



# Výstup č. 1.1.7: Analýza toku údajov

Realizačná zmluva o poskytnutí služieb a o dielo č. 445/2022

*Projekt:*  
**Zlepšenie využívania údajov vo verejnej  
správe**

*ITMS kód projektu:*  
**314011S979**

## Document review and approval

### Revision history

Version	Author	Date	Revision
1.0	Dagmar Ceľuchová Bošanská, Juraj Bárdy, Martin Janík	30.06.2023	

### This document has been reviewed by

Reviewer	Date reviewed
1	
2	
3	
4	
5	

### This document has been approved by

Subject matter experts		
Name	Signature	Date reviewed
1		
2		
3		
4		
5		

<b>ZOZNAM SKRATIEK</b>	
<b>Skratka</b>	<b>Význam</b>
API	Aplikačné programovacie rozhranie (Application Programming Interface)
CIP	Centrálne integračná platforma
CMÚ	Centrálny model údajov
CSV	comma-separated values
Databáza	Databáza je súbor vzájomne prepojených údajov, ktoré sa používajú na efektívne vyhľadávanie, vkladanie a mazanie údajov. Používa sa tiež na usporiadanie údajov vo forme tabuľky, schémy, pohľadov a zostáv atď.
DBMS	Systém správy databázy
DI	Dátová integrácia
DP	Dopytový projekt / Dopytové projekty
EDIW	Európska digitálna peňaženka s identitou (European Digital Identity Wallet)
eID	Elektronická identita
eIDAS	Nariadenie Európskej únie č. 910/2014 o elektronickej identifikácii a dôveryhodných službách pre elektronické transakcie na vnútornom európskom trhu.
ELT	Extract, load, transform
EK	Existujúci konzument
EP	Existujúci poskytovateľ
ESB	Enterprise Service Bus
ETL	Extract, transform, load
EU	Európska únia
GDBMS	Správa databázy grafov
GDPR	Všeobecné nariadenie o ochrane osobných údajov (General Data Protection Regulation)
HTML	HyperText Markup Language
iPaaS	Integration Platform as a Service
IS CSRÚ	Informačný systém centrálnej správy referenčných údajov
IS VS	Informačný systém verejnej správy
JSON	JavaScript Object Notation
JSON-LD	JSON pre linkované údaje (JSON for Linking Data)
JWE	JSON Web Encryption
JWS	JSON Web Signature
JWT	JSON Web Token
mID	Mobilná identita
MIRRI SR	Ministerstvo investícií, regionálneho rozvoja a informatizácie

## ZOZNAM SKRATIEK

MOU	Manažment osobných údajov
MV SR	Ministerstvo vnútra SR
NK	Nový konzument
NP	Nový poskytovateľ
OLAP	Online analytické spracovanie (Online analytical processing)
OVM	Orgán verejnej moci
PET	Technológie na zvýšenie súkromia (Privacy-enhancing Technologies)
PIMS	systemy na správu osobných informácií (Personal Information Management System)
RDF	Resource Description Framework
SDG	Jednotná digitálna brána (Single Digital Gateway)
SVG	Scalable Vector Graphics
SvM	Slovensko v mobile
OOTs	Technický systém pre jedenkrát a dosť (Once-Only Technical System)
OVM	Orgán verejnej moci
RDBMS	System na manažment relačných databáz (Relational database management system)
TRUSTS	Trusted Secure Data Sharing Space
URI	Jednotný referencovateľný identifikátor
VC	Overiteľné poverenia (Verifiable Credentials)
W3C	World Wide Web Consortium
XDM	XPath Data Model
XHTML	Extensible HyperText Markup Language
XML	Extensible Markup Language
XSD	XML Schema Definition
XSL	eXtensible Stylesheet Language
XSLT	XSL Transformations

## Obsah

1	Úvod a zhrnutie	7
1.1	Kontext	8
1.1.1	Zhromažďovanie dát prostredníctvom DI	8
1.1.2	Zoznam dátových entít a číselníkov.	9
1.1.3	Údaje obsiahnuté v dopytových výzvach zameraných na manažment údajov inštitúciei verejnej správy	10
1.2	Metodika realizácie výstupu	10
2	Definícia množín poskytovateľov a konzumentov údajov	12
2.1	Nový poskytovateľ údajov	12
2.2	Existujúci poskytovateľ – rozšírené poskytovanie údajov	12
2.3	Nový konzument údajov	12
2.4	Existujúci konzument – rozšírené konzumovanie údajov	13
2.6	Mimo rozsah dokumentu Analýza toku údajov	23
3	Určenie dátových objektov	24
3.1.1	Vysvetlenie významu určenia dátových objektov v analytickom procese.	25
3.1.2	Súčasný kontext: množstvo dostupných dát a potreba účinných nástrojov na ich analýzu	26
3.1.3	Cieľ úvodného postupu: poskytnúť štruktúru a organizáciu pre ďalšie spracovanie a analýzu	27
3.2	Identifikácia cieľov a požiadaviek	28
3.2.1	Stanovenie jasných cieľov pre analýzu dát a ich dosiahnutie	28
3.2.2	Definovanie požadovaných informácií a výstupov z analýzy.	29
3.2.3	Určenie typov dát potrebných na dosiahnutie cieľov	30
3.2.4	Rozlišovanie štruktúrovaných a neštruktúrovaných dát	31
3.3	Zhromaždenie a evaluácia dostupných dát	32
3.3.1	Identifikácia zdrojov dát, vrátane interných, externých a verejne dostupných zdrojov	32
3.3.2	Analýza dostupných dát a ich relevancie pre ciele analýzy	33
3.3.3	Vykonanie auditu dát pre určenie ich kvality, spoľahlivosti a dostupnosti	34
3.3.4	Vyhodnotenie vhodnosti dát na základe ich kvality, relevancie a dostupnosti	35
3.4	Identifikácia dátových objektov	36
3.4.1	Identifikácia hlavných kategórií dátových objektov na základe stanovených cieľov	36
3.4.2	Detailný popis každej kategórie dátových objektov a ich účelov v analýze	37
3.4.3	Rozdelenie každej kategórie na konkrétne atribúty, ktoré budú analyzované	38
3.4.4	Vyhodnotenie dátových objektov na základe ich schopnosti poskytnúť požadované informácie	39
3.5	Vytvorenie schémy dátových objektov	40

3.5.1	Návrh schémy dátových objektov na základe identifikovaných objektov a atribútov	40
3.5.2	Definovanie vzťahov medzi objektmi a ich atribútmi pomocou primárnych a cudzích kľúčov	42
3.5.3	C. Vizualizácia schémy dátových objektov pomocou diagramov a grafických reprezentácií	42
3.6	Implementácia a správa dátových objektov	43
3.6.1	Implementácia navrhutej schémy dátových objektov do vhodného systému alebo databázového riešenia	43
3.6.2	Overenie presnosti a správnosti načítania dát do jednotlivých objektov	45
3.6.3	Správa dátových objektov, vrátane udržiavania aktuálnych údajov, indexovania a zálohovania	46
4	Model komunikačných väzieb	48
4.1	Spojitosť s Dátovou integráciou	49
4.2	Aký to má súvis s Dátovou kanceláriou ?	50
4.3	Prepojenie modelu komunikačných väzieb so znalostnými grafmi	52
5	Meranie komunikačných väzieb	61
5.1	Možné softvérové nástroje pre meranie	61
5.1.1	Prepojenie IS CSRÚ a Neo4j	62
5.1.2	Nástroj Graphlytic	64
5.1.3	Základné komponenty Graphlyticu	66
5.2	Architektúra riešenia	67
5.2.1	Zdroje údajov použité v grafových modeloch	68
5.3	Grafové modely a ich použitie	68
5.3.1	Hlavný grafový model	69
5.3.2	Agendový grafový model – Zdieľanie datasetov medzi OVM	70
5.3.3	Mapovanie grafových vzorov	70
5.4	Práca s aplikáciou Graphlytic	71
5.4.1	Prístup do aplikácie	71
5.4.2	Práca s hlavným modelom	73
5.4.3	Vizualizácia dát konkrétneho OVM.	73
5.5	Práca s agendovým modelom „Zdieľanie datasetov medzi OVM“	74
5.5.1	Vizualizácia zdieľania konkrétneho Datasetu	75
5.6	Synergia Graphlytic s Neo4J	76
6	Identifikácia priestoru na zlepšenie	78
6.1	Využitie znalostných grafov	78
6.2	Využitie znalostných grafov v kontexte E-Government-e	79
6.3	Využitie znalostných grafov v kontexte Dátovej kancelárie	81
6.4	Využitie znalostných grafov v kontexte dátovej integrácie	82
6.5	Integrácia Neo4j s nadstavbou Graphlytic na IS CSRÚ	84
6.6	Odporúčane odborné role pre prácu so znalostnými grafmi	86
6.7	Publikácia výsledkov meraní pre oblasti využitia grafových nástrojov vo Verejnej správe	87

## Zoznam obrázkov

Obrázok 1: Aplikačná architektúra Graphlyticu .....	67
Obrázok 2: Schéma architektúry riešenia.....	67
Obrázok 3: Hlavný grafový model .....	69
Obrázok 4: Agendový grafový model .....	70
Obrázok 5: Mapovanie grafových vzorcov .....	71
Obrázok 6: Prostredie pre prístup do Graphlyticu .....	72
Obrázok 7: Náhľad na vizualizáciu možnosti rozdelenia projektov .....	73
Obrázok 8: vizualizácia údajov konkrétneho OVM .....	74
Obrázok 9: 5.5.1 Vizualizácia zdieľania konkrétneho Datasetu.....	75

## Zoznam tabuliek

Tabuľka 1: Rozdelenie OVM.....	16
Tabuľka 2: Rozdelenie integrácii podľa fáz v projekte DI .....	19
Tabuľka 3: Inštitúcie zahrnuté v analýze .....	22
Tabuľka 4: Výhody / nevýhody RDBMS .....	58
Tabuľka 5Výhody / nevýhody GBDMS .....	59

# Error! Reference source

## 1 Úvod a zhrnutie

Ministerstvo investícií, regionálneho rozvoja a informatizácie Slovenskej republiky (ďalej len „MIRRI SR“) aktuálne realizuje systematický dátový program (ukončenie 2023), ktorý je zameraný na zavedenie manažmentu údajov verejnej správy, t. j. nielen pre MIRRI SR.

Dátový program smeruje k praktickému a efektívnemu napĺňaniu cieľov Programového vyhlásenia vlády Slovenskej republiky na obdobie rokov 2021 – 2024, ako aj platných, účinných a pripravovaných národných strategických dokumentov a to najmä:

- Národnej koncepcie informatizácie verejnej správy,
- Iniciatívy pre otvorené vládnutie
- Európskych strategických dokumentov (najmä Európskej dátovej stratégie), právne záväzných aktov Európskej únie (GDPR, PSI, DGA, EDA)
- Slovenskej legislatívy
  - zákon č. 305/2013 Z. z. o elektronickej podobe výkonu pôsobnosti orgánov verejnej moci a o zmene a doplnení niektorých zákonov (zákon o e-Governmente) v znení neskorších predpisov,
  - zákon č. 177/2018 Z. z. o niektorých opatreniach na znižovanie administratívnej záťaže využívaním informačných systémov verejnej správy a o zmene a doplnení niektorých zákonov (zákon proti byrokracii) v znení neskorších predpisov,
  - zákon č. 95/2019 Z. z. o informačných technológiách vo verejnej správe a o zmene a doplnení niektorých zákonov v znení neskorších predpisov, návrh zákona o údajoch).

Nosnou časťou dátového programu je zavádzanie systematického manažmentu údajov vo verejnej správe prostredníctvom tzv. dopytových výziev, z ktorých vzišlo cca 90 projektov a národného projektu Dátová integrácia.

Účelom realizácie dátovej integrácie je integrácia na centrálnu platformu (vrátane úpravy integračných služieb agendových informačných systémov) na strane IS VS a to buď formou sprístupnenia nových objektov evidencie pre zdieľanie voči oprávneným orgánom verejnej moci alebo formou konzumovania už dostupných údajov nevyhnutných vo vzťahu ku svojej úradnej činnosti.

Národný projekt DI prispeje výraznou mierou k naplneniu princípu „jedenkrát a dosť“ v rámci verejnej správy tým, že dôjde k sprístupneniu infraštruktúry pre zdieľanie kritickej masy údajov medzi orgánmi verejnej moci. Projekt zároveň podporí optimalizáciu agend verejnej správy, digitalizáciu životných situácií, zníženie byrokracie, zvýšenie kvality údajov vedených v registroch, zabezpečí dostupnosť otvorených datasetov a dátovú interoperabilitu.



# Error! Reference source

## 1.1 Kontext

**Detailný výstup č. 1.1.7: Analýza toku údajov vznikol ako aktualizácia výstupov dostupných z projektu Dátová integrácia (DI), zo zoznamu dátových entít a z analýzy dopytových projektov zameraný na Manažment údajov inštitúcii verejnej správy.**

Dokument bol pripravený v rámci projektu „Zlepšenie využívania údajov vo verejnej správe“. Tento projekt má ambíciu transformovať fungovanie inštitúcií verejnej správy tak, aby dokázali maximálne efektívne spravovať a zdieľať údaje, využívať údaje pre lepšie rozhodovanie na základe faktov a dôkazov, pre zlepšenie efektivity a adresnosti služieb na základe lepšieho využívania dát.

Projekt Zlepšenie využívania údajov vo verejnej správe realizuje Dátová kancelária verejnej správy ako špeciálna jednotka Ministerstva investícií, regionálneho rozvoja a informatizácie (ďalej aj MIRRI SR).

Vstupmi do analýzy toku údajov sú tri oblasti zoznamov údajov:

### 1.1.1 Zhromažďovanie dát prostredníctvom DI

Realizácia približne 800 nových integračných väzieb medzi informačnými systémami verejnej správy, prostredníctvom modulu procesnej integrácie a integrácie údajov ako centrálnemu komponentu eGovernmentu (centrálne integračná platforma)

- zapojenie 31 orgánov verejnej moci v roli nových poskytovateľov zdieľaných údajov,
- zapojenie 39 orgánov verejnej moci v roli nových konzumentov zdieľaných údajov,
- sprístupnenie 223 nových typov údajov – objektov evidencie (datasetov),
- vykonanie 512 meraní dátovej kvality v rámci objektov evidencie (datasetov) spravovaných v
- informačných systémoch verejnej správy,
- identifikácia kľúčových registrov verejnej správy a možností zdieľania údajov z týchto registrov,
- vrátane otvorených údajov, prostredníctvom centrálnemu komponentu eGovernmentu,
- zmapovanie logických dátových modelov informačných systémov verejnej správy vrátane ich
- možného zosúladenia s centrálnym modelom verejnej správy.

# Error! Reference source

## 1.1.2 Zoznam dátových entít a číselníkov.

Zoznam obsahuje nasledovné rozdelenie:

### 1.1.2.1 *Objekty evidencie, identifikované ako kľúčové údaje, ktoré sa nachádzajú v CSRÚ alebo je plánované ich pripojenie pre účely integrácie a zdieľania*

- Spracovaný je aktuálny stav ako aj nadchádzajúce plánované integrácie poskytovateľov dát
- Pre jednotlivé objekty evidencie sú identifikované prislúchajúce dátové entity
- Definuje, či sa entita nachádza v Centrálnom modeli údajov
- Na základe predpokladaného počtu konzumentov ako aj významu samotnej entity je definovaná miera prepoužiteľnosti dátovej entity čím sa zvyšuje jej potenciál na zaradenie do Centrálného modelu údajov
- Zoznam zahŕňa analýzu jednotlivých objektov evidencie a dátových entít z pohľadu kandidátov na Moje dáta
- Samotné prepojenie je odstupňovanie podľa dôležitosti

### 1.1.2.2 *Dáta, z pohľadu verejnosti (dáta vyžadované verejnosťou a dáta s vysokým potenciálom na prepoužitie)*

- Všetky datasey/objekty evidencie sú analyzované z pohľadu potenciálu na zaradenie do CMÚ
- Pri jednotlivých záznamoch je definovaná ich súčasná prístupnosť verejnosti
- Pre všetky objekty je definovaná ich dôležitosť z pohľadu zaradenie dátovej entity do CMÚ
- Jednotlivé záznamy majú prepojenie na plán CSRÚ (bod 1- Objekty evidencie, identifikované ako kľúčové údaje)

### 1.1.2.3 *Kľúčové číselníky*

- Kompletná analýza číselníkov, ktoré sú prostredníctvom integrácií v CSRÚ používané
- Analýza číselníkov z ďalších, dnes využívaných zdrojov očistená o prípadné duplicity
- Pre vyhlásené základné číselníky sa v zozname nachádzajú aj schválené/navrhované jednotné referencovateľné identifikátory
- Pri číselníkoch je uvádzaný spôsob poskytovania číselníkov tretím stranám
- Zoznam obsahuje a kategorizuje odporúčania pre všetky číselníky, ktoré nie sú základné s definovaním ich potenciálu stať sa základným číselníkom

# Error! Reference source

## 1.1.3 Údaje obsiahnuté v dopytových výzvach zameraných na manažment údajov inštitúcie verejnej správy

- Manažment údajov inštitúcie VS (OPII\_2018\_7\_3\_DOP)
- Manažment údajov inštitúcie VS (OPII-2019/7/6-DOP)
- Manažment údajov inštitúcie VS (OPII-2021/7/15-DOP)

## 1.2 Metodika realizácie výstupu

Stručne povedané, analýza toku údajov vo vláde zahŕňa sledovanie štruktúrovaného procesu, ktorý zahŕňa identifikáciu relevantných údajov, zhromažďovanie a čistenie údajov, výber správneho analytického nástroja, analýzu údajov, identifikáciu oblastí na zlepšenie a odporúčanie zlepšení na základe poznatkov zistených z analýzy. Tento proces môže byť náročný, ale so správnymi nástrojmi, metodológiou a odbornými znalosťami je možné zefektívniť vládne procesy tak, aby priniesli efektívnosť operácií, zlepšili tvorbu politik a verejné služby.

