo použité na zostavenie sieťového grafu. Okrem toho boli určené základné štatistické ukazovatele a časové charakteristiky a nájdená kritická cesta.

Tabuľka č. 11: Legenda PERT
(Zdroj: Vlastné spracovanie)

i	Bezprostredne predchádzajúca činnosť
j	Nasledujúca činnosť
a_{ij}	Optimistický odhad trvania činnosti
m_{ij}	Najpravdepodobnejší odhad trvania činnosti
b_{ij}	Pesimistický odhad trvania činnosti
y_{ij}	Stredná doba trvania
σ²	Rozptyl
σ	Smerodajná odchýlka
ZM	Začiatok možný
KM	Koniec možný
ZP	Začiatok prípustný
KP	Koniec prípustný
RC	Rezerva celková
RV	Rezerva voľná

Tabuľka č. 12: Časová analýza PERT (Zdroj: Vlastné spracovanie)

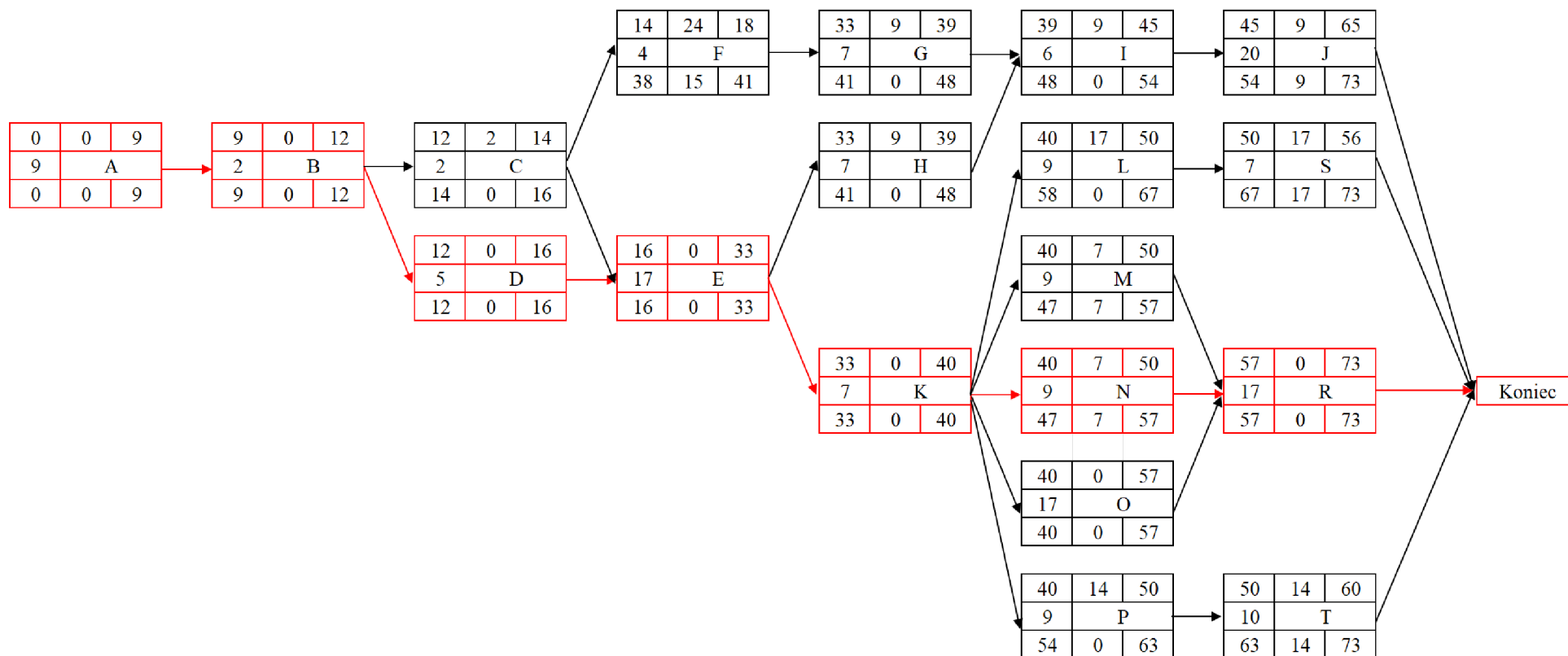
Údaje o činnosti projektu				Doba trvania [MD]				Štatistické ukazovatele		Termíny				
Činnosť	Popis činnosti	i	j	a _{ij}	m _{ij}	b _{ij}	y _{ij}	σ ²	σ	ZM	KM	ZP	KP	RC
A	Vymenovanie sponzora DG programu	-	B	5	9	15	9	2,78	1,67	0	9	0	9	0
B	Menovanie vedúceho tímu DG	A	C, D	1	2	5	2	0,44	0,67	9	12	9	12	0
C	Pridelenie rozpočtu na DG	B	E, F	1	2	5	2	0,44	0,67	12	14	14	16	2
D	Definícia dátovej stratégie	B	E	2	4	10	5	1,78	1,33	12	16	12	16	0
E	Prijatie špecialistu DG	C, D	H, K	15	16	20	17	0,69	0,83	16	33	16	33	0
F	Zriadenie výboru pre DG (DGSC)	C	G	2	4	5	4	0,25	0,50	14	18	38	41	24
G	Definovanie základnej politiky	F	I	5	6	10	7	0,69	0,83	33	39	41	48	9
H	Vymedzenie politiky vlastníctva dát	E	I	5	6	10	7	0,69	0,83	33	39	41	48	9
I	Definovanie vlastníctva na úrovni domén	G, H	J	12	15	20	15	1,78	1,33	39	45	48	54	9
J	Identifikácia DO, DS a TDS	I	-	12	20	25	20	4,69	2,17	45	65	54	73	9
K	Vytvorenie politiky a postupov v oblasti metadát	E	L, M, N, O, P	5	7	11	7	1,00	1,00	33	40	33	40	0
L	Zabezpečenie platformy pre business glossary	K	S	5	9	15	9	2,78	1,67	40	50	58	67	17
M	Zabezpečenie dátového katalógu	K	R	5	9	15	9	2,78	1,67	40	50	47	57	7
N	Zabezpečenie platformy pre data lineage	K	R	5	9	15	9	2,78	1,67	40	50	47	57	7
O	Zabezpečenie Confluence ako centrálného CMS	K	R	9	15	30	17	12,25	3,50	40	57	40	57	0
P	Zabezpečenie report katalógu	K	T	5	9	15	9	2,78	1,67	40	50	54	63	14
R	Počítačové zhromažďovanie metadát	M, N, O	-	10	15	30	17	11,11	3,33	57	73	57	73	0
S	Zhromažďovanie pojmov do business glossary	L	-	5	6	10	7	0,69	0,83	50	56	67	73	17
T	Zhromažďovanie reportov do report katalógu	P	-	7	10	15	10	1,78	1,33	50	60	63	73	14

Pre sieťový graf časovej analýzy metódy PERT slúži nasledujúca legenda:

Tabuľka č. 13: Legenda pre sieťový graf.

(Zdroj: Vlastné spracovanie)

ZM	RC	KM
yij	Činnosť	
ZP	RV	KP



Obrázok č. 18: Sieťový graf

(Zdroj: Vlastné spracovanie)

Kritická cesta, ktorá určuje najdlhšiu možnú dĺžku návrhu zmeny, bola určená pomocou metódy PERT v tabuľke č. 12. Na grafe uvedenom vyššie je možné vidieť zobrazenie kritickej cesty. Kritická cesta prechádza činnosťami: A, B, D, E, K, N, R. Oneskorenie jednej z identifikovaných kritických činností spôsobí oneskorenie celého návrhu zmeny. Odhaduje sa, že dokončenie projektu bude trvať 73 dní.

3.6. Analýza rizík

V kapitole analýza rizík sú definované riziká, ktoré plynú zo zmeny a implementácie novej dátovej stratégie pomocou Data Governance. Najprv budú identifikované potenciálne riziká spojené so zavedením programu Data Governance. Po identifikácii rizík sa každé riziko ohodnotí z hľadiska pravdepodobnosti, s akou môže nastať, a závažnosti jeho dopadu. Tieto informácie sa potom použijú na výpočet hodnoty rizika. Kvantifikácia sa vykoná pomocou bodovacej metódy. Pravdepodobnosť výskytu a miera vplyvu rizika sa vyjadri pomocou hodnôt na stupnici od 0 do 10. Hodnota rizika sa vypočíta ako súčin pravdepodobnosti výskytu a miery vplyvu a zistí sa na stupnici od 0 do 100. Nakoniec sa prijímú opatrenia na minimalizáciu rizík a ich vplyvu na zmenu. Hodnotiacu tabuľku pre pravdepodobnosť a pre dopad je možné vidieť nižšie.

Tabuľka č. 14: Riziková politika
(Zdroj: Vlastné spracovanie)

Subjektívna miera pravdepodobnosti			Dopad	
Hodnota	Percentuálne	Slovné hodnotenie	Hodnota	Slovné hodnotenie
1-2	0-19 %	Veľmi nepravdepodobná	1-2	Bezvýznamný
3-4	20-39 %	Npravdepodobná	3-4	Málo významný
5-6	40-59 %	Pravdepodobná	5-6	Významný
7-8	60-79 %	Viac pravdepodobná	7-8	Veľmi významný
9-10	80-100 %	Veľmi pravdepodobná	9-10	Kritický

3.6.1. Identifikácia a ohodnotenie rizík

V nasledujúcej tabuľke sú identifikované a kvantifikované riziká.

Tabuľka č. 15: Identifikácia a ohodnotenie rizík
(Zdroj: Vlastné spracovanie)

Analýza rizík					
#	Hrozba	Scenár	P-st	Dopad	Hodnota rizika
A	Výpoveď kľúčového zamestnanca	Implementácia zmeny sa oneskorí, vzniknú ďalšie náklady na výberové konanie.	2	9	18
B	Nedostatočné znalosti zamestnancov riadiacich zmenu	Zmeny nie sú efektívne riadené, čo spôsobuje zdržania v aktivitách a neefektívnu alokáciu zdrojov.	5	5	25
C	Vnútorne politické konflikty	Konflikty medzi oddeleniami alebo vedením znižujú efektivitu projektu.	7	6	42
D	Neochota k zmene	Odpor zamestnancov voči zmene a novým politikám.	7	7	49
E	Nedostatočné zapojenie zainteresovaných strán	Zainteresované strany sa nezúčastňujú na plánovacích a hodnotiacich stretnutiach.	8	8	64
F	Technická nekompatibilita starších systémov	Zastaralé technológie bránia efektívnemu napojeniu nového nástroja, je potrebné vyvinúť konektor.	3	3	9
G	Riešenie neodpovedá požiadavkám	Nástroj implementovaný v rámci zmeny nemá požadované funkcie, resp. ich použitie je obmedzené.	3	5	15
H	Zavádzané procesy a politiky sú nejasné	Procesy a politiky je možné si vyložiť viacerými spôsobmi, nie sú dostatočne komunikované.	4	8	32
I	Nejasné ciele projektu	Pre zmenu je zvolených priveľa dátových prvkov a aktív, implementácia sa spomaľuje na operatívne, aktivity majú oneskorenie a ich finalizácia nie je prioritná	9	5	45

3.6.2. Mapa rizík

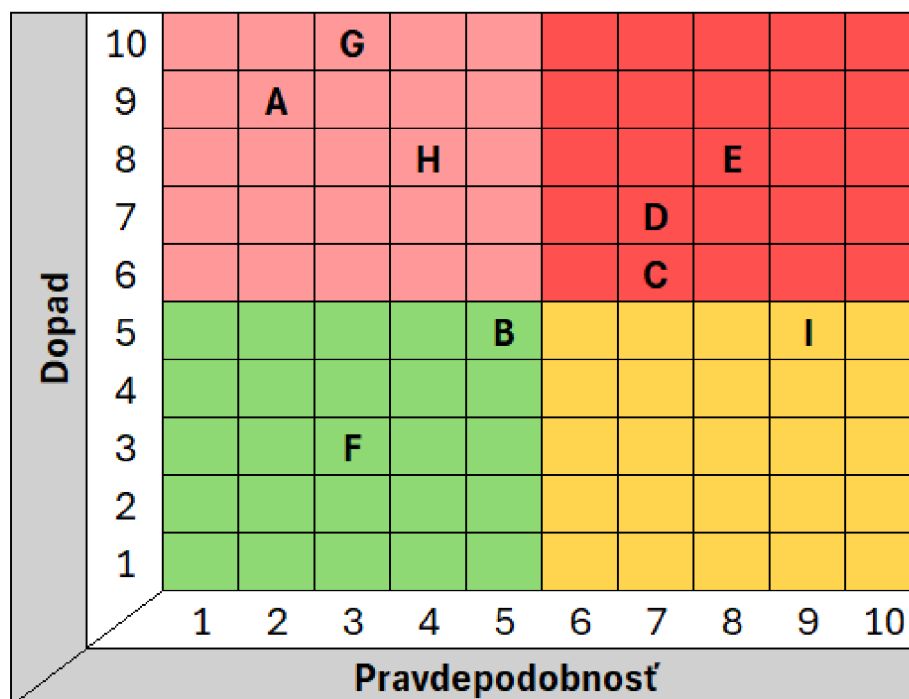
Mapa rizík zobrazuje riziká uvedené v predchádzajúcej tabuľke. Mapa je rozdelená do štyroch kvadrantov:

Červený kvadrant je kvadrant kritických rizík, ktorému je potrebné venovať maximálnu pozornosť a pre tieto riziká je potrebné vytvoriť také opatrenia, ktoré by znížili pravdepodobnosť aj hodnotu dopadu.

Ružový kvadrant je kvadrantom významných rizík, ktoré majú malú pravdepodobnosť, ale veľkú hodnotu dopadu. U týchto rizík je potrebné nájsť také opatrenia, ktoré znížia hodnotu dopadu.

Oranžový kvadrant je kvadrantom bežných rizík s veľkou pravdepodobnosťou, ale s nízkou hodnotou dopadu. U týchto rizík je možné znížiť pravdepodobnosť alebo je možné ich akceptovať.

Zelený kvadrant je kvadrant bezvýznamných rizík. Obsahuje hrozby s malou pravdepodobnosťou a s nízkou hodnotou dopadu. Tieto riziká môžeme buď zmierniť alebo akceptovať.



Obrázok č. 19: Mapa rizík pred zavedením opatrení
(Zdroj: Vlastné spracovanie)

3.6.3. Návrh opatrení

V nasledujúcej tabuľke sú pre identifikované riziká navrhnuté opatrenia. Dôraz je kladený na riziká v červenom a ružovom kvadrante.

Tabuľka č. 16: Návrh opatrení

(Zdroj: Vlastné spracovanie)

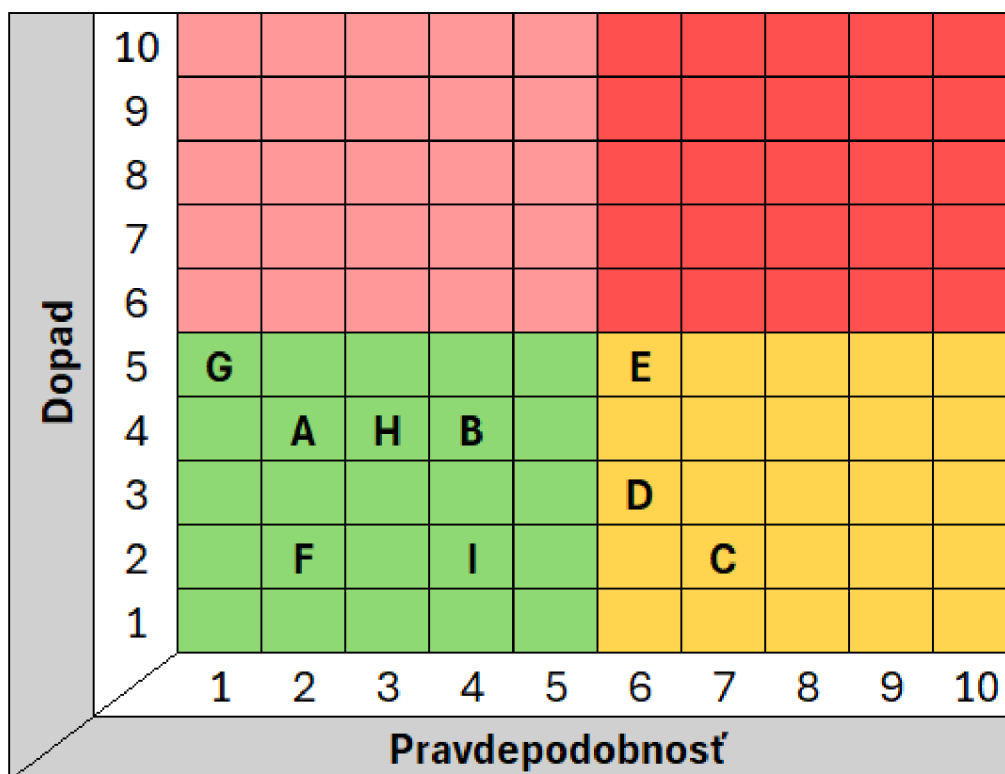
#	Hrozba	Opatrenie	P-st	Dopad	Hodnota rizika
A	Výpoveď kľúčového zamestnanca	Zavedenie procesu nástupníctva a školenia pre kľúčové pozície.	2	4	8
B	Nedostatočné znalosti zamestnancov riadiacich zmenu	Posilnenie interných školení a mentoringových programov pre zamestnancov riadiacich zmenu.	4	4	16
C	Vnútročné politické konflikty	Implementácia pravidelných stretnutí a team-buildingových aktivít na zlepšenie spolupráce a na základe odhadovaného trvania činností ich naplánovanie v dostatočnom predstihu.	7	2	14
D	Neochota k zmene	Zvýšenie transparentnosti rozhodovacích procesov a zavedenie benefitov pre zapojené strany. Zavedenie školiacich programov pre zamestnancov.	6	3	18
E	Nedostatočné zapojenie zainteresovaných strán	Intenzívnejšia komunikácia s hlavnými zainteresovanými stranami a pravidelné informovanie o pokroku projektu. Zavedenie motivačného systému.	6	5	30
F	Technická nekompatibilita starších systémov	Zameranie sa na 5 až 10% najdôležitejších a najpoužívanejších dátových prvkov aktív a po úspešnom zavedení prejsť na menej dôležité prvky. Vyvinutie konektora alebo adaptérov pre integráciu starších systémov.	2	2	4
G	Riešenie neodpovedá požiadavkám	Revízia a úprava požiadaviek s možnosťou testovania prototypu pred nasadením.	1	5	5
H	Zavádzané procesy a politiky sú nejasné	Rozvoj jasne definovaných štandardov pre procesy a politiky podľa best practices a nadnárodných noriem. Zavedenie štandardizovaných šablón a jasnejších riadiacich dokumentov na zabezpečenie konzistencie.	3	4	12
I	Nejasné ciele projektu	Precízne definovanie a obmedzenie počtu cieľov na základe prioritizácie a dôležitosti pre podnik. Implementácia KPI sledujúcich postup a pravidelného informovania o dosiahnutých cieľoch a ďalších krokoch.	4	2	8

3.6.4. Riziká po opatreniach

Nasledujúca tabuľka zobrazuje mapu rizík po navrhnutých opatreniach. Najväčšiu hodnotu rizika malo riziko E a teda nedostatočné zapojenie zainteresovaných strán. Toto riziko bolo identifikované ako najvýznamnejšie aj z toho dôvodu, že celá zmena je veľmi závislá na jej kladnom prijatí zamestnancami, ako bolo spomenuté v ostatných kapitolách. Zavedením opatrenia sa dopad aj pravdepodobnosť znížili a hodnota rizika sa presunula do kvadrantu bežných rizík a teda je možné riziko na tejto hodnote akceptovať.

Druhým najvýznamnejším rizikom bolo riziko D, ktoré opäť súvisí so zamestnancami spoločnosti. Ako je možné vidieť, aj toto riziko bolo po zavedení opatrení možné ohodnotiť tak, že sa presunulo do kvadrantu bežných rizík a je možné ho na súčasnej hladine akceptovať.