Pri definícii množiny poskytovateľov údajov sa zameriame na najdôležitejšie verejné inštitúcie a priority projektu DI podľa špecifikácii požiadaviek tak, aby analýza tokov údajov mohla byť prakticky použiteľná pri riadení dátovej integrácie po skončení samotného projektu (čo vnímame ako hlavný účel dokumentu).

Pre určenie dátových objektov a modelovanie komunikačných väzieb bol vytvorený jednoduchý excel, ktorý obsahuje:

- Názov orgánu verejnej moci - OVM
- Názov informačného systému - ISVS
- Kód informačného systému - ISVS\_kod
- Typ integračnej väzby
  - Poskytovateľ
  - Konzument
- Spôsob integrácie
  - Priame integrácie
  - Integrácia na CSRÚ
- Status
  - DNR – množina údajov definovaný / identifikovaných v DNR v projekte DI
  - DOP – Množina poskytovateľských väzieb identifikovaných v dopytových projektoch

# Error! Reference source

- ID8 - obsahujúci cieľový zoznam poskytovaných a konzumovaných OE
- IMPL-test – Prebieha integrácia datasetu
- PLAN – To be stav, po skončení projektu DI
- PROD – OE je už v produkcii
- PROD-DI – OE bude dostupný v produkcii počas projektu DI
- Výzva – OE je zaradený v backlogu integrácii v projekte DI
- Register / Objekt evidencie
- Dataset

Objekty evidencie budú prepojené s Centrálnym modelom údajov. Vznikne tak konsolidovaný zoznam objektov evidencie, s ktorým bude možné ďalej pracovať (pri ostatných úlohách Dátovej kancelárie, ako je vyhlasovanie referenčných údajov alebo služba Moje dáta).

Dokument tak vychádzať z vstupov definovaných v kapitole 1.1. Analýza tokov bude publikovaná na webovom sídle Dátovej kancelárie.

Vzhľadom na rozsiahlosť údajov, ktorými verejná správa disponuje bude obsah tohto dokumentu zameraný na spôsob spracovania veľkej masy údajov prívetivejším spôsobom z pohľadu prevádzkovateľa znalostných grafov a ich aktualizácii skrz vizuálnu reprezentáciu, ktorý bude viesť ku praktickým a efektívnejším pre definovanie nových cieľov. Výstupom znalostného grafu je obvykle reprezentácia dôležitých informácií, vzťahov a vzorov vo forme grafového modelu.

# Error! Reference source

## 2 Definícia množín poskytovateľov a konzumentov údajov

V rámci výstupu boli analyzované realizované integrácie OVM v troch fázach z projektu DI pre nasledujúce z nižšie uvedenej množiny OVM.

V projekte DI bolo plánované zapojenie 60 OVM, pričom každá zo zapojených OVM môže byť súčasne v jednej z nasledujúcich roli:

### 2.1 Nový poskytovateľ údajov

OVM nie je integrovaný na IS CSRÚ v roli poskytovateľa. Súčasťou realizácie integračnej väzby na strane IS CSRÚ sú nasledujúce aktivity:

- Špecifikácia OE pre účely pripojenia
- Príprava a podpis DIZ (použité služby, štruktúra OE, harmonogram)
- Vypracovanie technicko-implentačného projektu
- Registrácia OVM v roli poskytovateľa
- Nastavenie a otestovanie sieťového prepojenia
- Registrácia OE
- Integračné testy
- Príprava dokumentu SLA

### 2.2 Existujúci poskytovateľ – rozšírené poskytovanie údajov

OVM je integrovaný na IS CSRÚ v roli poskytovateľa. Súčasťou realizácie integračnej väzby na strane IS CSRÚ sú nasledujúce aktivity:

- Príprava a podpis dodatku DIZ (použité služby, štruktúra OE, harmonogram)
- Aktualizácia technicko-implentačného projektu
- Registrácia OE
- Integračné testy
- Aktualizácia dokumentu SLA

### 2.3 Nový konzument údajov

OVM nie je integrovaný na IS CSRÚ v roli konzumenta. Súčasťou realizácie integračnej väzby na strane IS CSRÚ sú nasledujúce aktivity:

- Špecifikácia OE pre účely pripojenia
- Príprava a podpis DIZ (použité služby IS CSRÚ, štruktúra OE, harmonogram)
- Vypracovanie technicko-implentačného projektu

# Error! Reference source

- Registrácia OVM v roli konzumenta
- Nastavenie a otestovanie sieťového prepojenia
- Nastavenie práv pre poskytované OE
- Integračné testy
- Príprava dokumentu SLA

## 2.4 Existujúci konzument – rozšírené konzumovanie údajov

OVM je integrovaný na IS CSRÚ v roli konzumenta. Súčasťou realizácie integračnej väzby na strane IS CSRÚ sú nasledujúce aktivity:

- Príprava a podpis dodatku DIZ (použitie služby IS CSRÚ, štruktúra OE, harmonogram)
- Aktualizácia technicko-implentačného projektu
- Nastavenie práv pre poskytované OE
- Integračné testy
- Aktualizácia dokumentu SLA

Počet inštitúcií ktoré používajú CSRÚ (konzumenti(57), poskytovatelia (38))

- Počet ISVS ktorý používajú CSRÚ (konzumenti, poskytovatelia)
- Počet dostupných objektov evidencie
- Počet dopytových projektov, ktoré sa budú pripájať v roku 2023
- Počet inštitúcií, ktoré používajú služby dátovej kvality, výsledky merania dátovej kvality
- Výsledky zákona proti byrokracii, koľko výpisov a odpisov už nie je potrebné vymenovať -
- Počet dátových zdrojov pripojených na KAV a počet prípadov použitia, ktoré sa v rámci KAV - 0

# Error! Reference source

P.č.	Názov OVM	Postavenie OVM k začatiu realizačnej fáze Projektu	
		Poskytovateľ	Konzument
1	Ministerstvo pôdohospodárstva a rozvoja vidieka SR	NP	EK
2	Národné lesnícke centrum	NP	NK
3	Štátny veterinárny ústav Zvolen	NP	NK
4	Štátna veterinárna a potravinová správa Slovenskej republiky	NP	NK
5	Štátny veterinárny a potravinový ústav	NP	NK
6	Ústredný kontrolný a skúšobný ústav poľnohospodársky	NP	NK
7	Hydromeliorácie, štátny podnik	NP	NK
8	Plemenárske služby SR, š. p.	EP	NK
9	Ministerstvo zdravotníctva SR	NP	NK
10	Úrad verejného zdravotníctva SR	NP	NK
11	Štatistický úrad SR	EP	NK
12	Finančné riaditeľstvo SR	EP	EK
13	Datacentrum elektronizácie územnej samosprávy	NP	EK
14	Ministerstvo vnútra SR	EP	NK
15	Ministerstvo investícií, regionálneho rozvoja a informatizácie SR	EP	EK
16	Ministerstvo školstva, vedy, výskumu a športu SR	EP	EK
17	Notárska komora	NP	EK
18	Ministerstvo životného prostredia	NP	EK
19	Ministerstvo kultúry SR	NP	EK
20	Úrad geodézie, kartografie a katastra SR	EP	EK
21	Ministerstvo spravodlivosti SR	EP	EK
22	Pamiatkový úrad	NP	NK
23	Ministerstvo hospodárstva SR	NP	EK
24	VUJE, a. s.	NP	EK
25	Úrad pre dohľad nad zdravotnou starostlivosťou	NP	EK
26	Slovak Business Agency (SBA)	NP	NK
27	Slovenská inovačná a energetická agentúra	NP	EK
28	Úrad pre reguláciu sieťových odvetví	NP	NK

# Error! Reference source

29	Generálna prokuratúra SR	EP	NK
30	Úrad vlády SR	NP	NK
31	Národný inšpektorát práce	NP	NK
32	Úrad pre verejné obstarávanie*	NP	NK
33	Ministerstvo financií SR	EP	EK
34	Pôdohospodárska platobná agentúra	NP	NK
35	Národná agentúra pre sieťové a elektronické služby	NP	NK
36	Ministerstvo dopravy a výstavby SR	NP	NK
37	Úrad pre reguláciu elektronických komunikácií a poštových služieb (ÚPREKaPS)	NP	NK
38	Slovenská konsolidačná, a.s.	NP	EK
39	Dopravný úrad	NP	NK
40	Ministerstvo práce, sociálnych vecí a rodiny SR a	EP	EK
41	Ústredie práce, sociálnych vecí a rodiny (ÚPSVaR)	EP	EK
42	Prezídium policajného zboru SR	NP	NK
43	Ministerstvo zahraničných vecí a európskych záležitostí SR	NP	NK
44	Úrad priemyselného vlastníctva SR	NP	NK
45	Zbor väzenskej a justičnej stráže SR	NP	NK
46	Štátny fond rozvoja bývania	NP	NK
47	Slovenský pozemkový fond	NP	NK
48	Národné centrum zdravotníckych informácií	NP	EK
49	Protimonopolný úrad SR	EP	NK
50	Sociálna poisťovňa	EP	EK
51	Ministerstvo obrany SR	NP	NK
52	Národná banka Slovenska	NP	NK
53	EXIMBANKA SR	NP	NK
54	Najvyšší kontrolný úrad	NP	NK
55	Slovenská obchodná inšpekcia	NP	NK
56	Environmentálny fond	NP	NK
57	Ústavný súd SR	NP	NK
58	Slovenská agentúra pre rozvoj investícií a obchodu	NP	NK
59	DÔVERA zdravotná poisťovňa, a. s.	EP	EK
60	Všeobecná zdravotná poisťovňa, a. s.	EP	EK



# Error! Reference source

61	Union zdravotná poisťovňa, a. s.	EP	EK
62	Bratislavský samosprávny kraj	NP	NK
63	Trnavský samosprávny kraj	NP	NK
64	Trenčiansky samosprávny kraj	NP	EK
65	Nitriansky samosprávny kraj	NP	NK
66	Žilinský samosprávny kraj	NP	NK
67	Banskobystrický samosprávny kraj	NP	NK
68	Prešovský samosprávny kraj	NP	NK
69	Košický samosprávny kraj	NP	NK
70	Slovenská banková asociácia	NP	NK
71	Centrum právnej pomoci	NP	NK
72	Slovenská agentúra životného prostredia	NP	NK
73	Agentúra štátom podporovaného nájomného bývania	NP	NK
74	Národné osvetové centrum	NP	NK
75	Národný bezpečnostný úrad	NP	EK
76	Slovenská banková asociácia	NP	NK
77	Slovenská asociácia poisťovní	NP	NK
78	Rada pre mediálne služby	NP	NK
79	Kancelária najvyššieho súdu SR	NP	NK

**Legenda:**  
NP - Nový poskytovateľ (Zapojenie nového OVM v roli poskytovateľa údajov)  
EP - Existujúci poskytovateľ (Rozšírenie poskytovania údajov už zapojeného OVM)  
NK - Nový konzument (Zapojenie nového OVM v roli konzumenta údajov)  
EK - Existujúci konzument (Rozšírenie konzumovania údajov už zapojeného OVM)

Tabuľka 1: Rozdelenie OVM

# Error! Reference source

Názov subjektu (OVM)	Počet integrácií ako konzument údajov	Počet integrácií ako poskytovateľ údajov	Počet integrácií spolu
<b>Fáza 1 - spolu</b>	<b>294</b>	<b>180</b>	<b>474</b>
Datacentrum elektroniz. územ. samosprávy	38	17	55
Finančné riaditeľstvo SR	14	5	19
Generálna prokuratúra SR	12	3	15
Ministerstvo dopravy a výstavby SR	5	7	12
Ministerstvo financií SR	6	5	11
Ministerstvo kultúry SR	22	3	25
Ministerstvo pôdohosp. a rozvoja vidieka SR	33	28	61
Ministerstvo školstva, vedy, výskumu a športu SR	61	25	86
Ministerstvo spravodlivosti SR	40	28	68
Ministerstvo vnútra SR	12	34	46
Úrad geodézie, kartografie a katastra SR	24	7	31
Úrad práce, sociálnych vecí a rodiny SR	13	10	23
Úrad pre verejné obstarávanie	14	8	22
<b>Fáza 2 - spolu</b>	<b>82</b>	<b>24</b>	<b>106</b>
Dôvera – zdravotná poisťovňa	1	0	1
European (eJustice) portál	1	0	1
Exportno-importná banka SR	11	0	11
Kancelária ústavného súdu SR	5	0	5
Krajský súd BA	5	0	5
Ministerstvo práce, sociálnych vecí a rodiny	1	1	2
Ministerstvo zdravotníctva	3	1	4
Najvyšší kontrolný úrad SR	3	0	3
Národná diaľničná spoločnosť	4	0	4
Národné centrum zdravot. informácií	0	4	4

# Error! Reference source

Pôdohospodárska platob. Agentúra	2	0	2
Protimonopolný úrad	13	2	15
Rada pre vysielanie a retransmisiu	2	0	2
Rozhlas a televízia slovenska	6	0	6
Slovak Credit Bureau	0	1	1
Slovenská lekárska komora	1	0	1
Sociálna poisťovňa	6	6	12
Štatistický úrad SR	3	0	3
UNION zdravotná poisťovňa	7	0	7
MIRRI	5	2	7
Úrad pre dohľad nad zdrav. starostlivosťou	0	2	2
Úrad pre normalizáciu metrológiu a skúšob.	1	0	1
Úrad pre reguláciu el. komunikácií a pošt. sl.	1	0	1
Úrad vlády SR	0	4	4
Všeobecná zdravotná poisťovňa	1	1	2
<b>Fáza 3 - spolu</b>	<b>117</b>	<b>16</b>	<b>133</b>
Advokátska komora	0	1	1
Banskobystrický samosprávny kraj	10	1	11
Bratislavský samosprávny kraj	0	1	1
Košický samosprávny kraj	15	1	16
Mesto Košice	8	0	8
Mesto Nitra	3	0	3
Mesto Žilina	3	0	3
Ministerstvo hospodárstva	16	0	16
Ministerstvo obrany	0	1	1
Ministerstvo zahraničných vecí a EÚ záležit.	4	1	5
Ministerstvo životného prostredia	5	1	6
Národná agentúra pre sieťové a el. služby	4	1	5
Nitriansky samosprávny kraj	8	1	9
Notárska komora	0	2	2

# Error! Reference source

Policajný zbor SR	1	0	1
Prešovský samosprávny kraj	11	1	12
Slovenská exekútorská komora	0	1	1
Trenčiansky samosprávny kraj	6	0	6
Trnavský samosprávny kraj	10	1	11
Úrad priemyselného vlastníctva	3	0	3
Zbor väzenskej a justičnej stráže SR	0	1	1
Žilinský samosprávny kraj	10	1	11

*Tabuľka 2: Rozdelenie integrácií podľa fáz v projekte DI*

V tabuľke uvedenej nižšie je prehľad inštitúcií z projektu DI a dopytových projektov, ktoré sú zmapované z pohľadu poskytovania údajov do IS CSRÚ. Detailnejší pohľad sa nachádza v prílohe 1.1.7 Analýza tokov údajov.xlsx.

# Error! Reference source

ID	Organizácia
1	Agentúra na podporu výskumu a vývoja
2	Agentúra štátom podporovaného nájomného bývania
3	Archeologický ústav Slovenskej akadémie vied
4	Banskobystrický samosprávny kraj
5	Bratislavský samosprávny kraj
6	Centrum právnej pomoci
7	Centrum vedecko-technických informácií SR
8	Datacentrum elektronizácie územnej samosprávy
9	Dopravný úrad
10	DÓVERA zdravotná poisťovňa, a. s.
11	Environmentálny fond
12	EXIMBANKA SR
13	Finančné riaditeľstvo SR
14	Generálna prokuratúra SR
15	Generálne riaditeľstvo Zboru väzenskej a justičnej stráže
16	Hydromeliorácie, štátny podnik
17	Kancelária najvyššieho súdu SR
18	Kancelária Ústavného súdu Slovenskej republiky
19	Košický samosprávny kraj
20	Mesto Bánovce nad Bebravou
21	Mesto Banská Bystrica
22	Mesto Komárno
23	Mesto Levice
24	Mesto Levoča
25	Mesto Liptovský Mikuláš
26	Mesto Lučenec
27	Mesto Myjava
28	Mesto Nové Mesto nad Váhom
29	Mesto Nové Zámky
30	Mesto Prešov
31	Mesto Prievidza
32	Mesto Rimavská Sobota
33	Mesto Ružomberok
34	Mesto Senica
35	Mesto Snina
36	Mesto Svidník
37	Mesto Šamorín
38	Mesto Topoľčany
39	Mesto Trenčín
40	Mesto Trnava
41	Mesto Trstená
42	Mesto Tvrdošín

# Error! Reference source

43	Mesto Zlaté Moravce
44	Mesto Zvolen
45	Ministerstvo dopravy a výstavby SR
46	Ministerstvo financií SR
47	Ministerstvo hospodárstva SR
48	Ministerstvo investícií, regionálneho rozvoja a informatizácie SR
49	Ministerstvo kultúry SR
50	Ministerstvo obrany SR
51	Ministerstvo pôdohospodárstva a rozvoja vidieka SR
52	Ministerstvo práce, sociálnych vecí a rodiny SR
53	Ministerstvo spravodlivosti SR
54	Ministerstvo školstva, vedy, výskumu a športu SR
55	Ministerstvo vnútra SR
56	Ministerstvo zahraničných vecí a európskych záležitostí SR
57	Ministerstvo zdravotníctva SR
58	Ministerstvo životného prostredia SR
59	Najvyšší kontrolný úrad
60	Národná agentúra pre sieťové a elektronické služby
61	Národná banka Slovenska
62	Národné centrum zdravotníckych informácií
63	Národné lesnícke centrum
64	Národné osvetové centrum
65	Národné športové centrum
66	Národný bezpečnostný úrad
67	Národný inšpektorát práce
68	Nitriansky samosprávny kraj
69	Notárska komora
70	Obec Cífer
71	Obec Horná Súča
72	obec Lubina
73	Pamiatkový úrad
74	Pamiatkový úrad Slovenskej republiky
75	Plemenárske služby SR, š. p.
76	Pôdohospodárska platobná agentúra
77	Prešovská univerzita v Prešove
78	Prešovský samosprávny kraj
79	Prezídium policajného zboru SR
80	Protimonopolný úrad SR
81	Rada pre mediálne služby
82	Slovak Business Agency (SBA)
83	Slovenská agentúra pre rozvoj investícií a obchodu
84	Slovenská agentúra pre rozvoj investícií a obchodu
85	Slovenská agentúra životného prostredia
86	Slovenská asociácia poisťovní

# Error! Reference source

87	Slovenská banková asociácia
88	Slovenská inovačná a energetická agentúra
89	Slovenská inšpekcia životného prostredia
90	Slovenská komora daňových poradcov
91	Slovenská konsolidačná, a.s.
92	Slovenská obchodná inšpekcia
93	Slovenský pozemkový fond
94	Sociálna poisťovňa
95	Štatistický úrad SR
96	Štátna veterinárna a potravinová správa SR
97	Štátny fond rozvoja bývania
98	Štátny veterinárny a potravinový ústav
99	Štátny veterinárny ústav Zvolen
100	Technická univerzita v Košiciach
101	Trebišov
102	Trenčiansky samosprávny kraj
103	Trnavský samosprávny kraj
104	Union zdravotná poisťovňa, a. s.
105	Univerzita Komenského v Bratislave
106	Univerzita sv. Cyrila a Metoda v Trnave vysoká škola
107	Úrad geodézie, kartografie a katastra SR
108	Úrad pre dohľad nad zdravotnou starostlivosťou
109	Úrad pre reguláciu elektronických komunikácií a poštových služieb (ÚPREKaPS)
110	Úrad pre reguláciu hazardných hier
111	Úrad pre reguláciu sieťových odvetví
112	Úrad pre verejné obstarávanie
113	Úrad priemyselného vlastníctva SR
114	Úrad verejného zdravotníctva SR
115	Úrad vlády SR
116	Ústavný súd SR
117	Ústredie práce, sociálnych vecí a rodiny (ÚPSVaR)
118	Ústredný kontrolný a skúšobný ústav poľnohospodársky
119	Všeobecná zdravotná poisťovňa, a. s.
120	VUJE, a. s.
121	Zbor väzenskej a justičnej stráže SR
122	Žilinská univerzita v Žiline
123	Žilinský samosprávny kraj

*Tabuľka 3: Inštitúcie zahrnuté v analýze*

# Error! Reference source

## 2.6 Mimo rozsah dokumentu Analýza toku údajov

Pre účely analýzy, nie sú zmapované projekty z dopytovej výzvy zameranej na Lepšie využívanie údajov inštitúciami verejnej správy, Malé zlepšenia e-gov.



# Error! Reference source

## 3 Určenie dátových objektov

Správne určenie dátových objektov prináša viaceré výhody. Po prvé, umožňuje efektívne spracovanie dát a zameranie sa na relevantné informácie. Pri presnom vymedzení kategórií a atribútov je možné získavať presné a relevantné výsledky. Po druhé, správne určené dátové objekty uľahčujú interpretáciu výsledkov a zlepšujú ich interpretovateľnosť pre používateľov. Štruktúra a organizácia dátových objektov pomáhajú jasne definovať, ako sú dáta vzájomne prepojené a ako sú používané v analytickom procese.

Okrem toho, správne určenie dátových objektov zvyšuje efektívnosť a opakovateľnosť analýzy dát. Rovnaká štruktúra a organizácia môžu byť aplikované na rôzne sady dát, čo uľahčuje opakovateľnosť a porovnateľnosť výsledkov. Navyše, kvalitné dátové objekty zlepšujú aj správu a spracovanie dát, vrátane udržiavania aktuálnych údajov, indexovania a zálohovania.

Celkovo možno povedať, že správne určenie dátových objektov je kritické pre úspešnú analýzu dát a získanie hodnotných informácií. Poskytuje štruktúru, organizáciu a jasnosť v analytickom procese, čo vedie k presnejším výsledkom a lepšiemu porozumeniu dát. Správne určené dátové objekty tiež zvyšujú efektívnosť, opakovateľnosť a správu dát. Preto je nevyhnutné klásť dôraz na správne určenie dátových objektov ako kľúčový krok pri analýze dát.

Zhodnotenie postupu pre určenie dátových objektov má prínos aj pre analytický proces. Postup pre určenie dátových objektov je dôležitým krokom v analytickom procese, pretože poskytuje štruktúru, organizáciu a jasnosť v spracovaní a analýze dát. Tento postup zahŕňa identifikáciu cieľov a požiadaviek analýzy, zhromaždenie a evaluáciu dostupných dát, identifikáciu dátových objektov, vytvorenie schémy dátových objektov, implementáciu a správu týchto objektov.

Taktiež je dôležité zdôrazniť, že postup pre určenie dátových objektov prispieva k efektívnosti a efektívnosti analytického procesu. Jasne definované dátové objekty umožňujú lepšiu organizáciu dát a zameranie sa na relevantné informácie. To vedie k úspore času a zdrojov pri analýze dát.

Ďalším prínosom tohto postupu je zlepšenie interpretácie výsledkov analýzy. Kvalitne určené dátové objekty poskytujú štruktúru a vzťahy medzi dátami, čo uľahčuje ich interpretáciu a porozumenie. Tým sa zvyšuje hodnota analýzy a umožňuje sa lepšie využívať výsledky pre rozhodovanie a stratégiu podniku.

Celkovo možno povedať, že postup pre určenie dátových objektov je kľúčovým krokom v analytickom procese. Jeho správne vykonanie zabezpečuje štruktúru, organizáciu, efektívnosť a vysokú hodnotu analýzy dát. Je nevyhnutným nástrojom pre efektívne využívanie dát v podnikovom prostredí a zabezpečuje lepšie rozhodovanie a výkon organizácie.

nasledujúcich častiach postupu sa budeme venovať identifikácii cieľov a požiadaviek, zhromaždeniu a evaluácii dostupných dát, identifikácii konkrétnych dátových objektov a ich atribútov, vytvoreniu schémy dátových objektov a nakoniec implementácii a správe týchto objektov.

Celkový postup pre určenie dátových objektov je komplexný proces, ktorý vyžaduje spoluprácu a komunikáciu medzi členmi tímu, doménovými expertmi a zainteresovanými stranami. Je dôležité získať správne požiadavky a porozumieť konkrétnej doméne, aby sa vytvoril dátový model, ktorý bude presne zodpovedať požiadavkám projektu a zabezpečí efektívne a kvalitné spravovanie dát. Podrobný postup pre určenie dátových objektov môžeme rozdeliť do nasledovných bodov:

- I. Identifikácia cieľov a požiadaviek
- II. Zhromaždenie a evaluácia dostupných dát
- III. Identifikácia dátových objektov
- IV. Vytvorenie schémy dátových objektov
- V. Implementácia a správa dátových objektov

Tento podrobný postup na určenie dátových objektov zabezpečuje systematický a štruktúrovaný prístup k analýze dát

### 3.1.1 Vysvetlenie významu určenia dátových objektov v analytickom procese.

Určenie dátových objektov je kľúčovým krokom v analytickom procese, pretože poskytuje štruktúru a organizáciu pre spracovanie a analýzu dát. Tieto objekty predstavujú jednotlivé entity, ktoré sú relevantné pre danú analýzu a majú svoje vlastné atribúty a vzťahy s inými objektmi.

Význam určenia dátových objektov je nasledovný:

1. **Organizácia dát:** Dátové objekty umožňujú logické usporiadanie a organizáciu dát. Pomáhajú nám zoskupovať príbuzné údaje a identifikovať vzťahy medzi nimi. To zlepšuje prehľadnosť a štruktúru dát, čo uľahčuje prácu s nimi.
2. **Jasná identifikácia a definícia entít:** Určenie dátových objektov znamená jasne identifikovať dôležité entity, ktoré budeme analyzovať. Definícia týchto entít a ich atribútov nám umožňuje presne špecifikovať, aké informácie budeme zbierať a analyzovať.

# Error! Reference source

3. **Vzťahy medzi objektmi:** Dátové objekty zachytávajú vzťahy medzi rôznymi entitami. Tieto vzťahy nám pomáhajú pochopiť spojitosti a interakcie medzi dátami. Identifikovanie týchto vzťahov je dôležité pre vykonávanie komplexnejších analýz, napríklad pri zostavovaní spojenej analýzy viacerých objektov.
4. **Jednotný rámec pre analýzu:** Určenie dátových objektov poskytuje jednotný rámec a štruktúru pre analytický proces. Definovaním objektov a ich atribútov získavame jasný prehľad o tom, čo presne skúmame a aké údaje sú pre nás relevantné. To zlepšuje efektivitu a presnosť analýzy.
5. **Efektívne spracovanie a manipulácia s dátami:** Identifikácia dátových objektov nám umožňuje efektívne spracovávať a manipulovať s dátami. S definovanými objektmi vieme presne určiť, aké operácie budeme vykonávať a aké výsledky očakávame. To nám umožňuje optimalizovať analytické procesy a dosahovať lepšie výsledky.

Určenie dátových objektov je základným krokom pri príprave a realizácii analytického procesu. Poskytuje nám štruktúru, organizáciu a jasný rámec pre spracovanie a analýzu dát. Týmto spôsobom sa zabezpečuje, že analýza je presná, efektívna a produkuje relevantné výsledky potrebné pre rozhodovanie a dosahovanie stanovených cieľov.

## 3.1.2 Súčasný kontext: množstvo dostupných dát a potreba účinných nástrojov na ich analýzu

V súčasnom kontexte sme svedkami rýchleho nárastu množstva dostupných dát. S rozvojom digitálnej transformácie, internetu vecí, sociálnych médií a ďalších technologických pokrokov sa generuje obrovské množstvo štruktúrovaných aj neštruktúrovaných dát. Tieto dáta sú cenným zdrojom informácií pre organizácie, a preto je dôležité mať účinné nástroje na ich analýzu.

Množstvo dostupných dát prináša so sebou výzvy. Veľké objemy dát môžu byť zložité na spracovanie a analyzovanie tradičnými metódami. Okrem toho, dáta často pochádzajú z rôznych zdrojov a majú rôzne formáty a štruktúry, čo komplikuje ich integráciu a konzistentné spracovanie.

Pre účinnú analýzu týchto dát je potrebné mať nástroje, ktoré sú schopné riešiť tieto výzvy. Moderné analytické nástroje a platformy ponúkajú vysokovýkonné spracovanie dát, paralelné spracovanie, distribuované databázy a iné technológie, ktoré zvládajú spracovať veľké objemy dát rýchlo a efektívne.

# Error! Reference source

Okrem toho, nástroje na analýzu dát musia byť schopné pracovať s rôznymi typmi dát, vrátane štruktúrovaných, neštruktúrovaných. Musia ponúkať možnosti pre extrakciu, transformáciu a načítanie (ETL) dát, vizualizáciu, štatistické analýzy, strojové učenie a ďalšie pokročilé analytické techniky.

Účinné nástroje na analýzu dát by mali byť aj ľahko použiteľné a poskytovať intuitívne rozhranie, aby boli prístupné nielen odborníkom na dáta, ale aj širšiemu okruhu používateľov. Tieto nástroje by mali umožňovať rýchle a flexibilné modelovanie dát, prispôsobiteľné vizualizácie a interaktívne možnosti pre prieskum dát.

V súčasnom kontexte je potreba účinných nástrojov na analýzu dát veľmi významná. S množstvom dostupných dát sa organizácie snažia získať hodnotné informácie a podporiť rozhodovanie na základe dôkazov. Účinné nástroje umožňujú organizáciám extrahovať hodnotu z dát a odhaliť vzorce, trendy a informácie, ktoré môžu viesť k inováciám, konkurenčnej výhode a lepším výsledkom.

### 3.1.3 Cieľ úvodného postupu: poskytnúť štruktúru a organizáciu pre ďalšie spracovanie a analýzu

Cieľom úvodného postupu pre určenie dátových objektov je poskytnúť štruktúru a organizáciu pre ďalšie spracovanie a analýzu dát. Tento postup je dôležitým prvým krokom v analytickom procese, kde sa identifikujú relevantné dáta a vytvára sa rámec pre ich spracovanie.

Poskytnutím štruktúry a organizácie sa dosahuje niekoľko hlavných cieľov:

1. **Kategorizácia a zoskupenie dát:** Úvodný postup pomáha kategorizovať a zoskupovať dáta na základe ich charakteristík a vzťahov. Týmto spôsobom sa dáta delia do relevantných skupín, čo uľahčuje ich následné spracovanie a analýzu. Napríklad, dáta môžu byť rozdelené podľa tematickej oblasti, časového obdobia, geografickej polohy alebo iných faktorov.
2. **Identifikácia dôležitých entít:** Úvodný postup pomáha identifikovať dôležité entity alebo objekty, ktoré majú záujem v analytickom procese. Týmto spôsobom sa zabezpečuje, že sa zameriavame na relevantné informácie a vlastnosti, ktoré budú slúžiť ako základ pre ďalšiu analýzu. Identifikované entity môžu byť napríklad zákazníci, produkty, transakcie alebo iné relevantné subjekty.
3. **Definovanie atribútov a vzťahov:** Úvodný postup umožňuje definovať atribúty (vlastnosti) pre každú identifikovanú entitu a analyzovať vzťahy medzi nimi. Týmto spôsobom sa presne určujú dôležité informácie, ktoré chceme

# Error! Reference source

zbierať a analyzovať, a taktiež sa rozpoznávajú vzťahy, ako sú napríklad hierarchie, asociácie alebo závislosti medzi entitami.

4. **Vytvorenie štruktúry dátového modelu:** Úvodný postup umožňuje vytvorenie štruktúry dátového modelu, ktorý poskytuje organizovaný rámec pre spracovanie a analýzu dát. Dátový model môže obsahovať tabuľky, relácie, schémy a iné prvky, ktoré reprezentujú identifikované entity, ich atribúty a vzťahy. Týmto spôsobom sa vytvára štruktúrovaný prehľad o dátach, čo uľahčuje ďalšie spracovanie, dotazovanie a analýzu.

Celkovo má úvodný postup pre určenie dátových objektov za cieľ poskytnúť štruktúru a organizáciu pre ďalšie spracovanie a analýzu dát. Týmto spôsobom sa zabezpečuje efektívna manipulácia s dátami, získavanie relevantných informácií a dosahovanie cieľov analytického procesu.

## 3.2 Identifikácia cieľov a požiadaviek

Identifikácia cieľov a požiadaviek je kľúčovým krokom pri určovaní dátových objektov. Stanovenie jasných cieľov a požiadaviek pomáha určiť, aké informácie potrebujeme získať a čo chceme dosiahnuť pomocou analýzy dát. Tu je podrobnejší postup na identifikáciu cieľov a požiadaviek:

### 3.2.1 Stanovenie jasných cieľov pre analýzu dát a ich dosiahnutie

Stanovenie jasných cieľov pre analýzu dát je kľúčovým prvkom analytického procesu. Tu je vysvetlenie významu stanovenia cieľov a spôsob, ako ich dosiahnuť:

1. **Koncepčné porozumenie:** Stanovenie cieľov umožňuje získať koncepčné porozumenie toho, čo chceme dosiahnuť pomocou analýzy dát. Definovanie konkrétnych cieľov nám pomáha zameriavať sa na podstatné otázky a informácie, ktoré chceme z dát získať. Týmto spôsobom sa zabráni strate času a zdrojov na nezmyselnú analýzu a sústredíme sa na to, čo je skutočne dôležité.
2. **Vytvorenie merateľných ukazovateľov:** Stanovenie cieľov nám umožňuje vytvoriť merateľné ukazovatele, ktoré slúžia na hodnotenie úspešnosti analýzy dát. Merateľné ciele sú konkrétne, kvantifikovateľné a časovo ohraničené. Napríklad, cieľom môže byť zvýšenie predaja o 10% v priebehu nasledujúceho štvrťroka alebo zlepšenie zákazníckej spokojnosti o 15% do konca roka. Tieto ukazovatele nám umožňujú sledovať pokrok a hodnotiť efektivitu našej analytického procesu.

# Error! Reference source

- 3. Definovanie výskumných otázok:** Stanovenie cieľov pomáha definovať konkrétne výskumné otázky, ktoré chceme zodpovedať pomocou analýzy dát. Výskumné otázky nám umožňujú sústrediť sa na konkrétne aspekty a získať potrebné informácie. Napríklad, ak máme cieľ zvýšiť zákaznícku vernosť, výskumné otázky by mohli byť: Aké faktory ovplyvňujú zákaznícku vernosť? Aké sú najúčinnšie opatrenia na zvýšenie vernosti zákazníkov.
- 4. Orientácia na výsledky:** Stanovenie cieľov nám umožňuje byť orientovaní na výsledky. Kľúčové je definovať jasné očakávania a očakávané výsledky, ktoré chceme dosiahnuť pomocou analýzy dát. Týmto spôsobom sa zabezpečuje merateľný a zmysluplný výstup z analytického procesu, ktorý slúži ako základ pre rozhodnutia a akcie.