V tabuľke je ďalej možné vidieť, že sme eliminovali pravdepodobnosti a hodnotu aj ostatných rizík tak, že sa riziká presunuli do kvadrantu bežných alebo bezvýznamných rizík. V oboch prípadoch je možné riziká akceptovať.



Obrázok č. 20: Mapa rizík po zavedení opatrení
(Zdroj: Vlastné spracovanie)

3.7. Ekonomické zhodnotenie

Táto kapitola a príslušné podkapitoly hodnotia navrhnuté riešenie z ekonomického hľadiska. V úvode je zobrazený prehľad hodnotení vykonaných pomocou ROI a ROSI. Jednotlivé podkapitoly poskytujú detailné informácie pre vstupné hodnoty, pričom prvá podkapitola sa venuje nákladom, druhá ekonomickým benefítom pre ROI a tretia podkapitola poskytuje detail pre výpočet ROSI.

ROI je predmetom **tabuľky č. 18**, ktorá prezentuje dve rozličné výpočty, pričom každý z nich využíva odlišné metodologické prístupy a časové horizonty. V prvom prípade sú zahrnuté všetky náklady vrátane hodnoty práce v ôsmych doménach (tabuľka č. 20), čím je dosiahnuté ROI 16%. Tento výpočet zároveň zohľadňuje krátke obdobie, ako je uvedené v kapitole 3.7.2. V druhom prípade je nákladová položka hodnoty práce vynechaná a je zohľadnené dlhé obdobie, čo výrazne zvyšuje výsledné ROI na 464%.

Tabuľka č. 17: Prehľad návratnosti pomocou ROI v %
(Zdroj: Vlastné spracovanie)

Položka	Prístup č. 1	Prístup č. 2
Náklady	13 800 524 Kč	3 904 420 Kč
Benefity	16 054 944 Kč	22 012 802 Kč
ROI	16%	464%

ROSI je predmetom **obrázku č. 19**, ktorý je uvedený nižšie a ktorý poskytuje prehľad pre výpočet, pričom sú zohľadnené rôzne percentuálne pravdepodobnosti realizácie rizika, úrovne zníženia nebezpečia dopadu rizika a samotnej výšky očakávanej straty. Výška úrovni straty je v zmysle porušenia GDPR. V tretej podkapitole sú poskytnuté aj príklady rozširujúce stratu o % objem tržieb spoločnosti.

SLE	ARO	0,05	0,06	0,07	0,08	0,09	0,10
10 mil. €		-20%	15%	57%	105%	159%	220%
12 mil. €		-4%	38%	88%	146%	211%	284%
14 mil. €		12%	61%	120%	187%	263%	348%
16 mil. €		28%	84%	151%	228%	315%	412%
18 mil. €		44%	107%	182%	269%	367%	476%
20 mil. €		60%	131%	214%	310%	419%	540%
	MR	25%	30%	35%	40%	45%	50%

Obrázok č. 21: Prehľad návratnosti pomocou ROSI
(Zdroj: Vlastné spracovanie)

Ako bolo na začiatku kapitoly spomenuté, bol zobrazený prehľad návratnosti podľa ROI a ROSI a nasledujúce podkapitoly poskytnú detailnejší pohľad pre položky, ktoré do výpočtov vstupovali. V tomto bode je potrebné podotknúť, že na základne poslednej dostupnej výročnej správy spoločnosti z roku 2022 bolo vypočítané, že priemerné ročné mzdové náklady na jedného zamestnanca boli vo výške **1 009 806,57 Kč**. S touto hodnotou je ďalej pracované ako s „hodnotou práce“ v nasledujúcich výpočtoch. Táto informácia je kľúčová najmä pre podkapitolu 3.7.2., ktorá s ňou vo veľkej miere pracuje. S hodnotou sa v určitej miere pracuje aj v nasledujúcej podkapitole, pričom je rozpísaný spôsob výpočtu, ako aj argumenty pre zahrnutie, resp. vynechanie tejto hodnoty pri výpočte ROI.

Alternatívne bolo pre kapitolu 3.7.2. a teda pre výpočet ROI uvažované s obratom na jedného zamestnanca, pričom bol opäť východiskový rok 2022. Bolo vypočítané, že priemerný ročný obrat na jedného zamestnanca bol vo výške 207 862 098,54 Kč. Ak by prínosy navrhovaného riešenia zvýšili tento obrat o 1% u jedného zamestnanca, išlo by o hodnotu 2 078 620,99 Kč. Spoločnosť zamestnáva jedenásť B-1 manažérov a teda ak by vyššia kvalita dát v podobe presnejšieho forecastingu zvýšila presnosť rozhodnutí tak, že by sa zvýšil priemerný obrat pripadajúci na týchto jedenástich zamestnancov o 1%, išlo by o výnos vo výške 22 864 830 Kč, čo je približne hodnota prístupu č. 2 v tabuľke č. 18. Avšak, ako je z predpokladu zjavné, pre veľkú mieru neistoty a nedostatočne uspokojujúceho postupu výpočtu bola ako základ pre výpočty použitá priemerná ročná mzda, ako už bolo spomenuté vyššie.

3.7.1. Náklady

Ako bolo uvedené v predchádzajúcich kapitolách, pre implementáciu návrhu je potrebné zostaviť tím špecialistov a zaistiť potrebné licencie. Prehľad týchto nákladov je možné vidieť v tabuľke nižšie.

Tabuľka č. 18: Ročné náklady na implementáciu návrhu v Kč
(Zdroj: Vlastné spracovanie)

Položka	Náklady
Tím Data Governance	3 029 420 Kč
Licencie	875 000 Kč
Spolu	3 904 420 Kč

Návrhová časť sa vo veľkej miere opiera o návrh zavedenia dátového vlastníctva v rámci spoločnosti, ktorému je venovaná samostatná kapitola 3.4. V odbornej literatúre sa často uvádza, napr. v DAMA (2017), že „the best Data Stewards are often found, not made.“ Toto tvrdenie hovorí o tom, že vo väčšine organizácií už sú ľudia, ktorí sa o dáta starajú aj napriek tomu, že nie je zavedená formálna stránka data governance programu. V analytickej časti v podkapitolách 2.2.1. a 2.4.3. je možné vidieť, že tomuto tvrdeniu odpovedá aj stav v analyzovanej spoločnosti. Formalizovanie ich zodpovednosti za vlastníctvo dát berie na vedomie prácu, ktorú vykonávajú a dovoľuje im byť viac úspešnými a prispievať k správe ešte viac. V tomto ohľade by bolo možné konštatovať, že žiadne ďalšie náklady pre implementáciu návrhu nevznikajú.

Na druhú stranu je možné argumentovať, nakoľko v kapitole 3.4.6. je uvedený odhad FTE jednotlivých rolí, že by bolo vhodné vyčíslit' hodnotu práce jednotlivých rolí. Vyjadrenie tejto hodnoty je možné vidieť v nasledujúcej tabuľke. Je vyjadrená v rozmedzí MIN a MAX, čo odpovedá dolnej a hornej hranici odhadu FTE na konkrétnu rolu. Pre lepšiu orientáciu sú odhady FTE uvedené aj v tejto tabuľke.

Tabuľka č. 19: Hodnota práce jednotlivých rolí v rámci domény v Kč
(Zdroj: Vlastné spracovanie)

Rola	FTE		Hodnota práce	
	MIN	MAX	MIN	MAX
Vlastník domény	0,05	0,1	50 490,33 Kč	100 980,66 Kč
Data Owner	0,1	0,3	100 980,66 Kč	302 941,97 Kč
Business Data Steward	0,3	0,5	302 941,97 Kč	504 903,28 Kč
Technical Data Steward	0,3	0,8	302 941,97 Kč	807 845,26 Kč
Doména (Spolu)	-	-	757 354,93 Kč	1 716 671,17 Kč

Z tabuľky je možné vyčítať, že minimálna hodnota práce venovanej na jednu doménu je vo výške približne 760 tisíc Kč a maximálna vo výške 1,7 milióna Kč ročne. Priemerná hodnota na doménu je približne **1,24 milióna Kč**. Táto priemerná hodnota bola použitá ako vstup pre výpočet v ROI v úvode kapitoly ekonomického zhodnotenia. Pri výpočte, ktorý túto hodnotu zohľadňuje, sa uvažuje s ôsmymi doménami, ako bolo uvedené v podkapitole 3.4.1 a s minimálnou obsadenosťou rolí uvedenou v tabuľke č. 8 v podkapitole 3.2.4.

3.7.2. Benefits

Ako tvrdí Smallwood (2020, str. 21), dobre zavedená data governance pomáha vyhnúť sa negatívnym efektom nesprávne udržiavaných dát a zabezpečuje, že reporty, analýzy a rozhodnutia sú založené na kvalitných a spoľahlivých dátach. Tento efekt sa očakáva po zavedení návrhu a teda používatelia budú:

- môcť získať prístup k prehľadu metadát bez ohľadu na pozíciu a oddelenie,
- môcť robiť presnejšie rozhodnutia s menším množstvom kompromisov a aproximácií,
- schopní spájať obchodné terminológie a tok výpočtovej logiky,
- môcť organizovať a triediť dáta efektívnejšie,
- schopní vykonávať spoľahlivejšie analýzy a vytvárať presnejšie modely (BI, dashboards atď.) a analyzovať dopad svojich zmien dát (dopadová analýza),
- schopní pochopiť cestu dát s minimálnym zasahovaním systémových správcov, čo im umožní rýchlejšie spracovať dáta,
- môcť ušetriť čas pri manuálnom zosúladení s dátami nízkej kvality.

Prínos vyššie uvedených benefitov je zjavný, pretože každý manažér rozumie tomu, že mať prístup k dátam vysokej kvality a teda aj vyššiu dôveru v nich je lepšie, než mať obmedzený prístup k dátam nízkej kvality. Avšak vyjadriť ich monetárnu hodnotu je podstatne ťažšie, pretože je zložitá odhadnúť dopad napr. zavedenia dátových kontrol nad CRM systémom D3soft, na ktorom v súčasnosti žiadne kontroly nie sú. Z tohto dôvodu bude peňažná hodnota týchto dopadov, podobne ako tomu je v predchádzajúcej kapitole, vyjadrená najmä ako hodnota práce a ako vyhnutie sa nákladom. Peňažné zhodnotenie bude demonštrované na troch konkrétnych prípadoch, ktoré odpovedajú trom problémovým oblastiam, ktoré vzišli z výsledku analýzy v kapitole 2.4.