Pre dosiahnutie stanovených cieľov pre analýzu dát je dôležité mať jasný plán a postup. To zahŕňa výber správnych analytických metód a nástrojov, zhromažďovanie relevantných dát, správnu interpretáciu výsledkov a efektívnu komunikáciu výsledkov. Každý krok a rozhodnutie by malo byť smerované k dosiahnutiu stanovených cieľov a poskytnutiu hodnotných informácií pre organizáciu.

## 3.2.2 Definovanie požadovaných informácií a výstupov z analýzy.

Definovanie požadovaných informácií a výstupov z analýzy je dôležitým krokom pri stanovovaní cieľov analýzy dát. Tu je vysvetlenie významu a spôsobu definovania požadovaných informácií a výstupov:

- 5. Identifikácia potrebných informácií:** Je dôležité presne určiť, aké informácie potrebujeme z analýzy dát získať. Toto sa dosahuje tým, že si kladieme otázky, ktoré sú relevantné pre náš analytický cieľ. Napríklad, ak sa zaoberáme analýzou zákazníckeho správania, môžeme potrebovať informácie o preferenciách zákazníkov, ich nákupnom správaní, demografických údajoch a podobne. Identifikácia požadovaných informácií nám umožňuje sústrediť sa na získavanie konkrétnych dát, ktoré sú relevantné pre náš analytický cieľ.
- 6. Stanovenie typu výstupov:** Výstupy z analýzy dát môžu byť rôznorodé a závisia od konkrétneho cieľa a potrieb organizácie. Môžu to byť štatistické prehľady, vizualizácie, predikcie, modely, odporúčania alebo iné formy informácií. Je dôležité definovať, aký typ výstupov je pre nás najužitočnejší a najrelevantnejší pre rozhodovanie a akcie. Napríklad, ak analyzujeme predajné údaje, výstupom môže byť zoznam najpredávanejších produktov, predpovede predaja na nasledujúce obdobie, alebo segmentácia zákazníkov na základe ich nákupných preferencií.

# Error! Reference source

- 7. Špecifikácia formátu a prezentácie výstupov:** Okrem samotných informácií je dôležité špecifikovať aj formát a spôsob prezentácie výstupov z analýzy dát. To zahŕňa rozhodnutie o tom, či preferujeme tabuľkové prehľady, grafy, prezentácie, správy alebo iné formy vizualizácie a prezentácie výsledkov. Je potrebné zvážiť preferencie a potreby cieľovej skupiny, ktorá bude výstupy používať, aby sme zabezpečili ich efektívne pochopenie a využitie.
- 8. Kontrola kvality výstupov:** Pri definovaní požadovaných informácií a výstupov je dôležité zohľadniť aj aspekty kvality. Je potrebné stanoviť kritériá pre overenie správnosti a spoľahlivosti výstupov, ako aj ich relevanciu pre riešenie daného problému alebo podporu rozhodovania. Kontrola kvality zahŕňa overenie úplnosti, presnosti a spoľahlivosti dát, ako aj overenie správnosti a interpretability výsledkov analýzy.

Definovanie požadovaných informácií a výstupov z analýzy dát je dôležitým krokom, ktorý pomáha usmerňovať analytický proces a zabezpečiť, že výsledky sú relevantné, užitočné a spoľahlivé pre rozhodovanie a akcie.

## 3.2.3 Určenie typov dát potrebných na dosiahnutie cieľov

Určenie typov dát potrebných na dosiahnutie cieľov je kritickým krokom pri analýze dát. Tu je vysvetlenie významu a spôsobu určenia typov dát:

- 1. Identifikácia potrebných typov dát:** Prvým krokom je presne určiť, aké typy dát sú potrebné na dosiahnutie stanovených cieľov. To vyžaduje bližší pohľad na ciele analýzy a výskumné otázky, ktoré sa snažíme zodpovedať. Napríklad, ak analyzujeme zákaznícku spokojnosť, potrebujeme možno demografické údaje, preferencie, nákupné správanie a spätnú väzbu od zákazníkov. Identifikácia správnych typov dát je kľúčová pre získanie relevantných informácií.
- 2. Štruktúrované vs. neštruktúrované dáta:** Dáta môžu byť štruktúrované alebo neštruktúrované. Štruktúrované dáta sú organizované v tabuľkách alebo databázach s definovanými atribútmi a vzťahmi medzi nimi. Príklady zahŕňajú číselné hodnoty, textové polia, dátumy a iné štruktúrované formáty. Na druhej strane, neštruktúrované dáta majú menej formálneho usporiadania a môžu zahŕňať textové dokumenty, e-maily, správy zo sociálnych médií, obrázky a videá. Je dôležité určiť, aký typ dát je pre naše ciele relevantný.
- 3. Interné vs. externé dáta:** Ďalším hľadiskom je zdroj dát. Interné dáta sú dáta, ktoré organizácia vlastní a sú zozbierané v rámci jej prevádzky, ako napríklad predajné údaje, zákaznícke informácie alebo interné procesy. Externé dáta sú dáta, ktoré organizácia získava z vonkajších zdrojov, ako napríklad verejné štatistiky, trhové trendy alebo dáta od externých poskytovateľov. Identifikácia

# Error! Reference source

potrebných typov dát nám pomáha určiť, či potrebujeme sústrediť sa na interné alebo externé dáta.

4. **Kvantitatívne vs. kvalitatívne dáta:** Ďalším rozlišovaním typov dát je medzi kvantitatívnymi a kvalitatívnymi dátami. Kvantitatívne dáta sú merateľné a vyjadrujú sa numericky. Príklady zahŕňajú číselné hodnoty, počty, percentá a iné kvantitatívne atribúty. Kvalitatívne dáta sú popisné a vyjadrujú sa slovne alebo kategoricky. Príklady zahŕňajú textové popisy, hodnotenia, klasifikácie a iné kvalitatívne atribúty. Zohľadnenie potrebných typov dát nám umožňuje zvoliť vhodné analytické metódy a nástroje.

Určenie správnych typov dát je kľúčové pre úspešnú analýzu a dosiahnutie cieľov. Správna identifikácia umožňuje zameranie sa na získavanie relevantných dát a ich efektívne využitie v analytickom procese.

## 3.2.4 Rozlišovanie štruktúrovaných a neštruktúrovaných dát

Rozlišovanie medzi štruktúrovanými a neštruktúrovanými dátami je dôležité pre ich efektívne spracovanie a analýzu. Tu je vysvetlenie rozdielov medzi týmito dvoma typmi dát:

1. **Štruktúrované dáta:** Štruktúrované dáta sú dáta, ktoré majú jasne definovanú štruktúru a organizáciu. Sú organizované v tabuľkách, databázach alebo iných formátoch, ktoré používajú riadky a stĺpce na uchovávanie dát. Každý atribút má preddefinovaný typ a hodnoty sú uložené v špecifikovanom formáte. Príklady štruktúrovaných dát zahŕňajú číselné hodnoty, dátumy, mená, identifikátory a iné jasne definované údaje. Štruktúrované dáta sú ľahko interpretovateľné a umožňujú jednoduchú manipuláciu a analýzu pomocou štandardných databázových dotazov a analytických nástrojov.
2. **Neštruktúrované dáta:** Neštruktúrované dáta sú dáta, ktoré nemajú preddefinovanú štruktúru a organizáciu. Sú to často textové dáta, ako sú dokumenty, e-maily, správy zo sociálnych médií, obrázky, zvukové a videozáznamy. Neštruktúrované dáta môžu obsahovať komplexné a rôznorodé informácie, ktoré nie sú jednoducho štruktúrované do tabuliek alebo databáz. Ich analýza vyžaduje spracovanie prirodzeného jazyka, vzťahového a obrazového spracovania a iných pokročilých metód na extrakciu, transformáciu a porozumenie obsahu. Neštruktúrované dáta môžu poskytovať cenné informácie, ktoré nie sú dostupné v štruktúrovaných dátach, ako napríklad názory zákazníkov, obrazové vzorce alebo zvukové signály.

Rozlišovanie medzi štruktúrovanými a neštruktúrovanými dátami je dôležité pri príprave a spracovaní dát pre analýzu. Pre štruktúrované dáta môžeme využiť tradičné databázové technológie a analytické nástroje, zatiaľ čo pre



# Error! Reference source

neštruktúrované dáta sú potrebné špeciálne techniky na spracovanie a extrakciu informácií. Mnohokrát sa v analýze využíva kombinácia oboch typov dát, aby sa získali komplexnejšie a komplexnejšie poznatky.

## 3.3 Zhromaždenie a evaluácia dostupných dát

### 3.3.1 Identifikácia zdrojov dát, vrátane interných, externých a verejne dostupných zdrojov

Identifikácia zdrojov dát je kľúčovým krokom pri určovaní, odkiaľ budeme získavať potrebné dáta pre analýzu. Tu je vysvetlenie a rozlišovanie rôznych typov zdrojov:

- 1. Interné zdroje dát:** Interné zdroje dát sú dáta, ktoré organizácia vlastní a sú generované, zaznamenávané alebo zhromažďované v rámci jej prevádzky. Tieto zdroje dát môžu zahŕňať interné databázy, CRM systémy, ERP systémy, predajné údaje, zákaznícke informácie, interné správy a ďalšie. Interné zdroje dát sú často cenné, pretože obsahujú špecifické informácie o organizácii, jej procesoch a zákazníkoch. Identifikácia interných zdrojov dát zahŕňa preskúmanie interných systémov a databáz a určenie, aké dáta sú dostupné a relevantné pre analytické ciele.
- 2. Externé zdroje dát:** Externé zdroje dát sú dáta, ktoré organizácia získava z vonkajších zdrojov mimo svojej organizácie. Tieto zdroje dát môžu zahŕňať údaje od dodávateľov, partnerov, výskumných spoločností, verejné údaje, štatistiky, demografické údaje a iné. Externé zdroje dát môžu poskytnúť širší kontext, benchmarkingové údaje, trendy a ďalšie informácie, ktoré sú užitočné pri analýze a rozhodovaní. Identifikácia externých zdrojov dát zahŕňa identifikáciu relevantných poskytovateľov dát, vyhľadávanie dostupných údajov a získavanie prístupu k týmto zdrojom.
- 3. Verejne dostupné zdroje dát:** Verejne dostupné zdroje dát sú dáta, ktoré sú voľne dostupné pre verejnosť a môžu byť využité pre analytické účely. Tieto zdroje dát môžu zahŕňať vládne portály, štatistické úrady, výskumné organizácie, verejné databázy, sociálne médiá a ďalšie. Verejne dostupné zdroje dát môžu poskytnúť rôznorodé informácie o populácii, ekonomike, životnom prostredí, spoločenských trendoch a ďalších oblastiach. Identifikácia verejne dostupných zdrojov dát zahŕňa vyhľadávanie relevantných databáz, zdrojov a portálov, ktoré poskytujú dáta, ktoré sú relevantné pre analýzu.

# Error! Reference source

Identifikácia a zhromažďovanie zdrojov dát je kritické pre úspešnú analýzu. Zabezpečenie prístupu k relevantným interným, externým a verejne dostupným zdrojom dát umožňuje získanie komplexných a relevantných dát, ktoré sú nevyhnutné pre analytický proces.

## 3.3.2 Analýza dostupných dát a ich relevancie pre ciele analýzy

Analýza dostupných dát a ich relevancie je dôležitým krokom pri určovaní, aké dáta sú relevantné pre dosiahnutie stanovených cieľov analýzy. Tu je postup pre analýzu dostupných dát a ich relevancie:

1. **Identifikácia dostupných dát:** Prehľadajte interné, externe a verejne dostupné zdroje dát, ktoré ste identifikovali v predchádzajúcom kroku. Zhromaždite relevantné informácie o týchto dátových zdrojoch, ako je ich obsah, štruktúra, formát a dostupnosť.
2. **Hodnotenie relevancie:** Pre každý identifikovaný dátový zdroj vykonajte hodnotenie jeho relevancie pre ciele analýzy. Zamerajte sa na to, či obsahuje potrebné atribúty, informácie a vzťahy, ktoré sú kľúčové pre vaše analýzy. Zvážte aj kvalitu a spoľahlivosť týchto dát.
3. **Príprava dát:** Ak ste identifikovali dáta, ktoré sú relevantné pre vaše ciele analýzy, pripravte ich na ďalšie spracovanie a analýzu. To môže zahŕňať procesy ako čistenie dát, transformácia do potrebného formátu, normalizácia, agregácia a iné úpravy.
4. **Analýza dostupných dát:** Po príprave dát sa môžete pustiť do ich analýzy. Použite relevantné analytické metódy, techniky a nástroje na získanie poznatkov a informácií z týchto dát. Vyhodnoťte vzťahy, trendy, vzory alebo anomálie, ktoré sú relevantné pre vaše ciele analýzy.
5. **Porovnanie s cieľmi analýzy:** Porovnajte výsledky analýzy dostupných dát s pôvodnými cieľmi analýzy. Zhodnoťte, do akej miery sa dosiahli stanovené ciele a či sú dostupné dáta dostatočné a relevantné pre tieto ciele. Ak je to potrebné, môžete sa vrátiť k predchádzajúcim krokom a upraviť postup alebo zhromažďovať ďalšie dáta.

Analýza dostupných dát a ich relevancie umožňuje identifikovať najvhodnejšie zdroje a dáta pre dosiahnutie cieľov analýzy. Týmto spôsobom sa zabezpečuje, že vaša analýza je zameraná na relevantné a užitočné informácie, ktoré vám pomôžu pri rozhodovaní a dosahovaní stanovených cieľov.

# Error! Reference source

## 3.3.3 Vykonanie auditu dát pre určenie ich kvality, spoľahlivosti a dostupnosti

Vykonanie auditu dát je kritickým krokom pri posúdení kvality, spoľahlivosti a dostupnosti dát. Audit dát pomáha identifikovať potenciálne problémy a nedostatky v dátach, ktoré by mohli ovplyvniť ich presnosť a spoľahlivosť. Tu je postup pre vykonanie auditu dát:

1. **Definovanie kritérií auditu:** Identifikujte kritériá, na základe ktorých budete hodnotiť kvalitu, spoľahlivosť a dostupnosť dát. Tieto kritériá by mali zahŕňať aspekty ako presnosť, úplnosť, konzistencia, aktualita, dostupnosť, bezpečnosť a ďalšie relevantné faktory.
2. **Zozbieranie dát:** Zhromaždíte relevantné dáta, ktoré sú predmetom auditu. Tieto dáta môžu zahŕňať vzorky dát alebo celkovú súpravu dát, ktoré chcete hodnotiť.
3. **Analýza dát:** Vykonajte analýzu zozbieraných dát na základe stanovených kritérií auditu. Vyhodnoťte presnosť, úplnosť, konzistenciu a ostatné aspekty dát. Použite rôzne metódy a techniky, ako sú štatistické analýzy, overovanie hodnôt, porovnávanie s referenčnými údajmi a podobne.
4. **Identifikácia nedostatkov:** V priebehu auditu identifikujte nedostatky v dátach. Zamerajte sa na chyby, chýbajúce hodnoty, nekonzistencie, duplicity a iné problémy, ktoré by mohli ovplyvniť kvalitu a spoľahlivosť dát.
5. **Vyvodzovanie záverov:** Na základe výsledkov auditu vyvodte závery o kvalite, spoľahlivosti a dostupnosti dát. Identifikujte oblasti, ktoré vyžadujú zlepšenie, a navrhujte opatrenia na odstránenie zistených nedostatkov.
6. **Implementácia opatrení:** Ak boli identifikované nedostatky, prijmite opatrenia na ich odstránenie a zlepšenie kvality a spoľahlivosti dát. Tieto opatrenia môžu zahŕňať zmeny v procesoch zberu dát, zlepšenie kontrolných mechanizmov, opravy chýb, aktualizácie dátových zdrojov a ďalšie kroky na zabezpečenie kvalitných a spoľahlivých dát.

Vykonanie auditu dát pomáha zabezpečiť, že dáta používané v analytickom procese sú spoľahlivé, presné a relevantné. Týmto spôsobom sa zvyšuje dôvera výsledkov analýz a znižuje riziko rozhodovania na základe nedôveryhodných údajov.

# Error! Reference source

## 3.3.4 Vyhodnotenie vhodnosti dát na základe ich kvality, relevancie a dostupnosti

Vyhodnotenie vhodnosti dát je dôležitým krokom pri rozhodovaní, či sú dostupné dáta vhodné na dosiahnutie stanovených cieľov analýzy. Toto vyhodnotenie sa zvyčajne zakladá na kvalite, relevancii a dostupnosti dát. Tu je postup pre vyhodnotenie vhodnosti dát:

1. **Kvalita dát:** Posúďte kvalitu dát na základe ich presnosti, úplnosti, konzistencie a spoľahlivosti. Akékoľvek nedostatky, chyby alebo nekonzistencie v dátach by mohli ovplyvniť ich spoľahlivosť a presnosť. Zvážte, či sú dostupné dáta dostatočne kvalitné na podporu požadovaných analýz a rozhodovania.
2. **Relevancia dát:** Posúďte relevanciu dát vo vzťahu k cieľom analýzy. Skontrolujte, či dostupné dáta obsahujú potrebné atribúty, informácie a vzťahy, ktoré sú kľúčové pre vaše analýzy. Zvážte, či dostupné dáta poskytujú dostatočný pohľad na problém, s ktorým sa zaoberáte, a či sú relevantné pre vaše rozhodovacie procesy.
3. **Dostupnosť dát:** Skúmajte dostupnosť dát a prístup k nim. Zvážte, či máte potrebné oprávnenia a prostriedky na získanie a používanie týchto dát. Ak sú dáta obmedzené alebo nedostupné, môže to mať vplyv na vašu schopnosť ich využívať v analytickom procese.
4. **Porovnanie s požiadavkami:** Porovnajte dostupné dáta s požiadavkami a potrebami analýzy. Zhodnoťte, do akej miery tieto dáta spĺňajú stanovené kritériá kvality, relevancie a dostupnosti. Ak je nedostatok dát, nevyhovujúca kvalita alebo nesplnenie kritérií, zvážte ďalšie kroky, ako napríklad zhromaždenie ďalších dát alebo upravenie analytického prístupu.
5. **Vyhodnotenie rizík a prínosov:** Zhodnoťte riziká spojené s používaním dostupných dát, ako aj ich prínosy pre vaše analytické ciele. Zvážte, či riziká sú akceptovateľné a či prínosy z použitia dostupných dát prevyšujú potenciálne problémy.

Vyhodnotenie vhodnosti dát na základe ich kvality, relevancie a dostupnosti pomáha zabezpečiť, že v analytickom procese sa využívajú dáta, ktoré sú spoľahlivé, relevantné a dostupné pre dosiahnutie stanovených cieľov. Týmto spôsobom sa znižuje riziko chybných rozhodnutí a zabezpečuje sa presnejšia a účinnejšia analýza.

# Error! Reference source

## 3.4 Identifikácia dátových objektov

### 3.4.1 Identifikácia hlavných kategórií dátových objektov na základe stanovených cieľov

Identifikácia hlavných kategórií dátových objektov je dôležitým krokom pri určovaní štruktúry a organizácie dát pre analýzu. Táto identifikácia sa zakladá na stanovených cieľoch analýzy a pomáha zoskupiť dáta do relevantných kategórií pre ďalšie spracovanie a využitie. Tu sú niektoré hlavné kategórie dátových objektov, ktoré môžu byť identifikované na základe stanovených cieľov:

1. **Demografické údaje:** Táto kategória zahŕňa informácie o charakteristikách populácie, ako sú vek, pohlavie, vzdelanie, zamestnanie a podobne. Demografické údaje môžu byť dôležité pri analýze spotrebiteľského správania, segmentácii trhu, cielení reklamnej kampane a podobne.
2. **Finančné údaje:** Táto kategória zahŕňa informácie o finančných transakciách, príjmoch, výdavkoch, ziskoch, stratách a podobne. Finančné údaje sú dôležité pri analýze finančnej výkonnosti, prognózovaní tržieb, rozpočtovaní a podnikovom riadení.
3. **Produktové údaje:** Táto kategória zahŕňa informácie o produktoch alebo službách, ako sú vlastnosti produktu, cenové údaje, skladové zásoby, predajné údaje a podobne. Produktové údaje sú dôležité pri analýze spotrebiteľského správania, zlepšovaní produktových stratégií, riadení zásob a podobne.
4. **Zákaznícke údaje:** Táto kategória zahŕňa informácie o zákazníkoch, ako sú kontaktné údaje, histórie nákupov, preferencie, spokojnosť zákazníkov a podobne. Zákaznícke údaje sú dôležité pri analýze zákazníckej segmentácie, personalizovaných marketingových kampaní, riadení zákazníckeho vzťahu a podobne.
5. **Behové údaje:** Táto kategória zahŕňa informácie o správaní používateľov, ako sú online aktivity, kliky, hľadané frázy, sociálne médiá a podobne. Behové údaje sú dôležité pri analýze správania zákazníkov, personalizovaných odporúčaní, optimalizácii používateľskej skúsenosti a podobne.
6. **Procesné údaje:** Táto kategória zahŕňa informácie o interných procesoch a operáciách organizácie, ako sú časové údaje, výkonnosť procesov, zdroje a podobne. Procesné údaje sú dôležité pri analýze efektívnosti procesov, identifikácii slabých miest, riadení kvality a podobne.

# Error! Reference source

7. **Geografické údaje:** Táto kategória zahŕňa informácie o geografickej lokalizácii, ako sú adresy, GPS súradnice, regionálne štatistiky a podobne. Geografické údaje sú dôležité pri analýze geografického rozmiestnenia zákazníkov, identifikácii trendov v konkrétnych oblastiach, plánovaní trasy a geografickej segmentácii.
8. **Sociálne údaje:** Táto kategória zahŕňa informácie o sociálnych médiách, ako sú počet nasledovníkov, reakcie, zdieľania, komentáre a podobne. Sociálne údaje sú dôležité pri analýze reputácie značky, meraní angažovanosti, monitorovaní trendov v sociálnych médiách a podobne.
9. **Textové a obsahové údaje:** Táto kategória zahŕňa textové informácie, obsah webových stránok, blogy, recenzie zákazníkov a podobne. Textové a obsahové údaje sú dôležité pri analýze sentimentu zákazníkov, zisťovaní tém a trendov v texte, identifikácii kľúčových slov a podobne.
10. **Časové údaje:** Táto kategória zahŕňa informácie o časových údajoch, ako sú dátumy, časové značky, časové intervaly a podobne. Časové údaje sú dôležité pri analýze trendov, sezónnosti, časových radov a predikciách.

Pri identifikácii hlavných kategórií dátových objektov je potrebné zvážiť jedinečné požiadavky a ciele analýzy. Tieto kategórie slúžia ako vodítka pre organizáciu a štruktúru dát, čo umožňuje efektívne spracovanie a analýzu informácií.

## 3.4.2 Detailný popis každej kategórie dátových objektov a ich účelov v analýze

1. **Demografické údaje:** Táto kategória zahŕňa informácie o charakteristikách populácie, ako je vek, pohlavie, vzdelanie, zamestnanie a podobne. Účelom tejto kategórie v analýze je porozumieť demografickému profilu zákazníkov alebo populácie, čo umožňuje lepšie cielenie marketingových kampaní, segmentáciu trhu, personalizáciu produktov a služieb a identifikáciu trendov v závislosti od demografických faktorov.
2. **Finančné údaje:** Táto kategória zahŕňa informácie o finančných transakciách, príjmoch, výdavkoch, ziskoch, stratách a podobne. Účelom finančných údajov v analýze je hodnotiť finančnú výkonnosť organizácie, identifikovať trendy v nákladoch a príjmoch, robiť prognózy budúcich výnosov a nákladov, sledovať cash flow a riadiť finančné rozhodnutia.
3. **Produktové údaje:** Táto kategória zahŕňa informácie o produktoch alebo službách, ako sú vlastnosti produktu, cenové údaje, skladové zásoby, predajné údaje a podobne. Účelom produktových údajov v analýze je porozumieť preferenciám zákazníkov, sledovať predajné trendy, identifikovať

# Error! Reference source

najlepšie predávajúce sa produkty, zlepšovať produktovú stratégiu a riadiť skladové zásoby.

4. **Zákaznícke údaje:** Táto kategória zahŕňa informácie o zákazníkoch, ako sú kontaktné údaje, histórie nákupov, preferencie, spokojnosť zákazníkov a podobne. Účelom zákazníckych údajov v analýze je pochopiť správanie zákazníkov, vytvárať personalizované ponuky, zlepšovať zákaznícky servis, zvýšiť spokojnosť zákazníkov a vybudovať dlhodobé vzťahy so zákazníkmi.
5. **Behové údaje:** Táto kategória zahŕňa informácie o správaní používateľov, ako sú online aktivity, kliky, hľadané frázy, sociálne médiá a podobne. Účelom behových údajov v analýze je porozumieť preferenciám a potrebám zákazníkov, personalizovať obsah a odporúčania, optimalizovať používateľský zážitok, vyhľadávať trendy v správaní používateľov a zlepšovať marketingovú stratégiu.
6. **Konkurenčné údaje:** Táto kategória zahŕňa informácie o konkurencii, ako sú trhový podiel, cenová stratégia, marketingové aktivity a podobne. Účelom konkurenčných údajov v analýze je sledovať a porovnávať výkonnosť konkurencie, identifikovať konkurenčné výhody, získavať informácie pre strategické rozhodovanie a inovácie produktov.

Tieto sú len niektoré príklady kategórií dátových objektov a ich účelov v analýze. Každá kategória má svoju špecifickú hodnotu a prispieva k lepšiemu pochopeniu organizácie, zákazníkov a trhu. Je dôležité zvážiť potreby a ciele analýzy a vybrať relevantné kategórie dátových objektov, ktoré najlepšie slúžia účelu analýzy a podporujú rozhodovacie procesy.

### 3.4.3 Rozdelenie každej kategórie na konkrétne atribúty, ktoré budú analyzované

1. Demografické údaje:
  - Vek
  - Pohlavie
  - Vzdelanie
  - Zamestnanie
  - Príjmová skupina
  - Geografická lokalita
  - Rodinný stav
2. Finančné údaje:
  - Príjmy
  - Náklady
  - Zisk
  - Straty
  - Investície

# Error! Reference source

- Cash flow
  - Rozpočet
3. Produktové údaje:
- Názov produktu
  - Kategória produktu
  - Popis produktu
  - Cena
  - Dostupnosť
  - Predajné množstvo
  - Hodnotenia a recenzie
4. Zákaznícke údaje:
- Kontaktné údaje (meno, adresa, e-mail, telefón)
  - História nákupov
  - Preferované kanály komunikácie
  - Zákaznícka spokojnosť
  - Zákaznícky feedback
  - Segmentácia zákazníkov
5. Behové údaje:
- Webová aktivita (kliky, stránky navštívené, čas strávený na stránke)
  - Sociálne médiá (počet nasledovníkov, interakcie, zdieľania)
  - Hľadané frázy a kľúčové slová
  - Konverzie a ciele správania
  - Reakcie na marketingové kampane
6. Konkurenčné údaje:
- Trhový podiel konkurencie
  - Cenová stratégia konkurencie
  - Marketingové aktivity konkurencie
  - Inovácie a nové produkty konkurencie
  - Sledovanie recenzií a hodnotení konkurencie
  - Porovnanie ponuky a vlastností produktov konkurencie

Tieto atribúty predstavujú konkrétne charakteristiky v rámci každej kategórie dátových objektov, ktoré budú analyzované. Výber konkrétnych atribútov závisí od cieľov analýzy a potrieb organizácie. Je dôležité identifikovať relevantné atribúty, ktoré poskytnú relevantné a významné informácie pre analýzu a rozhodovanie.

### **3.4.4 Vyhodnotenie dátových objektov na základe ich schopnosti poskytnúť požadované informácie**

Pri vyhodnocovaní dátových objektov na základe ich schopnosti poskytnúť požadované informácie je dôležité zväžiť nasledujúce faktory:



# Error! Reference source

1. **Relevancia:** Zistite, či daný dátový objekt je relevantný pre dosiahnutie stanovených cieľov analýzy. Skontrolujte, či obsahuje potrebné informácie a či sa vzťahuje k problému alebo otázke, ktorú chcete analyzovať.
2. **Úplnosť:** Posúďte, či dátový objekt obsahuje kompletné a dostatočné informácie potrebné pre analýzu. Skontrolujte, či nechýbajú žiadne dôležité atribúty alebo údaje, ktoré by ovplyvnili správnosť alebo presnosť analýzy.
3. **Kvalita:** Vyhodnoťte kvalitu dátového objektu z hľadiska presnosti, spoľahlivosti a dôveryhodnosti informácií. Skontrolujte zdroj dát a overte, či dáta boli získané spoľahlivým spôsobom a či prešli kvalitnou kontrolou.
4. **Dostupnosť:** Zistite, či sú dátové objekty ľahko dostupné a získate k nim potrebné oprávnenia a povolenia. Skontrolujte, či sú dáta dostupné v požadovanom formáte a ak nie, či je možné ich transformovať alebo previesť do požadovaného formátu.
5. **Aktualita:** Posúďte časovú aktualitu dátových objektov. Skontrolujte, kedy boli údaje zozbierané a či sú dostatočne aktuálne na podporu rozhodovacích procesov a analýzy.
6. **Spojiteľnosť:** Skúmajte, či je možné spojiť dátové objekty s ďalšími relevantnými dátami. Vyhodnoťte, či existujú prepojenia a vzťahy medzi rôznymi dátovými objektmi, čo môže poskytnúť hlbší kontext a lepšie porozumenie analýze.
7. **Štruktúra:** Posúďte štruktúru dátových objektov a zistite, či je prehľadná a dobre organizovaná. Skontrolujte, či sú atribúty jasne definované a ak je potrebné, prispôbte štruktúru objektu tak, aby vyhovovala potrebám analýzy.

Na základe týchto kritérií môžete vyhodnotiť, aké informácie môžete získať z jednotlivých dátových objektov a či sú vhodné pre vaše potreby analýzy. Dôležité je vybrať objekty, ktoré najlepšie splnia požadované kritériá a poskytnú presné a relevantné informácie potrebné pre analytický proces.

## 3.5 Vytvorenie schémy dátových objektov

### 3.5.1 Návrh schémy dátových objektov na základe identifikovaných objektov a atribútov

Pri návrhu schémy dátových objektov je potrebné zohľadniť identifikované objekty a atribúty, aby sa zabezpečila správna organizácia a štruktúra dát. Tu je návrh schémy dátových objektov na základe uvedených informácií:

# Error! Reference source

1. Objekt: Zákazník
  - Atribúty:
    - ID zákazníka (unikátny identifikátor)
    - Meno
    - Adresa
    - E-mailová adresa
    - Telefónne číslo
    - Vek
    - Pohlavie
    - Vzdelanie
    - Zamestnanie
    - Príjmová skupina
    - Geografická lokalita
    - Rodinný stav
2. Objekt: Produkt
  - Atribúty:
    - ID produktu (unikátny identifikátor)
    - Názov produktu
    - Kategória produktu
    - Popis produktu
    - Cena
    - Dostupnosť
    - Predajné množstvo
    - Hodnotenia a recenzie
3. Objekt: Financie
  - Atribúty:
    - ID transakcie (unikátny identifikátor)
    - ID zákazníka
    - Časová pečiatka transakcie
    - Príjem/náklady
    - Suma
    - Popis transakcie
4. Objekt: Webová aktivita
  - Atribúty:
    - ID návštevy (unikátny identifikátor)
    - ID zákazníka
    - Časová pečiatka návštevy
    - Kliknuté odkazy
    - Navštívené stránky
    - Strávený čas na stránke
5. Objekt: Konkurencia
  - Atribúty:
    - ID konkurenta (unikátny identifikátor)
    - Názov konkurenta
    - Trhový podiel
    - Cenová stratégia

# Error! Reference source

- Marketingové aktivity
- Inovácie a nové produkty

Toto je iba príklad návrhu schémy dátových objektov a ich atribútov na základe identifikovaných informácií. Dôležité je prispôbiť schému podľa konkrétnych potrieb a cieľov analýzy.

## 3.5.2 Definovanie vzťahov medzi objektmi a ich atribútmi pomocou primárnych a cudzích kľúčov

Pri definovaní vzťahov medzi objektmi a ich atribútmi je možné využiť primárne a cudzie kľúče. Tu je príklad definovania vzťahov pomocou týchto kľúčov:

1. Objekt: Zákazník
  - Primárny kľúč: ID zákazníka
2. Objekt: Produkt
  - Primárny kľúč: ID produktu
3. Objekt: Financie
  - Primárny kľúč: ID transakcie
  - Cudzí kľúč: ID zákazníka (prepojenie na objekt Zákazník)
4. Objekt: Webová aktivita
  - Primárny kľúč: ID návštevy
  - Cudzí kľúč: ID zákazníka (prepojenie na objekt Zákazník)
5. Objekt: Konkurencia
  - Primárny kľúč: ID konkurenta

V tomto príklade sme vytvorili primárne kľúče pre každý objekt, ktoré jednoznačne identifikujú každý záznam v danom objekte. Navyše sme v niektorých objektoch použili cudzie kľúče, ktoré slúžia na prepojenie s inými objektmi. Napríklad v objekte Financie sme použili cudzí kľúč ID zákazníka na prepojenie s objektom Zákazník, čo umožňuje sledovať finančné transakcie konkrétneho zákazníka.

Tieto vzťahy medzi objektmi a ich atribútmi pomocou primárnych a cudzích kľúčov vytvárajú spojenie medzi rôznymi časťami dátového modelu a umožňujú efektívne vykonávanie analýz a dotazovania sa na dáta.