Tabuľka č. 20: Hodnota investovaných prostriedkov a benefitov
(Zdroj: Vlastné spracovanie)

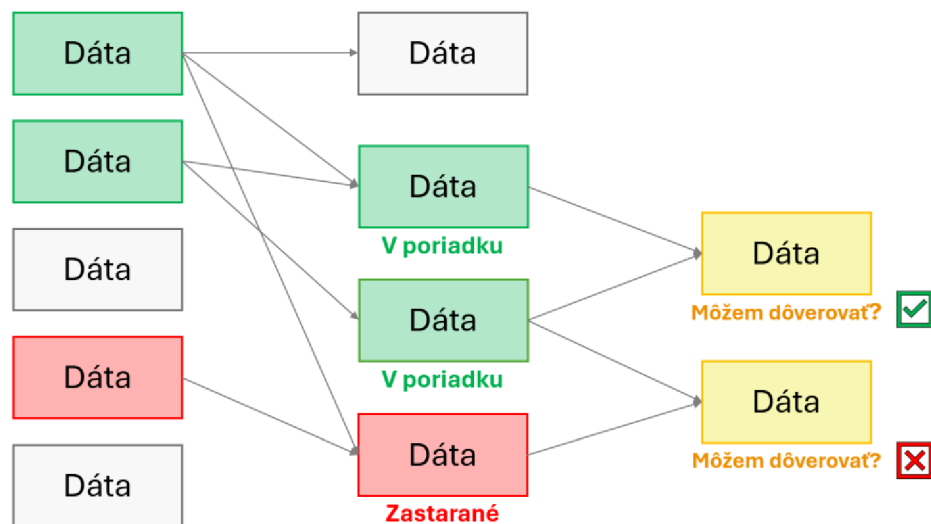
Benefit	Hodnota
Scenár použitia: Data Governance Framework	100 000 Kč
Scenár použitia: Data Quality	17 873 576 Kč
Scenár použitia: Metadáta	4 039 226 Kč
Celkom	22 012 802 Kč

Scenár použitia: Data Governance Framework

Prvý scenár sa týka už viac spomenutých zákazníckych údajov, pretože ak sa napríklad po e-mailovej kampani zistí, že kontakt je neplatný, nie je v DWH takto označený. V roku 2023 bolo takto nedoručených približne 300 000 ks direct marketing emailov, nakoľko za ich označenie nie je nikto zodpovedný. Po zavedení vlastníctva a príslušných politík by bol za tento nedostatok zodpovedný konkrétny zamestnanec, ktorý by mal povinnosť nastaviť nápravné opatrenia, aby tieto nesprávne emailové adresy boli označované a aby sa do kampaní v budúcnosti nedostali. Vyhnutie sa nákladom je možné vyčíslieť na približne 0,2 Kč na email + práca spojená so zaradením kontaktu do výberu a teda vo výške približne **100 000 Kč ročne**. V tomto bode by bolo možné uvažovať aj s potenciálnym predajom, ale pre rôznorodosť kampaní a produktov je výpočet príliš variabilný a teda sa s týmto variantom neuvažuje.

Scenár použitia: Data Quality

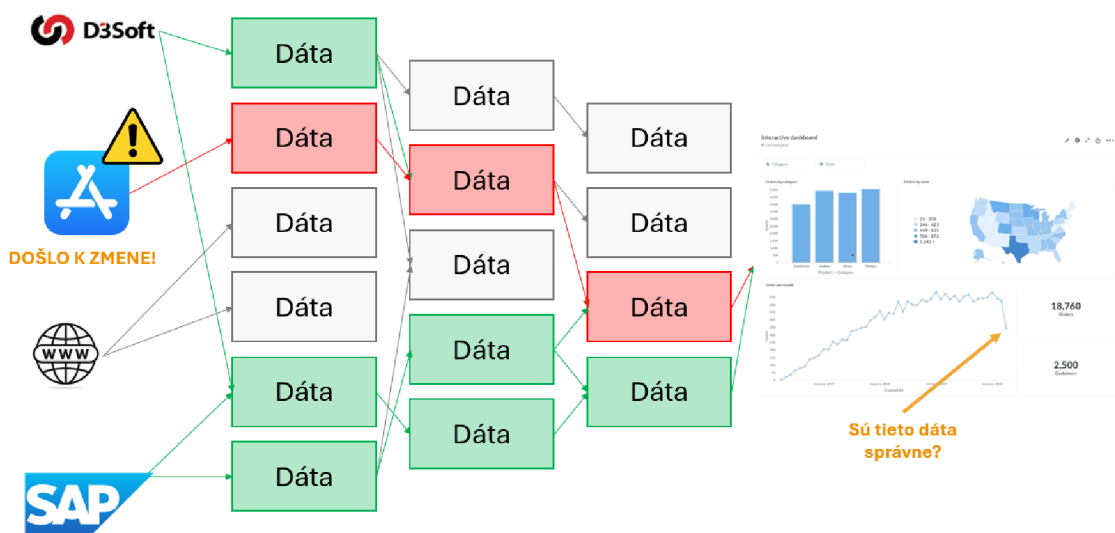
Zamestnanci v oblasti financií trávajú nadpolovičnú väčšinu času na manuálnych úpravách dát. Zároveň, za druhý najväčší problém (po nejasej prioritizácii) považujú nevenovanie dostatočnej pozornosti dátam kvalite a digitalizácii. Cieľom zavedených opatrení je znížiť toto percento o 20 p. b. v krátkom období a o 30 p. b. v dlhom období. Zlepšenie by znamenalo, že zamestnanci môžu venovať svoj čas hodnotnejším aktivitám, než je kontrola vstupných dát a ich úprava a teda v krátkom období by išlo o hodnotu **12 miliónov Kč** a v dlhom období o približne **18 miliónov Kč ročne**.



Obrázok č. 22: Príklad zabezpečenia spoľahlivosti dát prostredníctvom data lineage
(Zdroj: Vlastné spracovanie podľa Boisredon, 2023)

Scenár použitia: Metadáta

K zhodnoteniu implementácii nástrojov používaných pre vyhľadávanie dát, ako dátového katalógu, katalógu reportov alebo biznisového glosára je možné pristúpiť podobne, ako v predchádzajúcom scenári a teda že zamestnanci svoj čas využijú pre hodnotnejšiu prácu. Podľa skúmaní až 56% zamestnancov strávi jednu až tri hodiny denne hľadaním informácií potrebných pre ich prácu (Coveo, 2023). Ak by sme z 56% celkového počtu zamestnancov odpočítali tých z predchádzajúceho scenára, išlo by o 94 zamestnancov. Pre konzervatívny odhad bude použitý počet 80 zamestnancov a bude predpoklad, že trávajú 60 minút denne hľadaním dát. Znížením tohto času o 40% implementáciou nástrojov by bola získaná hodnota práce vo výške **4 miliónov Kč ročne**.



Obrázok č. 23: Príklad hľadania problému/analýzy dopadu prostredníctvom manažmentu metadát
(Zdroj: Vlastné spracovanie podľa Boisredon, 2023)

3.7.3. Zhodnotenie z hľadiska bezpečnosti

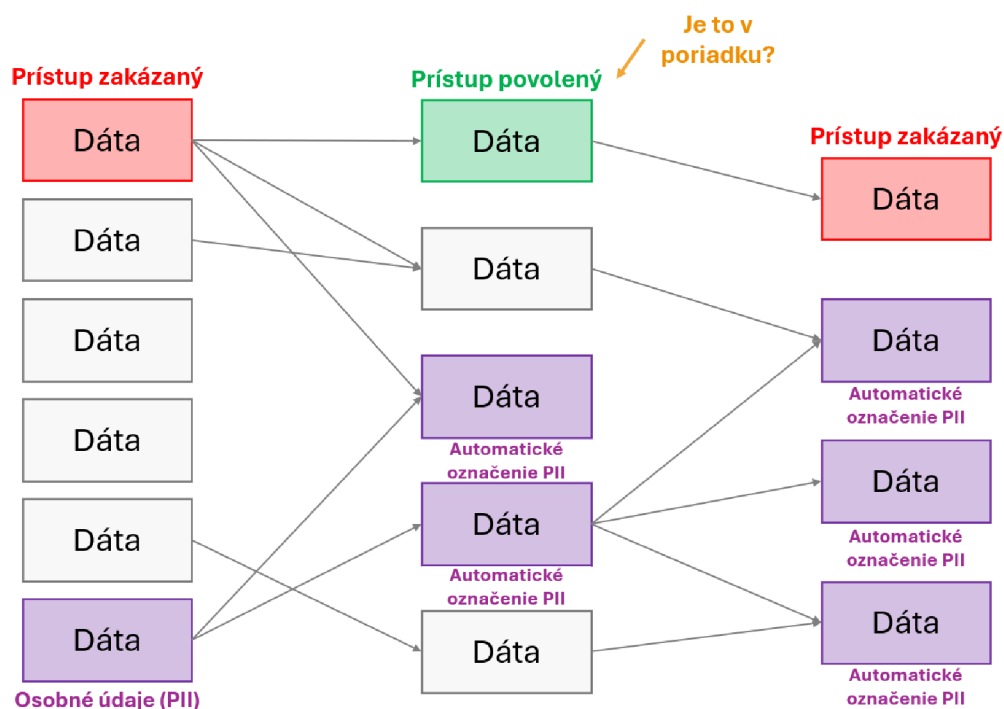
Nové zodpovednosti benefity, ktoré tento prístup prináša, sú:

- Zmapované dátové prostredie spoločnosti. Identifikované dátové aktíva, ich zhodnotenie podľa dôležitosti. Dátové aktíva majú svojich vlastníkov.
- Označenie citlivosti dátových aktív s granularitou až na konkrétne tabuľky a stĺpce. Pomoc pri lokalizácii osobných dát v spoločnosti.
- Návrh kontrolného, bezpečnostného a technologického systému a aktivít s ním spojených.
- Ohodnotenie rizík, definícia dátovo orientovaných kontrol a manažment rizík.