## 3.5.3 C. Vizualizácia schémy dátových objektov pomocou diagramov a grafických reprezentácií

# Error! Reference source

Vizualizácia schémy dátových objektov je dôležitým krokom, ktorý pomáha lepšie porozumieť štruktúre a vzťahom medzi jednotlivými objektmi. Tu sú niektoré spôsoby, ako vizualizovať schému dátových objektov pomocou diagramov a grafických reprezentácií:

1. **Entitno-relačný diagram (ER diagram):** Tento typ diagramu je široko používaný na vizualizáciu štruktúry dátových objektov a vzťahov medzi nimi. Na diagrame sú znázornené entitné objekty (napr. Zákazník, Produkt) a vzťahy medzi nimi. Priraďujú sa primárne a cudzie kľúče a ukazuje sa, ako sú objekty medzi sebou prepojené.
2. **UML diagramy:** Unified Modeling Language (UML) poskytuje rôzne typy diagramov, ktoré môžu byť použité na vizualizáciu dátových objektov. Napríklad triedny diagram UML je vhodný na zobrazenie objektov, ich atribútov a vzťahov medzi nimi. Ďalšie UML diagramy, ako napríklad aktivitný diagram alebo sekvenčný diagram, môžu byť použité na reprezentáciu procesov a toku dát v rámci dátových objektov.
3. **Grafické nástroje na modelovanie dátových objektov:** Existuje množstvo špecializovaných nástrojov a softvérov, ktoré umožňujú vytváranie a vizualizáciu dátových objektov pomocou grafických prostredí. Tieto nástroje poskytujú intuitívne rozhranie, ktoré umožňuje vytvárať objekty, pridávať atribúty, definovať vzťahy a generovať vizuálne reprezentácie schémy.
4. **Flowcharts:** Flowchart je typ diagramu, ktorý sa často používa na zobrazovanie postupov a tokov dát. Môže sa použiť na zobrazenie toku dát medzi dátovými objektmi a procesmi v analýze. Flowcharty môžu byť jednoduché a intuitívne prostriedky na vizuálnu reprezentáciu štruktúry a toku dátových objektov.

Pri vizualizácii schémy dátových objektov je dôležité vybrať si vhodný typ diagramu alebo grafického nástroja, ktorý najlepšie vyhovuje vašim potrebám a poskytuje jasný a prehľadný pohľad na štruktúru dátových objektov a ich vzťahy.

## 3.6 Implementácia a správa dátových objektov

### 3.6.1 Implementácia navrhutej schémy dátových objektov do vhodného systému alebo databázového riešenia

Implementácia navrhutej schémy dátových objektov do vhodného systému alebo databázového riešenia je dôležitým krokom, ktorý umožňuje vytvoriť fyzickú reprezentáciu dátových objektov a ich atribútov v reálnom prostredí. Tu sú niektoré kroky, ktoré sa môžu vykonať pri implementácii:

# Error! Reference source

1. **Výber vhodného systému alebo databázového riešenia:** Na základe vašich potrieb a požiadaviek vyberte vhodný systém alebo databázové riešenie, ktoré podporuje modelovanie a správu dátových objektov. Medzi najbežnejšie riešenia patria relačné databázy, ako napríklad MySQL, PostgreSQL, Oracle, alebo nereklačné databázy, ako napríklad MongoDB, Cassandra.
2. **Vytvorenie tabuliek alebo kolekcí:** V systéme alebo databázovom riešení vytvorte tabuľky alebo kolekcie pre každý dátový objekt na základe navrhutej schémy. Definujte stĺpce alebo polia pre jednotlivé atribúty objektov a zabezpečte, aby každý objekt mal priradený unikátny identifikátor (primárny kľúč).
3. **Vytvorenie vzťahov medzi tabuľkami alebo kolekciami:** Ak objekty majú medzi sebou vzťahy, definujte ich pomocou primárnych a cudzích kľúčov. Priradzovanie cudzích kľúčov zabezpečuje prepojenie medzi objektami a umožňuje dotazovanie sa a manipuláciu s dátami na základe týchto vzťahov.
4. **Implementácia prístupových metód:** V systéme alebo databázovom riešení implementujte prístupové metódy alebo funkcie, ktoré umožňujú manipuláciu s dátami. Tieto metódy môžu zahŕňať vytváranie, čítanie, aktualizáciu a mazanie (CRUD) operácie pre jednotlivé dátové objekty.
5. **Zabezpečenie a optimalizácia:** Pri implementácii je dôležité zabezpečiť prístup k dátam pomocou rôznych bezpečnostných opatrení, ako napríklad autentifikácia a autorizácia. Okrem toho sa môžu vykonávať optimalizácie databázových dotazov a indexácia, aby sa zlepšila výkonnosť pri manipulácii s veľkým objemom dát.
6. **Testovanie a ladenie:** Po implementácii je dôležité vykonať testovanie a ladenie systému alebo databázového riešenia. Overte, či sú dátové objekty správne uložené a vzťahy medzi nimi fungujú správne. Testujte aj výkonnosť a bezpečnosť systému.
7. **Správa a údržba:** Po úspešnej implementácii je dôležité pravidelne spravovať a udržiavať dátové objekty a databázové riešenie. To zahŕňa zálohovanie dát, monitorovanie výkonu a opravu prípadných chýb alebo problémov.

Implementácia navrhutej schémy dátových objektov vyžaduje dobre štruktúrovaný postup a starostlivosť o správne modelovanie a implementáciu. Dôkladný prístup k tomuto kroku zabezpečí správnu a efektívnu správu dát v rámci analytického procesu.

# Error! Reference source

## 3.6.2 Overenie presnosti a správnosti načítania dát do jednotlivých objektov

Overenie presnosti a správnosti načítania dát do jednotlivých objektov je dôležitým krokom, ktorý zabezpečuje, že dáta boli správne a úplne importované do cieľových objektov. Tu je niekoľko spôsobov, ako môžete overiť presnosť a správnosť načítania dát:

1. **Kontrola úplnosti dát:** Skontrolujte, či boli všetky potrebné dáta načítané do jednotlivých objektov. Porovnajete počet záznamov, ktoré boli importované, s očakávaným počtom záznamov zo zdroja dát. Ak je rozdiel, preskúmajte dôvod a vykonajte potrebné kroky na doplnenie chýbajúcich údajov.
2. **Kontrola kvality dát:** Skontrolujte kvalitu načítaných dát v jednotlivých objektoch. Skontrolujte formát dát, hodnoty atribútov a integritu dát. Zistíte, či sú dáta v súlade s očakávanými štandardmi a pravidlami. Ak zistíte nekonzistentné alebo neplatné dáta, vykonajte potrebné opravy alebo čistenie dát.
3. **Porovnanie s pôvodnými dátami:** Ak máte k dispozícii pôvodné dáta, porovnajete ich s načítanými dátami v jednotlivých objektoch. Skontrolujte zhodu dátových hodnôt, atribútov a vzťahov. Ak sa zistia rozdiely, preskúmajte príčinu a vykonajte korekčné opatrenia.
4. **Testovanie dotazov a operácií:** Vyhodnoťte správnosť načítaných dát pomocou testovania dotazov a operácií, ktoré sa vykonávajú nad dátovými objektmi. Overte, či dotazy vrátia očakávané výsledky a či operácie nad dátami fungujú správne. Ak zistíte nesprávne výsledky alebo chybné operácie, opravte ich a overte presnosť dát.
5. **Kolaborácia a revízia:** Zapojte do overovania presnosti a správnosti načítania dát ďalších členov tímu alebo odborníkov, ktorí majú znalosti o dátach a analytickom procese. Spoločne preskúmajte dáta, porovnajete ich s očakávaniami a diskutujte o prípadných odchýlkach či chybách. Tímovo zhodnotenie zlepšuje celkovú presnosť a správnosť načítania dát.
6. **Dokumentácia a záznamy:** Dôležité je aj správne dokumentovať výsledky overovania presnosti a správnosti načítania dát. Vytvorte si záznamy o identifikovaných chybách, opravách a vylepšeniach. Tieto dokumenty budú užitočné pre budúce referencie a údržbu dátových objektov.

# Error! Reference source

Overenie presnosti a správnosti načítania dát je kritickým krokom pri implementácii dátových objektov. Dôkladná kontrola a overovanie pomôžu zabezpečiť kvalitu a spoľahlivosť dátového modelu a analytického procesu.

### 3.6.3 Správa dátových objektov, vrátane udržiavania aktuálnych údajov, indexovania a zálohovania

Správa dátových objektov je dôležitým aspektom pre udržiavanie dátového prostredia v aktuálnom a bezpečnom stave. Zahrnuje nasledujúce činnosti:

1. **Udržiavanie aktuálnych údajov:** Je dôležité udržiavať dátové objekty v súlade s aktuálnymi dátami. To zahŕňa aktualizáciu údajov, keď sa zmenia zdrojové dáta, a zabezpečenie, aby boli všetky zmeny zachytené a odzrkadlené v dátových objektoch. Automatické alebo pravidelné procesy aktualizácie dát môžu byť implementované na základe potreby.
2. **Indexovanie dátových objektov:** Indexovanie je dôležité pre rýchle a efektívne vyhľadávanie v dátových objektoch. Vytvorenie vhodných indexov na základe atribútov a typu vyhľadávania môže výrazne zlepšiť výkon a efektivitu dopytovania nad dátovými objektmi. Je potrebné pravidelne skúmať a aktualizovať indexy na základe potrieb a rastu dátového objektu.
3. **Zálohovanie dátových objektov:** Dôležité je pravidelné zálohovanie dátových objektov na zabezpečenie obnovenia údajov v prípade straty alebo poškodenia dát. Vykonávanie pravidelných zálohovacích procesov a ukladanie záloh na spoľahlivé úložisko je kritické pre zachovanie integrovaných a nezmenených dátových objektov.
4. **Správa oprávnení:** Správa oprávnení a prístupu k dátovým objektom je nevyhnutná pre zabezpečenie bezpečnosti a ochranu dát. Definovanie a riadenie oprávnení používateľov k dátovým objektom, zabezpečenie, aby mali iba oprávnení prístup k relevantným dátam, a pravidelná revízia oprávnení sú súčasťou správy dátových objektov.
5. **Monitorovanie a optimalizácia výkonu:** Monitorovanie výkonu dátových objektov je dôležité pre identifikáciu a riešenie možných problémov s výkonom. Monitorovanie by malo zahŕňať sledovanie využitia zdrojov, ako sú pamäť, procesor a diskový priestor, a identifikáciu oblastí s nízkym výkonom, ktoré vyžadujú optimalizáciu.

# Error! Reference source

6. **Aktualizácia a rozšírenie dátových objektov:** S rastom a vývojom potrieb analytického procesu môže byť nevyhnutné aktualizovať a rozšíriť dátové objekty. To môže zahŕňať pridávanie nových atribútov, zmeny štruktúry objektov alebo rozširovanie vzťahov medzi objektmi. Dôležité je mať proces a stratégiu na aktualizáciu a rozšírenie dátových objektov bez poškodenia existujúcich dát.

Správa dátových objektov je kontinuálny proces, ktorý vyžaduje systematické monitorovanie, údržbu a optimalizáciu. Tím, ktorý sa zaoberá správou dátových objektov, by mal mať jasne definované postupy a zodpovednosti, aby zabezpečil spoľahlivosť, dostupnosť a kvalitu dát v rámci analytického procesu.



# Error! Reference source

## 4 Model komunikačných väzieb

Modely komunikačných väzieb z pohľadu spracovania údajov predstavujú teoretické a analytické nástroje, ktoré sa používajú na štúdium interakcií a vzťahov medzi rôznymi subjektmi v komunikačnom procese. Tieto modely umožňujú lepšie porozumieť spôsobu, akým sa informácie šíria a vzťahy vyvíjajú v rámci komunikácie.

Jeden zo základných modelov je model prenosu informácií, ktorý sa zameriava na štruktúru a obsah prenášaných správ. Podľa tohto modelu komunikácia začína odoslaním správy od zdroja (odosielača) k prijímateľovi (prijímačovi) prostredníctvom komunikačného kanála. Tento model zohľadňuje rôzne faktory, ako je jasnosť správy, hluk v komunikačnom kanáli a vnímanie správy prijímateľom.

V súvislosti so spracovaním údajov je dôležité spomenúť aj modely komunikácie založené na analýze siete. Tieto modely sa zameriavajú na skúmanie vzťahov medzi jednotlivými subjektmi v sieti a identifikáciu kľúčových hráčov a vzorov komunikácie. Tieto modely často využívajú metódy grafovej analýzy a štatistického spracovania údajov na identifikáciu centrality, komunitných štruktúr a vzorov šírenia informácií v komunikačných sieťach.

V oblasti verejnej správy existuje mnoho príkladov, ako modely komunikačných väzieb z pohľadu spracovania údajov môžu byť využité na zlepšenie komunikačných procesov a efektívnosť verejnej správy. Tu sú niektoré z týchto príkladov:

1. **Analýza sociálnych médií:** Modely komunikačných väzieb môžu byť využité na analýzu správ a diskusií na sociálnych médiách s cieľom získať informácie o verejnej mienke, identifikovať kľúčových aktérov a sledovať trendy a nálady voči konkrétnym témam. Na základe týchto analýz môže verejná správa prispôbiť svoje komunikačné stratégie a zameranie.
2. **Sledovanie médií:** Modely komunikačných väzieb môžu byť použité na monitorovanie spravodajských médií a hodnotenie spravodajskej krytie. Pomocou týchto modelov je možné identifikovať vplyvných novinárov, médiá s najväčším dosahom a tematické trendy. To môže pomôcť verejnej správe lepšie porozumieť a reagovať na médiá a ich úlohu pri formovaní verejnej mienky.

# Error! Reference source

3. **Analýza otvorených údajov:** Modely komunikačných väzieb môžu byť využité na analýzu otvorených údajov, ako sú verejné záznamy, štatistiky a dáta o verejných službách. Tieto analýzy môžu poskytnúť vhl'ad do komunikačných vzťahov medzi rôznymi inštitúciami a pomôcť identifikovať oblasti, kde je potrebné zlepšiť komunikáciu a spoluprácu.
4. **Prediktívne modelovanie:** Modely komunikačných väzieb môžu byť využité na predikciu a modelovanie budúcich komunikačných vzťahov a správania sa verejnosti. Tieto modely môžu byť využité na predpovedanie účinnosti komunikačných kampaní, analýzu rizík a odhalenie potenciálnych problémových oblastí v komunikácii.

Tieto sú len niektoré príklady, ako modely komunikačných väzieb môžu byť využité v oblasti verejnej správy. Ich využitie môže pomôcť zlepšiť komunikačné stratégie, identifikovať kľúčových aktérov a zefektívniť komunikáciu so zapojenými stranami a verejnosťou.

## 4.1 Spojitosť s Dátovou integráciou

Samotná integrácia údajov zahŕňa viacero aspektov, pričom modely komunikačných väzieb môžu mať vplyv na niektoré z nich. Tu je ďalší detailnejší prehľad niektorých oblastí, kde modely komunikačných väzieb môžu prispieť k dátovej integrácii:

1. **Harmonizácia štruktúry dát:** Pri integrácii údajov z rôznych zdrojov často narazíme na rozdielne formáty, schémy a štruktúry dát. Modely komunikačných väzieb môžu pomôcť pri harmonizácii týchto štruktúr a poskytnúť rámec pre identifikáciu podobností a vzájomné prepojenia medzi jednotlivými dátovými sadami.
2. **Mapovanie a porovnávanie údajov:** Modely komunikačných väzieb môžu byť využité na mapovanie údajov z rôznych zdrojov na spoločné identifikátory alebo atribúty. Týmto spôsobom je možné porovnávať a analyzovať údaje z rôznych zdrojov a identifikovať vzťahy medzi nimi. Modely môžu pomôcť pri určovaní podobnosti, zhody alebo rozdielov v údajoch a umožniť ich následnú integráciu.
3. **Detekcia a riešenie nezhôd v údajoch:** Pri dátovej integrácii sa môžu vyskytnúť nezhody a rozdiely v údajoch z rôznych zdrojov. Modely

# Error! Reference source

komunikačných väzieb môžu slúžiť na detekciu týchto nezhôd a poskytnúť mechanizmy na ich riešenie a vyrovnanie. Tieto modely môžu zahŕňať algoritmy na porovnávanie údajov, odhaľovanie chýb a konfliktov, a následné rozhodovanie o tom, ako riešiť nezhôdne údaje.

4. **Sledovanie kvality údajov:** Modely komunikačných väzieb môžu byť využité na monitorovanie kvality údajov počas procesu integrácie. Tieto modely môžu analyzovať údaje, identifikovať anomálie, chýbajúce hodnoty alebo duplikáty a poskytnúť mechanizmy na ich korekciu a zlepšenie kvality údajov.
5. **Riadenie metadát:** Modely komunikačných väzieb môžu byť využité na riadenie metadát, ktoré poskytujú informácie o údajoch a ich charakteristikách. Tieto modely môžu pomôcť pri štandardizácii a správe metadát, čo zjednodušuje identifikáciu, vyhľadávanie a správu údajov v rámci integrovaného prostredia.

Je dôležité si uvedomiť, že modely komunikačných väzieb sú jedným z mnohých nástrojov, ktoré sa používajú pri dátovej integrácii. Ich využitie závisí od špecifických potrieb a kontextu danej organizácie alebo projektu.

## 4.2 Aký to má súvis s Dátovou kanceláriou ?

Dátové kancelárie poskytujú platformu pre správu a analýzu údajov, pričom modely komunikačných väzieb môžu byť integrované do týchto kancelárií s cieľom optimalizovať komunikačné procesy a zlepšiť výkon organizácie.

Jedným z hlavných prínosov modelov komunikačných väzieb v dátových kanceláriách je analýza komunikačných vzťahov. Tieto modely môžu identifikovať a analyzovať vzťahy medzi jednotlivými členmi organizácie, zákazníkmi, partnermi a ďalšími relevantnými aktérmi. Na základe týchto analýz je možné získať lepšie pochopenie preferencií, potrieb a správania sa týchto aktérov, čo môže viesť k personalizovanejšej komunikácii a lepšej spokojnosti zákazníkov.

Integrácia modelov komunikačných väzieb do dátových kancelárií umožňuje aj efektívne zosúladenie údajov z rôznych zdrojov. Pri integrácii údajov sa často stretávame s problémami nekonzistentnosti, rozdielnych formátov a štruktúr údajov. Modely komunikačných väzieb môžu poskytnúť nástroje a metódy na harmonizáciu týchto údajov a ich transformáciu do jednotného formátu, čo

# Error! Reference source

zjednodušuje ich správu a analýzu. Týmto spôsobom sa minimalizuje riziko chýb pri integrácii a zabezpečuje sa kvalita údajov.