- Podpora manažmentu prístupov a požiadaviek na bezpečnosť.
- Identifikácia zainteresovaných strán a zabezpečenie ich práv a zodpovedností.

Podľa štúdie IBM (2023) boli v roku 2023 najčastejšie zneužitým typom dát osobné údaje zákazníkov. 52 % všetkých porušení sa týkalo nejakej formy osobných údajov zákazníka. Druhým najčastejšie ohrozeným typom údajov boli PII zamestnancov, a to v 40 %.

V Českej republike sa uplatňujú pokuty za porušenie ochrany osobných údajov podľa Zákona o spracúvaní osobných údajov (zákon č. 110/2019 Sb.), ktorý reflektuje a implementuje predpisy GDPR. Tento zákon umožňuje Úradu pro ochranu osobních údajů ukladať pokuty a iné sankcie za nedodržiavanie pravidiel ochrany dát. Ten môže udeliť pokuty až do výšky 10 mil. € alebo 2 % z celosvetového obratu spoločnosti za menej vážne porušenia, a až do výšky 20 mil. € alebo 4 % z celosvetového obratu spoločnosti za vážnejšie porušenia. V Česku bola spoločnosť Avast na začiatku tohto roku pokutovaná vo výške približne 15,4 mil. € za nesprávnu manipuláciu s citlivými údajmi svojich zákazníkov. Data Governance pomáha prechádzať riziku výskytu týchto udalostí.



Obrázok č. 24: Príklad podpory manažmentu prístupu a označovania PII pomocou metadát
(Zdroj: Vlastné spracovanie podľa Boisredon, 2023)

Návratnosť investície do programu je potom možné vypočítať pomocou ROSI modelu. Nasledujúce tabuľky obsahujú štyri modelové príklady a príslušnú návratnosť podľa ROSI.

Hodnota 1% z obratu spoločnosti bola v roku 2022 vo výške **23 637 742 Kč**. Je s ňou uvažované v kontexte stratených predajov a zákazníkov, ktorý vypovedajú zmluvy z dôvodu straty goodwillu. V tabuľkách, ktoré berú v úvahu výšku pokuty + zníženie 1% obratu spoločnosti je možné vidieť, že aj pri hodnote ARO vo výške 0,05 a MR 10% je návratnosť tejto investície v zmysle bezpečnosti pozitívna.

Dolná hranica pokuty						
MR/ARO	0,05	0,06	0,07	0,08	0,09	0,10
10%	-68%	-62%	-55%	-49%	-42%	-36%
15%	-52%	-42%	-33%	-23%	-14%	-4%
20%	-36%	-23%	-10%	2%	15%	28%
25%	-20%	-4%	12%	28%	44%	60%
30%	-4%	15%	34%	54%	73%	92%
35%	12%	34%	57%	79%	102%	124%
40%	28%	54%	79%	105%	131%	156%
45%	44%	73%	102%	131%	159%	188%
50%	60%	92%	124%	156%	188%	220%

Horná hranica pokuty						
MR/ARO	0,05	0,06	0,07	0,08	0,09	0,10
10%	-36%	-23%	-10%	2%	15%	28%
15%	-4%	15%	34%	54%	73%	92%
20%	28%	54%	79%	105%	131%	156%
25%	60%	92%	124%	156%	188%	220%
30%	92%	131%	169%	207%	246%	284%
35%	124%	169%	214%	259%	303%	348%
40%	156%	207%	259%	310%	361%	412%
45%	188%	246%	303%	361%	419%	476%
50%	220%	284%	348%	412%	476%	540%

Dolná hranica pokuty + 1% z obratu spoločnosti						
MR/ARO	0,05	0,06	0,07	0,08	0,09	0,10
10%	8%	29%	51%	72%	94%	115%
15%	62%	94%	126%	158%	191%	223%
20%	115%	158%	202%	245%	288%	331%
25%	169%	223%	277%	331%	385%	438%
30%	223%	288%	352%	417%	482%	546%
35%	277%	352%	428%	503%	578%	654%
40%	331%	417%	503%	589%	675%	762%
45%	385%	482%	578%	675%	772%	869%
50%	438%	546%	654%	762%	869%	977%

Horná hranica pokuty + 1% z obratu spoločnosti						
MR/ARO	0,05	0,06	0,07	0,08	0,09	0,10
10%	40%	68%	96%	124%	151%	179%
15%	110%	151%	193%	235%	277%	319%
20%	179%	235%	291%	347%	403%	459%
25%	249%	319%	389%	459%	529%	599%
30%	319%	403%	487%	571%	654%	738%
35%	389%	487%	585%	682%	780%	878%
40%	459%	571%	682%	794%	906%	1018%
45%	529%	654%	780%	906%	1032%	1157%
50%	599%	738%	878%	1018%	1157%	1297%

Obrázok č. 25: Príklady výpočtu ROSI pri rôznych mierach SLE, ARO a MR
(Zdroj: Vlastné spracovanie)

3.8. Zhrnutie

Návrhová časť vychádzala z analýz vykonaných v predchádzajúcej časti diplomovej práce, kde bola úroveň vyspelosti spoločnosti z pohľady správy dát vyhodnotená na úrovni dva z päť. Samotný návrh bol zameraný na oblasti, u ktorých bol zistený najväčší potenciál pre zlepšenie. V úvode návrhovej časti bola najprv opísaná navrhovaná zmena, ku ktorej má z pohľadu správy dát nastať. Nakoľko ide o zmenu, nadväzujúcou témou bolo jej pojmávanie pomocou Lewinového modelu, v rámci ktorého bola vykonaná analýza silového poľa, boli identifikovaný sponzor, agent a obhajca zmeny, ako aj intervenčné oblasti. Ďalej boli v rámci fázy prechodu a aplikácie zmeny zoradené jednotlivé činnosti, ktoré je pre dosiahnutie zmeny potrebné vykonať. Nakoniec bol vo fáze zmrazenia opísaný požadovaný výsledný stav a zároveň boli definované metriky, ktorými je možné verifikovať dosiahnutie výsledkov nielen na konci projektu, ale aj počas neho.

Činnostiam, ktoré sú predmetom druhej fázy v Lewinovom modeli, bola venovaná samostatná kapitola a podkapitoly, v ktorých sú komplexne opísané z pohľadu ich významu a implementácie. Pre rozsiahlosť činnosti "identifikácie dátových vlastníkov" bola tejto činnosti venovaná samostatná kapitola a podkapitoly, v ktorých boli určené kompetencie, práva a povinnosti dátových vlastníkov a ich zaradenie v organizačnej štruktúre.

Pretože bola riadená zmena pojatá projektovo, bola vypracovaná časová analýza pomocou metódy PERT vrátane identifikácie kritickej cesty. Zároveň bola vykonaná analýza rizík, ktoré súvisia s vykonaním zmeny a boli navrhnuté metódy, ako tieto rizikové miesta prekonať, prípadne ako znížiť riziko na únosnú mieru.

Posledná kapitola návrhovej časti bola venovaná ekonomickému zhodnoteniu implementácie návrhu. Kapitola komplexne rozoberá pridanú hodnotu implementácie pomocou ROI a ROSI, vrátane konkrétnych príkladov, ktoré vyplynuli z analytickej časti.

ZÁVER

Táto práca si v úvode ako cieľ vytýčila zhodnotiť súčasný stav správy a riadenia dát vo vybranej spoločnosti a využitím poznatkov zo zavedených postupov následne zaviesť systém riadenia dát, ktorý povedie k zlepšeniu súčasnej situácie. Pre úspešné dosiahnutie cieľa boli stanovené štyri čiastkové ciele, ktoré práca dosiahla v rámci troch častí.

V prvej teoretickej časti bolo rešeršou odbornej literatúry vypracované teoretické a odborné zázemie, o ktoré sa opera analytická aj návrhová časť. Časť predstavuje hlavné koncepty riešenej problematiky, akými sú všeobecne uznávaný rámec DAMA a analytické postupy pre zhodnotenie súčasnej úrovne správy dát a pre povahu práce obsahuje aj základy manažmentu zmien, riadenia rizík a spôsoby pre ekonomické zhodnotenie zavedenia návrhu.

V nasledujúcej druhej časti bolo cieľom analyzovať súčasný stav správy a riadenia dát vo vybranej spoločnosti. Zároveň bol na základe osobnej skúsenosti vyslovený predpoklad, že budú v rámci analýzy zistené rozličné medzery, ktoré bude potrebné kategorizovať a prioritizovať. Súčasný stav spoločnosti bol skúmaný na operatívnej úrovni pomocou analýzy medzier a na taktickej a strategickej úrovni pomocou Maturity Assessmentu. Medzery boli zoskupené do skupín podľa konkrétnych oblastí, ktorých sa týkali, a to v rámci znalostných oblastí DAMA a oblastí PPTG. Maturity Assessmentu bol vykonaný rovnako pomocou rámca DAMA a PPTG. Výsledky týchto skúmaní boli skombinované a na ich základe boli identifikované oblasti s najväčším potenciálom pre zlepšenie.

V poslednej tretej časti boli využité poznatky získané z teoretickej a analytickej časti pre zavedenie systému riadenia dát, ktorý reflektuje špecifické požiadavky spoločnosti a zároveň je v súlade s uznávanými rámcami a najlepšimi postupmi. Pre bezproblémové zavedenie systému je súčasťou návrhu projektové plánovanie pomocou metódy PERT, vrátane Lewinovho modelu a analýzy rizík. Na záver kapitoly sú zhrnuté prínosy, ktoré z návrhu plynú a zároveň je návrh ekonomicky zhodnotený pomocou ROI a ROSI.

Na základe stručného zhrnutia možno teda konštatovať, že práca úspešne dosiahla svoj cieľ a zaviedla vo vybranej spoločnosti systém riadenia dát s využitím súboru procesov a oblastí znalostí DAMA, ktoré sú považované za najlepšie postupy v oblasti správy dát.