Ďalšou oblasťou, kde môžu modely komunikačných väzieb prispieť k dátovej integrácii, je zlepšenie procesu mapovania údajov. Pri integrácii údajov často dochádza k problémom s rovnakými entitami, ktoré sú reprezentované rôznymi identifikátormi alebo atribútmi. Modely komunikačných väzieb môžu pomôcť identifikovať a riešiť tieto problémy, umožňujúc presné mapovanie údajov a ich vzájomné prepojenie. To má za následok zlepšenú kvalitu a spoľahlivosť integrovaných údajov.

Okrem toho, modely komunikačných väzieb môžu tiež prispieť k detekcii a riešeniu nezhôd v údajoch. Pri integrácii údajov sa môžu vyskytnúť nekonzistentnosti, chýbajúce hodnoty alebo duplikáty údajov. Modely komunikačných väzieb môžu slúžiť ako nástroje na identifikáciu týchto problémov a poskytnúť mechanizmy na ich riešenie. To zahŕňa algoritmy na porovnávanie údajov, odhaľovanie chýb a konfliktov a rozhodovanie o tom, ako riešiť nezhôdne údaje. Týmto spôsobom sa zabezpečuje konzistentnosť a integrita údajov v dátových kanceláriách.

Modely komunikačných väzieb môžu tiež prispieť k automatizácii komunikačných procesov v dátových kanceláriách. Na základe týchto modelov je možné predikovať preferencie a správanie sa zákazníkov, personalizovať komunikáciu a poskytovať relevantné informácie a služby. Tým sa zvyšuje efektivita komunikácie a zlepšuje sa skúsenosť zákazníkov. Modely komunikačných väzieb môžu tiež pomôcť pri automatizácii správy údajov, napríklad pri automatickom zasielaní notifikácií, spracovaní žiadostí alebo aktualizácii profilov zákazníkov.

Sledovanie výkonu a správa údajov je ďalšou oblasťou, kde sa modely komunikačných väzieb a dátové kancelárie prepojili. Modely komunikačných väzieb môžu byť využité na zhromažďovanie a analýzu údajov o úspešnosti kampaní, interakciách so zákazníkmi a ďalších komunikačných metrikách. Tieto informácie sa potom môžu použiť na meranie výkonu, identifikáciu oblastí na zlepšenie a lepšie rozhodovanie na základe dát.

Výhody prepojenia modelov komunikačných väzieb s dátovými kanceláriami sú evidentné. Organizácie môžu získať lepšie porozumenie komunikačných vzťahov, zosúladiť údaje z rôznych zdrojov, zlepšiť mapovanie údajov, riešiť nezhôdne údaje a automatizovať komunikačné procesy. Tým sa zvyšuje efektivita, presnosť a kvalita spracovania údajov v organizácii.

# Error! Reference source

V závere môžeme povedať, že prepojenie modelov komunikačných väzieb s dátovými kanceláriami prináša významné výhody a možnosti pre efektívne spracovanie a správu údajov v organizáciách. Toto prepojenie poskytuje lepšie porozumenie komunikačným vzťahom, zlepšuje integritu a konzistentnosť údajov, zvyšuje efektivitu komunikačných procesov a umožňuje monitorovanie výkonu a správu údajov. S narastajúcim objemom údajov a potrebou ich efektívneho využitia je prepojenie modelov komunikačných väzieb a dátových kancelárií neodmysliteľnou súčasťou moderného dátového spracovania a správy údajov.

## 4.3 Prepojenie modelu komunikačných väzieb so znalostnými grafmi

Znalostné grafy majú tiež svoje miesto v kontexte spracovania údajov a dátovej integrácie. Znalostné grafy sú modely reprezentujúce znalosti a vzťahy medzi entitami vo forme grafových štruktúr. Tieto grafy sa skladajú z uzlov, ktoré reprezentujú jednotlivé entity, a hrán, ktoré reprezentujú vzťahy medzi týmito entitami.

Prepojenie modelov komunikačných väzieb a znalostných grafov môže poskytnúť silný nástroj na lepšie porozumenie komunikačných vzťahov a ich prepojenie s ďalšími relevantnými entitami. Znalostné grafy môžu byť použité na modelovanie a reprezentáciu vzťahov medzi jednotlivými aktérmi, ako aj ich vlastností a atribútov. Týmto spôsobom sa získava holistický pohľad na komunikačné vzťahy v rámci organizácie alebo projektu.

Znalostné grafy môžu tiež slúžiť ako úložisko pre metadáta a informácie o údajoch. Uzly grafu môžu reprezentovať rôzne typy údajov, ako napríklad tabuľky, dokumenty, obrázky alebo iné formáty. Hrany medzi týmito uzlami môžu reprezentovať vzťahy medzi týmito údajmi, ako napríklad referencie, prepojenia alebo vlastnícke vzťahy. Týmto spôsobom sa zabezpečuje jednotné a štruktúrované uloženie údajov a zlepšuje sa ich správa a vyhľadávanie.

V kombinácii s modelmi komunikačných väzieb môžu znalostné grafy poskytnúť bohatý kontext pre analýzu a spracovanie údajov. Analytické nástroje a algoritmy môžu byť aplikované na tieto grafy s cieľom identifikovať vzorce, odhaliť vzťahy medzi entitami a generovať nové poznatky. Týmto spôsobom sa posilňuje schopnosť organizácie porozumieť komunikačným vzťahom a využiť ich na lepšie rozhodovanie a optimalizáciu procesov.

Okrem toho, znalostné grafy môžu tiež pomôcť pri riešení problémov spojených s dátovou integráciou. Vďaka svojej grafovej štruktúre môžu znalostné grafy efektívne reprezentovať a spravovať vzťahy medzi údajmi z rôznych zdrojov. Modely komunikačných väzieb môžu poskytnúť cenné informácie o týchto

# Error! Reference source

vzťahoch, ktoré sa potom môžu integrovať do znalostných grafov. Týmto spôsobom sa zabezpečuje lepšia konzistencia a súvislosť medzi integrovanými údajmi.

V záverečnom dôsledku je prepojenie modelov komunikačných väzieb s dátovými kancelármi a znalostnými grafmi veľmi prospešné pre efektívne spracovanie a správu údajov. Spájanie týchto modelov umožňuje získať lepšie porozumenie komunikačných vzťahov, harmonizovať údaje z rôznych zdrojov, riešiť nezhodne údaje a vytvárať nové poznatky prostredníctvom analýzy znalostných grafov. Toto prepojenie poskytuje organizáciám silný nástroj na efektívne spravovanie, analýzu a využitie údajov v dnešnom dynamickom a informačne náročnom prostredí

Ďalšou dôležitou oblasťou, kde môžu modely komunikačných väzieb a znalostné grafy spolupracovať, je automatizácia spracovania a vyhľadávania informácií. Znalostné grafy môžu slúžiť ako základná štruktúra pre vyhľadávanie a indexovanie údajov. Modely komunikačných väzieb môžu byť využité na zlepšenie relevancie a presnosti vyhľadávania v rámci týchto grafov. Pomocou analýzy komunikačných vzťahov môžu byť identifikované kľúčové vzťahy a vzorce, ktoré sa potom môžu aplikovať na vyhľadávanie relevantných informácií v grafoch. Tým sa zvyšuje efektivita vyhľadávacieho procesu a zlepšuje sa prístup k dôležitým informáciám.

V kontexte znalostných grafov a dátových kancelárií môžu byť modely komunikačných väzieb využité aj na vytvorenie odporúčacích systémov. Analýza komunikačných vzťahov môže poskytnúť dôležité informácie o preferenciách, záujmoch a správaní používateľov. Tieto informácie sa následne môžu integrovať do znalostných grafov a využiť na generovanie personalizovaných odporúčaní. Napríklad, na základe analýzy komunikačných vzťahov a vzorcov správania sa zákazníkov, môžu byť odporúčané relevantné produkty, služby alebo ďalší relevantní aktéri. Týmto spôsobom sa zvyšuje hodnota a účinnosť odporúčacích systémov v dátových kanceláriách.

Okrem toho, kombinácia modelov komunikačných väzieb a znalostných grafov môže prispieť aj k lepšiemu riadeniu informácií a zabezpečeniu súladu s regulačnými požiadavkami. Modely komunikačných väzieb môžu byť použité na analýzu a monitorovanie toku informácií medzi rôznymi aktérmi. Tieto informácie sa potom môžu integrovať do znalostných grafov, kde môžu byť sledované a kontrolované v súlade s príslušnými regulačnými požiadavkami. Tým sa zlepšuje transparentnosť a zodpovednosť pri spracovaní a správe údajov v organizácii.

V konečnom dôsledku spojenie modelov komunikačných väzieb a znalostných grafov prispieva k lepšiemu porozumeniu vzťahov a dát v organizácii,

# Error! Reference source

efektívnemu vyhľadávaniu a spracovaniu informácií, personalizovanej komunikácii a zvýšenej bezpečnosti údajov. Táto kombinácia modelov poskytuje organizáciám výhodu konkurenčnej výhody v dnešnom datovo náročnom a dynamickom prostredí. Spracovanie údajov a správa informácií sú kľúčovými faktormi úspechu pre organizácie v súčasnej digitálnej ére, a preto je prepojenie modelov komunikačných väzieb a znalostných grafov významným a perspektívnym smerom v oblasti spracovania údajov.

V kontexte spracovania údajov a dátovej integrácie je dôležité zdôrazniť niektoré konkrétne príklady využitia modelov komunikačných väzieb a znalostných grafov v oblasti verejnej správy a dátovej integrácie.

Verejná správa zahŕňa široké spektrum organizácií a inštitúcií, ktoré spravujú a spracovávajú veľké množstvo údajov. Príklady využitia modelov komunikačných väzieb a znalostných grafov v verejnej správe môžu zahŕňať:

- 1. Analýza komunikačných vzťahov:** Modely komunikačných väzieb môžu byť využité na analýzu vzťahov medzi rôznymi inštitúciami, orgánmi verejnej správy a občanmi. Tieto vzťahy môžu byť zobrazené v znalostných grafoch, čo umožňuje lepšie porozumenie a vizualizáciu komplexných vzťahov. Napríklad, analýza komunikačných vzťahov medzi úradmi sociálneho zabezpečenia a občanmi by mohla pomôcť identifikovať oblasti s vysokým výskytom problémov a umožniť efektívnejšie riešenie týchto problémov.
- 2. Vyhľadávanie informácií:** Znalostné grafy môžu slúžiť ako úložisko údajov a metadát v rámci verejnej správy. Modely komunikačných väzieb môžu byť využité na zlepšenie vyhľadávania a prístupu k týmto informáciám. Napríklad, vyhľadávanie v grafoch môže umožniť efektívne vyhľadávanie relevantných dokumentov, legislatívy alebo iných dôležitých informácií v rámci verejných inštitúcií.
- 3. Dátová integrácia:** Pri dátovej integrácii je často potrebné spojiť údaje z rôznych zdrojov a formátov. Modely komunikačných väzieb môžu byť využité na zlepšenie mapovania a integrácie údajov. Napríklad, ak existujú rôzne databázy a systémy v rámci verejnej správy, modely komunikačných väzieb môžu pomôcť identifikovať relevantné vzťahy medzi týmito údajmi a umožniť ich efektívne spojenie.
- 4. Správa a sledovanie dátových tokov:** Vzhľadom na rozsiahlosť a zložitosť údajov v rámci verejnej správy je dôležité mať prehľad o tokoch údajov a ich využití. Modely komunikačných väzieb a znalostné grafy môžu byť využité na

# Error! Reference source

správu a sledovanie dátových tokov v rámci organizácie. Týmto spôsobom je možné identifikovať, kto sú relevantní aktéri, ako sa údaje premiestňujú a aké sú ich vzťahy a transformácie.

- 5. Riadenie a monitorovanie procesov:** Modely komunikačných väzieb a znalostné grafy môžu poskytnúť cenné informácie pre riadenie a monitorovanie procesov v verejnej správe. Tieto modely môžu odhaliť slabé miesta, nedostatky alebo neefektívne komunikačné vzťahy, čo umožňuje zlepšiť procesy a rozhodovanie v rámci organizácie.

Využitie modelov komunikačných väzieb a znalostných grafov v kontexte verejnej správy a dátovej integrácie môže prispieť k lepšiemu porozumeniu údajov, optimalizácii procesov, zlepšeniu prístupu k informáciám a zvýšeniu efektivity verejnej správy ako celku. Tieto modely môžu slúžiť ako silný nástroj pre zlepšenie výkonu a rozhodovania v rámci verejných inštitúcií a prispieť k vytvoreniu inteligentnej a efektívnej verejnej správy.

Okrem vyššie uvedených príkladov existuje ešte niekoľko ďalších spôsobov, ako môžu modely komunikačných väzieb a znalostné grafy prispieť k dátovej integrácii a efektívnemu spracovaniu údajov v kontexte verejnej správy.

- 6. Identifikácia dátových vzorcov a trendov:** Modely komunikačných väzieb môžu byť využité na identifikáciu vzorcov a trendov v údajoch. Napríklad, analýza komunikačných vzťahov medzi občanmi a verejnými inštitúciami môže odhaliť opakujúce sa vzorce žiadostí, sťažností alebo požiadaviek. Tieto informácie môžu byť následne využité na zlepšenie procesov, poskytnutie lepšej služby občanom a identifikáciu oblastí, ktoré vyžadujú zlepšenie.
- 7. Spájanie údajov z rôznych zdrojov:** Verejná správa často pracuje s údajmi z viacerých zdrojov, ako sú rôzne databázy, webové stránky, formuláre a ďalšie. Modely komunikačných väzieb môžu pomôcť pri spájaní a harmonizácii týchto údajov. Týmto spôsobom je možné vytvoriť jednotný pohľad na údaje a zabezpečiť konzistentné a spoľahlivé spracovanie a analýzu údajov.
- 8. Správa identít a prístupových práv:** Znalostné grafy môžu slúžiť aj na správu identít a prístupových práv v rámci verejnej správy. Modely komunikačných väzieb môžu pomôcť identifikovať vzťahy medzi rôznymi



# Error! Reference source

identitami, rolami a oprávneniami. Tieto informácie sa potom môžu využiť na správu a riadenie prístupov k dôležitým údajom a zdrojom v organizácii.

- 9. Prevencia a detekcia podvodov:** Modely komunikačných väzieb môžu byť využité na prevenciu a detekciu podvodov v rámci verejnej správy. Analyzovaním komunikačných vzťahov a vzorcov správania sa môžu identifikovať podozrivé vzory alebo nezrovnalosti. Tieto informácie môžu slúžiť ako východiskový bod pre ďalšie investigatívne činnosti a opatrenia na ochranu pred podvodnými aktivitami.

Využitie modelov komunikačných väzieb a znalostných grafov v dátovej integrácii a spracovaní údajov v verejnej správe prináša množstvo výhod. Zlepšuje sa porozumenie dát, zvyšuje sa efektivita spracovania a analýzy údajov, poskytujú sa lepšia personalizácia komunikácie a zabezpečuje sa vyššia bezpečnosť údajov. To všetko vedie k lepšej kvalite služieb pre občanov, zlepšeniu rozhodovacieho procesu a efektívnejšej verejnej správe ako celku.

Práca s údajmi inak povedané aj dátový manažment je širší, ktorý zahŕňa správu dát a informácií v organizácii. Zahŕňa rôzne aspekty, ako je zber, ukladanie, organizovanie, spracovanie a analýza dát s cieľom dosiahnuť efektívne využitie dát pre podporu rozhodovania a dosiahnutie obchodných cieľov.

Vo svete správy údajov existujú dva primárne typy systémov DBMS sú systém riadenia relačnej databázy (RDBMS) a systém riadenia grafovej databázy (GDBMS). Najmä relačné databázy sú vo verejnej správe značne rozšírenejšie a majú svoje opodstatnenie. V tejto kapitole podrobne rozoberieme rozdiely a podobnosti medzi týmito dvoma druhmi databázových systémov.

Systémy správy databáz (DBMS) sú všadeprítomné časti softvéru, ktoré umožňujú aplikáciám všetkých typov a veľkostí efektívne ukladať a získavať údaje.

Dnes existujú stovky rôznych systémov DBMS, z ktorých každý je optimalizovaný pre rôzne prípady použitia, od webových aplikácií po systémy riadenia vzťahov so zákazníkmi a všetko medzi tým.

Najzásadnejší rozdiel medzi RDBMS a GDBMS je spôsob, akým reprezentujú údaje.

V RDBMS sú údaje uložené v tabuľkách, pričom každá má stĺpce a riadky. Tieto tabuľky spolupracujú pomocou kľúčov na vytváranie vzťahov medzi tabuľkami.

# Error! Reference source

Na

rozdiel od toho GDBMS ukladá informácie so štruktúrou podobnou grafu. Dáta sú reprezentované pomocou uzlov, hrán a vlastností. Tieto uzly môžu byť čokoľvek od podnikateľského subjektu po osobu, miesto, produkt alebo dokonca koncept, zatiaľ čo okraje predstavujú vzťahy, ktoré majú so susednými uzlami.

Ďalším kritickým bodom, ktorý je potrebné zvážiť, sú možnosti dotazovania oboch systémov.

S RDBMS môžeme získavať údaje pomocou SQL dotazov a s GDBMS využívame grafové dotazy. SQL dotazy sú štruktúrované takým spôsobom, že vyžadujú jasnú, zrejmú a zdokumentovanú schému. To uľahčuje priame používanie systému pri manipulácii a spracovaní údajov pomocou SQL dotazov v administračnom rozhraní.