ZOZNAM POUŽITÝCH SKRATIEK

Skratka	Spojenie v angličtine	Spojenie v slovenčine
BI	Business Intelligence	Obchodné spravodajstvo
BOBJ	SAP Business Objects	(neprekladá sa)
CDE	Critical Data Elements	Kritické dátové prvky
CDO	Chief Data Officer	Dátový/á riaditeľ/ka
CFO	Chief Financial Officer	Finančný/á riaditeľ/ka
CISO	Chief Information Security Officer	Riaditeľ/ka pre informačnú bezpečnosť
CMS	Content Management System	System na správu obsahu
CRM	Customer Relationship Management	Riadenie vzťahov so zákazníkmi
DG	Data Governance	Správa dát
DGSC	Data Governance Steering Committee	Riadiaci výbor pre správu dát
DMBOK2	Data Management Body of Knowledge 2	(neprekladá sa)
DO	Data Owner	Dátový vlastník
DPO	Data Protection Officer	Vedúci/a pre ochranu dát
DS	Data Steward	Dátová správcu
DWH	Data Warehouse	Dátový sklad
E2E	End-to-End	(neprekladá sa)
FTE	Full-Time Equivalent	Ekvivalent plného pracovného úväzku
KPI	Key Performance Indicator	Kľúčový výkonnostný ukazovateľ
LoB	Line of Business	Obchodná línia
PERT	Program Evaluation and Review Technique	(neprekladá sa)
ROI	Return on Investment	Návratnosť investícií
ROSI	Return on Security Investment	Návratnosť investícií do bezpečnosti
SQL	Structured Query Language	(neprekladá sa)
TDS	Technical Data Stewards	Technickí dátový správcu

REGISTER OBECNE POUŽÍVANÝCH POJMOV

Adaptér	(Connector) Zariadenie alebo softvér, ktorý umožňuje kompatibilitu medzi rôznymi systémami alebo zariadeniami.
Ad-hoc	Analytické alebo operatívne úlohy vykonávané bez predchádzajúceho plánovania, často ako reakcia na okamžitú potrebu.
Agent zmeny	Osoba alebo skupina zodpovedná za riadenie a implementáciu zmien v organizácii.
Aktívum	Všetok hmotný a nehmotný majetok spoločnosti. Je to ekonomický zdroj, ktorý možno vlastniť alebo kontrolovať a ktorý má alebo vytvára hodnotu.
Analýza medzier	Identifikácia rozdielov medzi aktuálnym stavom a požadovanými/referenčným stavom
Analýza rizík	Proces identifikácie, hodnotenia a prioritizácie rizík spojených s konkrétnymi činnosťami alebo projektmi.
Audít	Systematické zhodnotenie a preskúmanie nezávislou kvalifikovanou osobou.
Brainstorming	Kreatívna technika na generovanie nápadov skupinou ľudí s cieľom riešiť špecifický problém alebo nájsť inovatívne riešenia.
Business Glossary	(Biznisový glosár) Nástroj na definovanie a sprístupňovanie podnikových pojmov a terminológie, znižujúci riziko nesprávneho využívania dát.
Business Intelligence	Procesy a technológie na analýzu dát, ktoré podporujú lepšie podnikové rozhodovanie.
Business Unit	Oddelenie alebo jednotka v organizácii.
CMS	(Content Management System) Systém na správu obsahu, ktorý umožňuje efektívne zhromažďovanie, organizovanie a vyhľadávanie obsahu.
CRM	(Customer Relationship Management) Systém na správu vzťahov so zákazníkmi, ktorý centralizuje dáta o zákazníkoch pre lepšiu službu a obchodné rozhodnutia.
Dashboard	Vizualizačný nástroj, ktorý poskytuje súhrnné informácie a metriky v grafickom formáte na jednom mieste.
Data Catalog	(Dátový katalóg) Organizovaný inventár dát poskytujúci informácie o dátových zdrojoch, ich metadátach a vlastníctve.

Data Discovery	Proces vyhľadávania a objavovania dát v rámci organizácie, ktorý umožňuje lepšie porozumenie a využívanie dát.
Data Governance	Proces, ktorý stanovuje politiky, postupy a štandardy na správu dát s cieľom zabezpečiť ich kvalitu, dostupnosť a bezpečnosť.
Data Lineage	(Dátová línia) Mapovanie pôvodu, cesty a použitia dát, ktoré umožňuje lepšie pochopenie a správu dátových aktív.
Data Management	Zhromažďovanie, organizovanie a sprístupňovanie dát na podporu produktivity, efektívnosti a rozhodovania.
Data Security	Ochrana digitálnych informácií pred neoprávneným prístupom, poškodením alebo krádežou počas ich celého životného cyklu.
Data Steward	Osoba zodpovedná za správu dátových aktív a zabezpečenie ich správneho využitia v rámci organizácie.
Data Domain	Špecifikovaná oblasť alebo kategória dát, ktorá má jednotné pravidlá a štandardy pre ich správu a využitie.
Dátová kultúra	Kultúrne aspekty organizácie, ktoré ovplyvňujú ako sú dáta vnímané, hodnotené a využívané v rámci firmy.
Dátové objekty	Štruktúra na opis dátovej entity zoskupením súboru súvisiacich polí (napríklad „zákazník“).
Dátoví architekti	Profesionáli zodpovední za návrh, štruktúru a správu dátových riešení v organizácii.
Dátový artefakt	Fyzický alebo digitálny objekt, ktorý obsahuje dáta alebo informácie, napríklad dokumenty alebo databázy.
Dátový model	Dátové modely sú vizuálne reprezentácie dátových prvkov podniku a prepojení medzi nimi.
Dátový sklad	(Data Warehouse) Centrálné úložisko integrovaných dát z jedného alebo viacerých zdrojov, navrhnuté tak, aby umožňovali a podporovali činnosti BI (najmä analytiku).
DMBOK2	Rozšírené vydanie 'Data Management Body of Knowledge', ktoré obsahuje pravidlá a najlepšie postupy pre správu dát.
Dopadová analýza	(Impact analysis) Hodnotenie predpokladaných vplyvov zmeny, ktoré pomáha identifikovať možné riziká a náklady.

End-to-End	Proces alebo riešenie pokrývajúce všetky etapy od začiatku po koniec, zvyčajne s dôrazom na úplnosť a integráciu.
FTE	(Full-Time Equivalent) Meranie pracovného úväzku, ktoré predstavuje prácu jedného zamestnanca na plný úväzok.
Goodwill	Hodnota nehmotných aktív firmy, ako sú značka, obchodné vzťahy a reputácia, ktorá prekračuje jej hmotné aktíva.
Granulácia	Detailnosť alebo úroveň rozlíšenia dát alebo informácií.
Hrozba	Potenciálna príčina škody alebo negatívneho vplyvu na systém alebo organizáciu.
Integrácia dát	Proces kombinovania dát z rôznych zdrojov do jedného koherentného celku.
Interoperabilita	Schopnosť rôznych systémov alebo organizácií efektívne a bezproblémovo komunikovať a pracovať spolu.
JIRA	Softvér na správu projektov a sledovanie chýb, ktorý je široko používaný v softvérovom inžinierstve.
Katalogizácia	Proces systematického usporiadania a popisovania informácií alebo dát v katalógu alebo databáze.
Kompetencia	Schopnosť alebo znalosť potrebná na vykonanie konkrétnej úlohy alebo činnosti na uspokojivú úroveň.
Kontextualizované	Dáta alebo informácie, ktoré sú poskytované alebo interpretované v konkrétnom kontexte, čo zvyšuje ich relevanciu alebo zrozumiteľnosť.
Kritická cesta	Najdlhšia postupnosť závislých krokov v projekte, ktoré určujú najkratší možný čas na jeho dokončenie.
Lead	Potenciálny zákazník alebo príležitosť, ktorá môže byť vyvinutá do obchodného vzťahu alebo predaja.
Lewinov model	Teória zmeny navrhnutá Kurtom Lewinom, ktorá opisuje proces zmeny v troch krokoch: rozmrazenie, zmena a zmrazenie.
Maturity Assessment	Hodnotenie, ktoré určuje úroveň vyspelosti organizácie v danej oblasti.
Metadáta	Dáta, ktoré poskytujú informácie o iných dátach, často používané na riadenie, popis alebo kontrolu zdrojov dát.
Metadata Management	Proces správy metadát na zabezpečenie ich správnosti, dostupnosti a relevantnosti.

Onboarding	Proces integrácie nových zamestnancov do firmy alebo získavania nových zákazníkov a ich začlenenia do systému služieb.
Operatíva	Denné činnosti potrebné na vedenie organizácie a dosahovanie jej cieľov.
Pasívny používateľ dát	Osoba, ktorá prijíma a používa dáta, ale nezúčastňuje sa ich aktívnej tvorby alebo spracovania.
PERT	Metóda na plánovanie a kontrolu projektov, ktorá identifikuje kroky potrebné na dokončenie projektu a odhaduje ich trvanie.
Pravidlá párovania	Nastavenia alebo smernice používané na spájanie alebo porovnávanie dátových záznamov z rôznych zdrojov.
Profilovanie dát	Analýza existujúcich dát na identifikáciu vzorcov, anomálií a príležitostí na zlepšenie kvality dát.
Projekt	Súbor činností s konkrétnymi cieľmi, ktoré majú byť dosiahnuté do stanoveného termínu a rozpočtu.
Procesná disciplína	Štruktúrovaný prístup k riadeniu procesov zameraný na ich efektívnosť a predvídateľnosť. Ide o dodržiavanie premyslených a dobre definovaných procesov, ktoré sa vykonávajú denne.
Rámec	(Framework) Štruktúrovaný systém alebo súbor pravidiel a usmernení, ktoré poskytujú základ alebo štruktúru pre vykonávanie konkrétnych úloh alebo projektov.
Riziko	Miera ohrozenia, že sa uplatní hrozba dôjde k nežiadúcemu výsledku.
ROI	Ukazovateľ ziskovosti, ktorý meria návratnosť investovaných zdrojov.
ROSI	Ukazovateľ efektívnosti investícií do bezpečnostných opatrení vzhľadom na zníženie rizík a strat.
Skript	Preddefinovaný súbor príkazov, ktorý sa používa na automatizáciu úloh v softvérových alebo databázových systémoch.
Sponzor	Osoba alebo skupina, ktorá poskytuje podporu a zdroje pre projekt alebo iniciatívu.
SQL príkaz	Príkaz používaný v štruktúrovanom dopytovacom jazyku na manipuláciu alebo dopytovanie dát v databázach.
Stakeholder	Osoba alebo skupina, ktorá má záujem alebo je ovplyvnená výsledkami projektu alebo podnikovej činnosti.

Stratégia	Plán alebo súbor usmernení určených na dosiahnutie dlhodobých alebo celkových cieľov.
Taktika	Konkrétne metódy alebo techniky používané na dosiahnutie stanovených cieľov alebo úloh v rámci širšej stratégie.
Technická inventarizácia	Proces zhromažďovania a analýzy informácií o technických zdrojoch, ako sú hardvér, softvér alebo dáta.
Technokraticky	Prístup alebo systém riadenia, ktorý zdôrazňuje technické znalosti a kompetencie v rozhodovacom procese.
Úroveň C	Vysoká manažérska úroveň v organizácii, zvyčajne obsahujúca pozície ako CEO, CFO, CIO atď.
Kľúčový výkonnostný ukazovateľ	Metrika, ktorá kvantifikuje efektívnosť alebo úspech činností v organizácii pre konkrétny cieľ.
Životný cyklus dát	Fázy, ktorými prechádzajú dáta od ich vytvorenia cez ukladanie, použitie až po konečné vymazanie alebo archiváciu.

ZOZNAM POUŽITÝCH ZDROJOV

AGUERRE a kol., 2024. *GLOBAL DIGITAL DATA GOVERNANCE Polycentric Perspectives*. New York, USA: Routledge. ISBN: 978-1-032-48311-5.

ATACCAMA, 2024. *Augmented data lineage: What it is and why it matters*. Dostupné z: <https://www.ataccama.com/blog/augmented-data-lineage-what-it-is-and-why-it-matters> [cit. 2024-04-22].

BOISREDON, X., 2023. *Data governance*. Dostupné z: <https://www.castordoc.com/blog-category/data-governance> [cit. 2024-05-07].

CHISHOLM, M. and Robey-Lee, D, 2010. *Definitions in Information Management: A Guide to Fundamental Semantic Metadata*. Kanada: Bydesign Media. ISBN 978-0615357546.

COVEO, 2023. *2023 Coveo workplace relevance report*. Dostupné z: <https://www.coveo.com/en/pdf?file=-/media/library/files/reports/2023-relevance-reports/2023-coveo-workplace-relevance-report.pdf> [cit. 2024-04-22].

CUPOLI, Patricia, Earlay, Susan a Henderson, Deborah, 2014. *DAMA-DMBOK Framework*. Dostupné z: <https://www.itmatters.com.au/resources/DAMA-DMBOK2-Framework-V2-20140317-FINAL.pdf> [cit. 2023-11-06].

ČSN EN ISO/IEC 27001 Informační bezpečnost, kybernetická bezpečnost a ochrana soukromí - Systémy managementu informační bezpečnosti - Požadavky, 2023. [Praha]: Česká agentura pro standardizaci.

ČSN EN ISO/IEC 27002 Informační bezpečnost, kybernetická bezpečnost a ochrana soukromí - Opatření informační bezpečnosti, 2023. [Praha]: Česká agentura pro standardizaci.

DAMA International, 2017. *The DAMA guide to the data management body of knowledge*. Bradley Beach, New Jersey, USA: Technics Publications. ISBN 978-09-77140-08-4.

DOUCEK, Petr; KONEČNÝ, Martin a NOVÁK, Luděk, 2019. *Řízení kybernetické bezpečnosti a bezpečnosti informací*. Praha: Professional Publishing. ISBN 978-80-88260-39-4.

ERWIN, 2022. *Erwin DI Suite*. Dostupné z: <https://content.dataversity.net/rs/656-WMW-918/images/erwin-DI-Suite.pdf> [cit. 2023-09-15].

EVANS, N. and Price, J., 2012. 'Barriers to the effective deployment of information assets: An executive management perspective.' *Interdisciplinary Journal of Information, Knowledge, and Management*, Volume 7. 177-199. 10.28945/1721.

ERYUREK a kol, 2021. *Data Governance The Definitive Guide*. Sebastopol, USA: O'reilly Media, S.L. ISBN 978-1-492-06342-1.

HOLT, A., 2021. *DATA GOVERNANCE Governing data for sustainable business*. Swindon, United Kingdom: BCS, The Chartered Institute for IT. ISBN: 978-1-78017-3757.

HOPPER, M.A., 2023. *Practitioner's guide to operationalizing data governance*. New Jersey, USA: John Wiley & Sons. ISBN: 9781119851424.

IBM, 2023. *Data breach report*. Dostupné z: <https://www.ibm.com/reports/data-breach> [cit. 2024-04-01].

Interný dokument, 2023. Jihočeský kraj, XYZ, a. s., 2023.

KAPOUNOVÁ, Jana a Pavel KAPOUN, 2017. *Bakalářská a diplomová práce: od zadání po obhajobu*. Praha: Grada, 134 stran. ISBN: 978-80-271-0079-8.

LADLEY, J., 2012. *Data Governance How to Design, Deploy, and Sustain an Effective Data Governance Program*. Waltham, USA: Morgan Kaufmann Publishers. ISBN: 978-0-12-415829-0.

MAHANTI, R., 2021. *Data governance and data management*. Singapur: Springer Nature Singapore. ISBN: 978-981-1635-85-4.

OLESEN-BAGNEUX, O., 2023. *The enterprise data catalog*. Sebastopol, USA: O'reilly Media, Inc. ISBN: 978-1-492-09871-3.

PLOTKIN, D., 2021. *Data stewardship: An actionable guide to effective data management and data governance*. London, United Kingdom: Academic Press. ISBN: 978-0-12-822132-7.

REICHTENTAL, J., 2022, *Data Governance for dummies*. New Jersey, USA: John Wiley & Sons Inc. ISBN: 978-1-119-90677-3

SEBASTIAN-COLEMAN, L., 2013. *Measuring Data Quality for Ongoing Improvement: A Data Quality Assessment Framework*. Waltham, USA: Morgan Kaufmann Publishers. ISBN: 978-0123970336.

SEDLÁK, P. & Konečný, M., 2021. *Kybernetická (ne)bezpečnost: problematika bezpečnosti v kyberprostoru* Vydání první., Brno: Akademické nakladatelství CERM. ISBN: 978-80-7623-068-2.

SEDLÁK, Petr a KONEČNÝ, Martin, 2023. *Přeměna ISMS v manažerské informatice*. Brno: CERM, akademické nakladatelství. ISBN: 978-80-7623-110-8.

SMALLWOOD, R.F., 2020. *Information governance: Concepts, strategies and best practices*, 2nd ed. New Jersey, USA: John Wiley & Sons. ISBN: 9781119491415.

SMEJKAL, Vladimír a RAIS, Karel, 2013. *Řízení rizik ve firmách a jiných organizacích*. 4., aktualiz. a rozš. vyd. Praha: Grada. ISBN 978-80-247-4644-9.

SOARES, S., 2014. *The Chief Data Officer Handbook for Data Governance*. Boise, USA: MC Press Online, LLC. ISBN: 978-1-58347-417-4.

ŠŤASTNÝ, M. *Data Governance expertíza* [interview]. XYZ. Brno 13. 9. 2023.

VERIZON, 2023. Summary of findings. Dostupné z: <https://www.verizon.com/business/resources/reports/dbir/2023/summary-of-findings/> [cit. 2024-04-12].

VIMEO, 2024. *ADASTRA workshop: Jak na data governance?* Dostupné z: <https://vimeo.com/888706309/26f415a4bc?share=copy> [cit. 2024-04-12].

Výročná správa za rok 2022. Jihočeský kraj, XYZ, a. s., 2023.

ZOZNAM OBRÁZKOV

Obrázok č. 1: DAMA Wheel.....	18
Obrázok č. 2: Data Governance a Data Management	22
Obrázok č. 3: Tabuľka v zdroji a jej viditeľnosť v data catalogu	28
Obrázok č. 4: Rozdiel medzi hľadanie dát a hľadanie v dátach.....	29
Obrázok č. 5: Príklad E2E data lineage systému od zdrojových dát do BI nástrojov.....	30
Obrázok č. 6: Úrovne vyspelosti	32
Obrázok č. 7: PPTG model	33
Obrázok č. 8: Lewinov model riadenej zmeny	34
Obrázok č. 9: Základné kroky procesu modelovania riadenej zmeny	34
Obrázok č. 10: Metódy znižovania rizika	37
Obrázok č. 11: Oblasti s najväčším priestorom pre zlepšenie.....	48
Obrázok č. 12: Zobrazenie PII v dátových aktívach pomocou Data Lineage.....	68
Obrázok č. 13: Proces automatizovaného zberu reportov	71
Obrázok č. 14: Mriežkový prístup k hlavným dátovým doménam	72
Obrázok č. 15: Model dátového vlastníctva v spoločnosti.....	75
Obrázok č. 16: Organizačne zaradenie Data Governance – prvý variant	77
Obrázok č. 17: Organizačne zaradenie Data Governance – druhý variant	78
Obrázok č. 18: Sieťový graf.....	81
Obrázok č. 19: Mapa rizík pred zavedením opatrení	84
Obrázok č. 20: Mapa rizík po zavedení opatrení.....	86
Obrázok č. 21: Prehľad návratnosti pomocou ROSI.....	87
Obrázok č. 22: Príklad zabezpečenia spoľahlivosti dát prostredníctvom data lineage	91
Obrázok č. 23: Príklad hľadania problému/analýzy dopadu prostredníctvom manažmentu metadát	92
Obrázok č. 24: Príklad podpory manažmentu prístupu a označovania PII pomocou metadát.....	93
Obrázok č. 25: Príklady výpočtu ROSI pri rôznych mierach SLE, ARO a MR	94

ZOZNAM TABULIEK

Tabuľka č. 1: Zaradenie medzier v rámci oblastí PPTG	41
Tabuľka č. 2: Zaradenie medzier v rámci znalostných oblastí.....	42
Tabuľka č. 3: Počet tém použitých v modeloch podľa znalostnej oblasti.....	45
Tabuľka č. 4: Výsledné hodnotenie vykonaných analýz.....	49
Tabuľka č. 5: Analýza silového poľa	59
Tabuľka č. 6: Jednotlivé činnosti pre zmenu.....	61
Tabuľka č. 7: Metriky pre meranie dosiahnutých výsledkov	62
Tabuľka č. 8: Polia pre business terms.....	69
Tabuľka č. 9: Polia pre report katalóg.....	70
Tabuľka č. 10: FTE alokácia na role definované v rámci DG	76
Tabuľka č. 11: Legenda PERT	79
Tabuľka č. 12: Časová analýza PERT (Zdroj: Vlastné spracovanie).....	80
Tabuľka č. 13: Legenda pre sieťový graf.	81
Tabuľka č. 14: Riziková politika.....	82
Tabuľka č. 15: Identifikácia a ohodnotenie rizík	83
Tabuľka č. 16: Návrh opatrení	85
Tabuľka č. 17: Prehľad návratnosti pomocou ROI v %.....	87
Tabuľka č. 18: Ročné náklady na implementáciu návrhu v Kč	88
Tabuľka č. 19: Hodnota práce jednotlivých rolí v rámci domény v Kč.....	89
Tabuľka č. 20: Hodnota investovaných prostriedkov a benefitov	90

ZOZNAM GRAFOV

Graf č. 1: Úroveň vyspelosti podľa DMOK	46
Graf č. 2: Úroveň vyspelosti podľa PPTG	47

ZOZNAM PRÍLOH

PRÍLOHA Č. 1: Prehľad oblastí a úrovne vyspelosti pre PPTG	110
PRÍLOHA Č. 2: Prehľad oblastí a úrovne vyspelosti pre DMBOK2	111

PRÍLOHA Č. 1: Prehľad oblastí a úrovne vyspelosti pre PPTG

Oblasť / Úroveň	0 Absencia	1 Počiatočná	2 Opakovateľná	3 Definovaná	4 Riadená	5 Optimalizovaná
Ludia	Žiadne role, žiadny experti.	Na základe ad-hoc údajov s ad-hoc výstupmi.	Neformálni experti na dáta zaručujú opakovateľné výstupy.	Dátové role sú adekvátne definované, ale nie plne priradené. Neformálni odborníci sú stále potrební.	Všetky role sú definované a dátoví experti sú vhodne identifikovaní a aktívni.	Optimalizované priradenie rolí a úloh.
Procesy	Chýbajúce procesy.	Procesy sú unikátne sady úloh explicitne definované pre každý prípad.	Procesy sú neformálne, ale v určitých prípadoch sa môžu opäť použiť.	Procesy sú štruktúrované, ale ich výkonnosť sa môže líšiť.	Procesy sú riadené a kontrolované s garantovanými výstupmi a SLA.	Nepretržitá optimalizácia procesov.
Technológie	Žiadne platformy pre dáta alebo metadáta.	Niektoré dátové platformy existujú s obmedzeným rozsahom.	Hlavné dátové platformy sú zriadené a integrácia údajov prebieha nepretržite.	Všetky platformy pre dáta a metadáta sú zriadené. Ich integrácia je obmedzená.	Všetky platformy pre dáta a metadáta sú zriadené a dobre integrované.	Všetky dátové platformy a platformy metadát sú bezproblémovo integrované a neustále optimalizované.
Správa	Chýbajúce politiky.	Implicitné politiky sú pevne zakódované v aplikáciách.	Existujú nejaké politiky, ale sú izolované.	Politiky sú definované, ale výkonné procedúry ešte nie sú zriadené.	Politiky a procedúry sú plne funkčné.	Neustála optimalizácia politik a postupov.

PRÍLOHA Č. 2: Prehľad oblastí a úrovne vyspelosti pre DMBOK2

#	Oblasť	Úroveň 1	Úroveň 2	Úroveň 3	Úroveň 4	Úroveň 5
0	Data Governance Core	Správa formálne definovaná ako súčasť podnikovej architektúry.	Niektoré procesy správy dát sú opakovateľné.	Dobre definovaná a plne funkčná správa údajov pokrývajúca kľúčové oblasti.	Správa dát je kvantifikovaná a kontrolovaná vo väčšine dátových domén a dátových aktív.	Správa dát optimalizuje všetky dátové aktíva.
1	Data Architecture	Ad-hoc dátová architektúra. Mnoho dočasných ad-hoc riešení bez akejkoľvek stratégie.	Existujú kľúčové dátové platformy a komponenty s určitými špecifickými nedostatkami.	Všetky dátové platformy a komponenty existujú s určitými špecifickými nedostatkami.	Všetky dátové platformy a komponenty existujú a sú primerane spravované bez závažných problémov.	Neustále zlepšovanie dátových platforiem. dáta sú vhodne distribuované a spravované v rámci architektúry.
2	Data Modelling & Design	Niektoré dátové modely existujú, ale nie sú aktualizované. Doménový dátový model existuje v obmedzenom rozsahu. Spoločný dátový model neexistuje.	Niektoré dátové modely existujú a sú primerane udržiavané. Existuje doménový dátový model. Spoločný dátový model neexistuje.	Všetky dátové modely sú definované a udržiavané. Doménový dátový model je rozšírený na spoločný dátový model.	Všetky dátové modely sú definované, udržiavané a spravované spoločne. Spoločný dátový model existuje s čiastočnou integráciou s ostatnými modelmi.	Existuje spoločný dátový model, ktorý integruje všetky dátové modely.
3	Data Storage & Operations	Ad-hoc operácie	Opakovateľné operácie. Základný technologický stack je dostatočne vyspelý na dlhodobé použitie.	Definované operačné procesy vrátane zodpovedností a automatizácie. Dobre štandardizovaná základný technologický stack pokrýva všetky hlavné potreby.	Spravované operácie, vrátane SLA. Komplexný technologický stack pre akékoľvek použitie s potrebnými zručnosťami.	Samo optimalizujúce operácie. Bezproblémový dátový technologický stack.
4	Data Security	Žiadna bezpečnosť	Niektoré dáta sú zabezpečené. Aplikujú sa niektoré bezpečnostné pravidlá.	Všetky bezpečnostné pravidlá sú definované. Dáta sú zabezpečené. User management je operatívny s nedostatkami.	Všetky bezpečnostné pravidlá sú úplne implementované a spravované. Bezpečnostné logy sú integrované a hodnotené.	Proaktívna bezpečnosť vrátane testovania bezpečnosti a neustáleho zlepšovania.

#	Oblasť	Úroveň 1	Úroveň 2	Úroveň 3	Úroveň 4	Úroveň 5
5	Data Integration & Interoperability	Ad-hoc integrácia dát	Opakovateľná integrácia dát bez explicitného vymedzenia hlavného prostredia. Ťažké získať prehľad alebo vykonať analýzu dopadu.	Explicitná integrácia dát. Dátové toky existujú a je možné získať prehľad alebo vykonať analýzu dopadu.	Spravovaná integrácia s plne operatívnu Dátovou kvalitou.	Neustále sa zlepšujúca integrácia dát s CI/CD
6	Document & Content Management	Ad-hoc dokumentácia bez hierarchie a odkazov. Kvalita a platnosť dokumentov sú príliš variabilné.	Aktuálne dokumentácia je povinná pre nové riešenia.	Všetky kritické riešenia a komponenty sú zdokumentované.	Centrálne riadenie dokumentov a obsahu vrátane vyhľadávania a objavovania.	Optimalizovaný centrálny dokument s automatizovanými funkciami.
7	Reference & Master Data	Referenčné dáta existujú, ale nie sú garantované.	Niektoré referenčné dáta sú definované a spravované.	Všetky referenčné dáta sú definované, MDM je ad-hoc.	Všetky súbory RDM & MDM sú spravované.	Plne operatívne RDM & MDM so všetkými prúdmi nahor a nadol.
8	Data Warehousing & Business Intelligence	Ad-hoc vývoj a prevádzka DWH a BI. Reťazec hodnoty dát je nespoľahlivý.	Opakovateľný vývoj, DWH a BI sú správne zriadené. Reporting je rozdelený do väčších skupín s niektorými nezrovnalosťami a neplánovanými prekryvami.	Dobre definovaný vývoj a prevádzka DWH a BI. Dátový hodnotový reťazec je plne funkčný so zvyčajne efektívnym riadením zmien	Dobre definovaný vývoj a prevádzka DWH a BI. Hodnotový reťazec dát je plne funkčný a spravovaný. Riadenie zmien zaručuje rýchlu reakciu na obchodné potreby.	Neustále a automatizované zlepšovanie. Dáta sú neustále pretransformované na hodnotu. Hodnota je optimalizovaná.
9	Metadata	Niektoré metadáta sú opísané, ale bez záruky.	Všetky metadáta sú opísané bez záruky. Manuálne je možné ich prepojiť.	Všetky metadáta sú opísané so zárukou. Môžu byť automaticky integrované. Dôveryhodnosť zlúčenia metadát je sporná.	Všetky metadáta sú opísané so zárukou. Je možné ich automaticky integrovať. Dôveryhodnosť metadát je prijateľná pre prípady použitia, ako je vývoj riadený metadátami s menšími problémami.	Všetky metadáta sú opísané so zárukou. Sú automaticky integrované a verzované. Dôveryhodnosť metadát je nespochybniteľná bez problémov. Pre metadáta existuje DQ.
10	Data Quality	Neexistuje žiadna dátová kvalita, alebo je dátová kvalita reaktívna a riadi sa informovanosťou používateľov dát.	Málo dátových položiek je pravidelne kontrolovaných. Povedomie užívateľov dát je kľúčovým faktorom pre akceptovateľnú kvalitu dát.	Kontroly DQ sú definované pre všetky relevantné dátové položky. DQ má reaktívne a proaktívne časti. Užívatelia hlásia zriedkavé problémy.	DQ kontroly implementované a spravované. DQ podporuje reaktívne a proaktívne režimy. Niektoré dátové položky sú automaticky čistené.	Neustále zlepšovanie kvality vrátane monitorovania a nástrojov DQ riadených umelou inteligenciou a automatizácie.