Na druhej strane, grafové dotazy sú flexibilnejšie, pokiaľ ide o získavanie údajov.

Pomocou GDBMS môžeme dopytovať údaje pomocou ľubovoľných kľúčov alebo atribútov na prechádzanie grafom bez ohľadu na to, či sú údaje uzamknuté do definovaných schém. Dotazy na grafy je možné prispôbiť tak, aby vyhovovali potrebám konkrétnych údajov.

Oba systémy dokážu spracovať obrovské množstvo údajov; škálujú sa však rôznymi spôsobmi.

Systémy RDBMS sa škálujú vertikálne, čo znamená inováciu hardvéru servera pridaním väčšieho procesora, pamäte RAM alebo úložného priestoru na zvýšenie výkonu.

Na druhej strane systémy GDBMS sa škálujú horizontálne, čo znamená pridanie ďalších serverov na rozloženie pracovného zaťaženia v rámci klastra. Toto sa nazýva sharding, ktorý zahŕňa rozdelenie údajov na menšie časti a replikáciu na viacerých počítačoch a umožňuje GDBMS prispôbiť sa väčšiemu pracovnému zaťaženiu.

V tabuľke uvedenej nižšie je zobrazené prehľadné porovnanie výhod /nevýhod RDBMS a GDBMS.

Výhody RDBMS	Nevýhody RDBMS
<b>Integrita a konzistentnosť údajov:</b> Systémy RDBMS sú navrhnuté tak, aby zabezpečili konzistentnosť a	<b>Obmedzená škálovateľnosť:</b> RDBMS sú zvyčajne vertikálne škálovateľné, čo znamená, že ich

# Error! Reference source

spoľahlivosť údajov, čím sa zabezpečí, že kritické prvky údajov zostanú nedotknuté a zachovávajú si svoje vzťahy v rámci celého systému.	možno zväčšiť iba pridaním ďalších zdrojov na rovnaký server. Ich škálovanie môže byť nákladné, pretože požiadavky na úložisko a pamäť presahujú kapacitu jedného servera.
<b>Zabudovaná bezpečnosť:</b> RDBMS majú komplexný bezpečnostný model s funkciami, ako je riadenie prístupu, autentifikácia používateľov a izolácia údajov.	Obmedzená flexibilita: RDBMS sa spoliehajú na preddefinované statické schémy, ktoré obmedzujú typ údajov, ktoré možno uložiť.
<b>Štruktúrované údaje:</b> Štruktúrované údaje sú užitočné, pretože sa ľahšie spravujú a analyzujú.	Obmedzený výkon: RDBMS nie sú optimalizované na spracovanie extrémne zložitých súborov údajov, pretože to môže byť časovo náročné.
<b>Štandardizovaný dotazovací jazyk:</b> SQL je štandardný jazyk na interakciu s RDBMS. Táto štandardizácia umožňuje organizácii potenciálne využívať viacero RDBMS s minimálnym časom vývoja.	
<b>Výkon:</b> RDBMS sú navrhnuté tak, aby pracovali s preddefinovanými typmi údajov, ktoré do určitej miery umožňujú optimalizáciu výkonu prostredníctvom efektívneho indexovania a príbuzných osvedčených postupov.	

Tabuľka 4: Výhody / nevýhody RDBMS

Výhody GDBMS	Nevýhody GDBMS
<b>Flexibilné dátové modelovanie:</b> Grafický model môže reprezentovať akúkoľvek dátovú schému a vyvíjať ju v reálnom čase, keď s ľahkosťou vznikajú nové vzťahy, čo poskytuje bezprecedentnú flexibilitu.	<b>Komplexný dotazovací jazyk:</b> Dopytovanie údajov z grafovej databázy môže byť zložitejšie ako dotazy SQL a vyžaduje si dobré pochopenie teórie grafov.
<b>Rýchlosť:</b> Grafický model je vhodný pre rýchle dopyty na údaje a môže	<b>Bezpečnostné výzvy:</b> Grafové databázy sú menej bezpečné v

# Error! Reference source

vykonávať rýchly prechod na malých aj veľkých súboroch údajov.	porovnaní so systémami RDBMS len kvôli ich relatívnej novosti a nedostatku vnímanej bezpečnosti.
<b>Veľké množiny údajov:</b> Systémy GDBMS vynikajú v manipulácii s veľkými množinami údajov, a preto dokážu zvládnuť vysokorýchlostné spracovanie transakcií a interakcií v zložitých sieťach.	
<b>Natívna vizualizácia údajov a vzťahov:</b> Vďaka svojej grafickej povahe uľahčujú databázy grafov vizualizáciu údajov a pochopenie vzťahov medzi údajovými bodmi, čo umožňuje efektívnu a rýchlu analýzu údajov.	
<b>Horizontálna škálovateľnosť:</b> GDBMS sú horizontálne škálovateľné a môžu byť distribuované na mnohých serveroch, vďaka čomu sú vybavené na spracovanie oveľa väčších súborov údajov ako RDBMS rovnakej veľkosti.	

*Tabuľka 5Výhody / nevýhody GDBMS*

RDBMS aj GDBMS ponúkajú rôzne silné stránky a kvality, ktoré môžu byť užitočné pre konkrétne prípady použitia. Vo všeobecnosti sú systémy RDBMS vhodné na spracovanie štruktúrovaných údajov pre aplikácie, ktoré mnohí ľudia vyžadujú. Na rozdiel od toho sú systémy GDBMS zostavené tak, aby zvládali efektívne, flexibilné a rýchle spracovanie zložitých údajov zapojených do rozsiahlych súborov údajov. Výber systému, ktorý sa má použiť, bude závisieť od potrieb dátových projektov organizácie. Aby sa organizácie mohli informovane rozhodnúť, mali by začať podrobne skúmať odlišné požiadavky konkrétnych prípadov použitia a vyhodnotiť možnosti systému, nákladovú efektívnosť, škálovateľnosť a architektúry a stratégie nasadenia.

## Prípady použitia

Nakoniec, pri hodnotení oboch typov databáz by analytici mali zvážiť ich optimálne aplikácie. RDBMS sú ideálne pre údaje s predvídateľnou štruktúrou, ako sú finančné alebo zákaznícke informácie; GDBMS sú optimálne pre neštruktúrované a pološtruktúrované údaje vrátane sociálnych sietí,

# Error! Reference source

odporúčacích nástrojov, systémov na detekciu podvodov a riadenia dodávateľského reťazca.

# Error! Reference source

## 5 Meranie komunikačných väzieb

### 5.1 Možné softvérové nástroje pre meranie

Pre meranie a spracovanie údajov v oblasti modelov komunikačných väzieb a znalostných grafov existuje niekoľko softvérových nástrojov, ktoré môžete použiť. Tu je niekoľko populárnych nástrojov v tejto oblasti:

1. **Gephi** je open-source nástroj na analýzu a vizualizáciu sietí a grafov. Poskytuje rôzne algoritmy na analýzu grafu a mnoho možností pre vizualizáciu a manipuláciu s dátami. Gephi je populárny open-source nástroj na analýzu a vizualizáciu sietí a grafov. Jeho hlavným cieľom je poskytnúť užívateľom prostriedky na objavovanie, analyzovanie a prezentovanie vzorov v komplexných grafových dátach. Gephi umožňuje importovať a exportovať dáta vo formátoch ako CSV, GML, GraphML, Excel a ďalších. Poskytuje širokú škálu nástrojov pre manipuláciu s grafmi, ako aj algoritmy pre detekciu komunit, centrality, prehľadávanie grafov a vizualizáciu. Používateľské rozhranie je intuitívne a umožňuje rýchle interaktívne skúmanie a manipuláciu s dátami. Gephi je často používaný v oblasti sociálnych sietí, bioinformatiky a výskumu komplexných sietí.
2. **NetworkX** je populárna Python knižnica na prácu s grafovými štruktúrami a analýzu sietí. Je navrhnutá tak, aby bola jednoduchá na používanie a pritom flexibilná a výkonná. NetworkX poskytuje rozsiahlu sadu nástrojov na vytváranie, manipuláciu a analýzu grafov. Umožňuje vytváranie grafu pomocou rôznych typov dátových štruktúr, ako sú zoznamy, matice susednosti alebo susedné zoznamy. Knižnica obsahuje mnoho algoritmov na detekciu komunit, centrality, vyhľadávanie najkratších ciest a ďalšie. NetworkX je rozšíriteľný a umožňuje vytváranie vlastných algoritmov a funkcionalít. Je často používaný v akademickom prostredí a vo výskume v oblasti komplexných sietí.
3. **Cytoscape** je silný nástroj na analýzu a vizualizáciu sietí, najmä v kontexte biologických sietí. Jeho hlavným zameraním je umožniť užívateľom interaktívne vizualizovať a manipulovať s veľkými a zložitými sieťami. Cytoscape ponúka bohaté možnosti vizualizácie, vrátane rôznych algoritmov na usporiadanie grafu, farbenie a zobrazenie atribútov vrcholov a hrán. Okrem vizualizácie poskytuje aj rozsiahlu sadu nástrojov na analýzu sietí, vrátane algoritmov na detekciu komunit, centrality, spájanie sietí a vyhľadávanie

# Error! Reference source

vzorov. Cytoscape je ľahko rozšíriteľný pomocou pluginov a rozšírení, ktoré pridávajú nové funkcie a možnosti analýzy.

4. **Neo4j** je grafová databáza založená na jazyku Cypher, ktorá je optimalizovaná pre ukladanie a spracovanie grafových dát. Jej silnou stránkou je vysoká výkonnosť pri manipulácii so sieťami a veľkými grafovými dátami. Neo4j poskytuje robustné dotazovacie jazyky a aplikačné rozhranie, ktoré umožňujú efektívne vykonávanie operácií s grafovými štruktúrami. Databáza obsahuje širokú škálu funkcií, vrátane dotazovania cez hĺbku grafov, vyhľadávania vzorov, detekcie komunit a centrality. Je vhodná pre prípady, kde je dôležité rýchle a efektívne spracovanie grafových dát, ako napríklad v oblasti sociálnych sietí, dopravy, geografie a podobne.
5. **Apache Spark GraphX** je komponent distribuovaného spracovania údajov v Apache Spark, ktorý je zameraný na prácu s grafovými dátami. GraphX poskytuje rozšírenia Sparku pre vytváranie, manipuláciu a analýzu veľkých grafových štruktúr. Poskytuje distribuovanú implementáciu mnohých grafových algoritmov, ako sú PageRank, spojenie komponentov, šírenie vlny a ďalšie. GraphX je optimalizovaný pre vysoký výkon a škálovateľnosť, čo umožňuje efektívne spracovanie a analýzu veľkých grafových dát na klastru serverov. Je to nástroj vhodný pre prípady, kde je potrebné spracovať obrovské objemy grafových dát a dosiahnuť vysokú výkonnosť.

Tieto nástroje poskytujú rôzne funkcie pre prácu s modelmi komunikačných väzieb a znalostnými grafmi. Výber konkrétneho nástroja závisí od vašich potrieb, technických požiadaviek a úrovne zložitosti vášho výskumu.

## 5.1.1 Prepojenie IS CSRÚ a Neo4j

Neo4j a IS CSRÚ (Informačný systém Centrálnej štátnej správy) sú dva rôzne nástroje, ktoré majú odlišné účely a zameranie v kontexte dátovej kancelárie. Nižšie uvedený je stručný prehľad oboch nástrojov:

**IS CSRÚ:** Informačný systém Centrálnej štátnej správy (IS CSRÚ) je informačný systém, ktorý je zameraný na správu a integráciu dát v rámci verejnej správy. Je navrhnutý tak, aby poskytoval podporu a funkcionality pre efektívne riadenie a správu verejných služieb, administratívne procesy a komunikáciu medzi rôznymi úradmi a inštitúciami. IS CSRÚ zahŕňa rôzne moduly a komponenty, ktorých cieľom je zjednodušiť a optimalizovať procesy a poskytovať efektívne služby pre občanov a inštitúcie.

# Error! Reference source

Neo4j: Neo4j je grafová databáza, ktorá je navrhnutá na ukladanie, správu a dotazovanie veľkých množstiev grafových dát. Je založená na grafovej štruktúre, kde dáta sú reprezentované ako uzly a hrany, a umožňuje modelovať vzťahy a atribúty medzi týmito entitami. Neo4j poskytuje bohatú funkcionálnu pre dotazovanie a analýzu grafových dát a je často využívaná pre rôzne aplikácie, ako sú sociálne siete, dopravné systémy, doporučovacie systémy a podobne. V kontexte dátovej kancelárie môže byť Neo4j využitá na ukladanie a analýzu vzťahov medzi rôznymi entitami a dátami, čo umožňuje efektívne vyhľadávanie a spracovanie informácií.

Pri pohľade na spojenie Neo4j a IS CSRÚ je možné vidieť potenciálnu synergickú výhodu. Využitie grafových databáz ako Neo4j v rámci IS CSRÚ by mohlo zlepšiť správu a analýzu vzťahov medzi rôznymi údajmi a entitami v systéme. Pomocou grafov by bolo možné efektívne modelovať a vyhľadávať vzťahy medzi obyvateľmi, právnymi predpismi, projekty a inými dátami. To by umožnilo rýchlejšie a presnejšie spracovanie informácií a poskytovanie služieb.

Prepojenie medzi Neo4j a IS CSRÚ (Informačný systém Centrálnej štátnej správy) môže priniesť viaceré výhody a prínosy v kontexte efektívnejšej správy a analýzy dát v rámci verejnej správy. Tu je prehľad možného prepojenia a jeho výhod:

- 1. Ukladanie vzťahových dát v Neo4j:** IS CSRÚ často pracuje s množstvom vzťahových dát, ako sú vzťahy medzi občanmi, organizáciami, inštitúciami a ďalšími entitami. Využitie Neo4j ako grafovej databázy umožní efektívne uloženie a správu týchto vzťahov, čo umožní rýchlejšie vyhľadávanie a analýzu informácií. Databázový model Neo4j je založený na uzloch a hranách, čo sa dobre hodí pre reprezentáciu a spracovanie vzťahových dát.
- 2. Integrácia s IS CSRÚ:** Neo4j môže byť integrovaná s existujúcimi komponentmi a modulmi IS CSRÚ. Týmto spôsobom je možné využívať silné stránky oboch nástrojov a poskytnúť celkovú lepšiu funkcionálnu. Napríklad, dáta uložené v Neo4j môžu byť prístupné a používané v rámci IS CSRÚ na vykonávanie rôznych úloh, ako sú vyhľadávanie, analýza a správa dát. Taktiež je možné využiť Neo4j na vytváranie grafových vizualizácií pre lepšiu prehľadnosť a pochopenie dát.
- 3. Zlepšenie interoperability:** Integrácia Neo4j s IS CSRÚ môže pomôcť zlepšiť interoperabilitu medzi rôznymi systémami, modulmi a zdrojmi dát v rámci e-governmentu. Neo4j môže slúžiť ako centrálné úložisko pre vzťahové dáta a poskytnúť flexibilný a efektívny spôsob na modelovanie a spracovanie týchto dát. Týmto spôsobom je možné zjednodušiť a optimalizovať výmenu a zdieľanie informácií medzi rôznymi komponentmi IS CSRÚ a inými systémami, čo prispieva k lepšej integrácii dát a ich správe.



# Error! Reference source

4. **Vytváranie analytických modelov:** Neo4j poskytuje silné nástroje pre analýzu a vizualizáciu grafových dát. Využitím Neo4j v kontexte IS CSRÚ je možné vytvárať analytické modely, ktoré umožňujú odhaľovať skryté vzory, získavať poznatky a vyvodiť závery z dát. Tieto analytické modely môžu poskytnúť dôležité informácie pre rozhodovacie procesy v rámci e-governmentu a pomôcť pri plánovaní a optimalizácii poskytovania verejných služieb.

Využitie Neo4j v rámci IS CSRÚ môže teda priniesť výhody ako lepšie modelovanie vzťahov, zlepšenie interoperability, možnosti vyhľadávania a analýzy vzťahov medzi dátami a vytváranie analytických modelov. Týmto spôsobom by sa mohlo dosiahnuť efektívnejšie riadenie a poskytovanie služieb v rámci e-governmentu.

## 5.1.2 Nástroj Graphlytic

Graphlytic je webová aplikácia Java poskytujúca užívateľsky prívetivé, vysoko prispôsobiteľné prostredie na vizualizáciu grafov na spoločné skúmanie a analýzu grafov.

Jednou z najväčších výhod Graphlyticu je, že je dostupný cez webový prehliadač, čo znamená, že používatelia k nemu môžu pristupovať z akéhokoľvek zariadenia s internetovým pripojením. Vďaka tomu je vhodný pre distribuované tímy alebo jednotlivcov, ktorí potrebujú kontrolovať a analyzovať údaje z grafov na diaľku. Okrem toho Graphlytic poskytuje podporu technickým aj netechnickým používateľom, čo umožňuje používateľom vyťažiť z platformy maximum bez ohľadu na ich technické zázemie.

Stručne povedané, Graphlytic je nástroj na vizualizáciu a analýzu grafov, ktorý poskytuje širokú škálu pokročilých možností grafovej analýzy spolu s rozsiahlou prispôsobiteľnosťou a intuitívnym rozhraním. Jeho webový charakter a užívateľsky prívetivý dizajn ho sprístupňujú širokému okruhu používateľov bez ohľadu na predchádzajúce technické znalosti. V konečnom dôsledku Graphlytic pomáha používateľom skúmať a porozumieť ich údajom novým a jedinečným spôsobom, čo v konečnom dôsledku vedie k lepším prehľadom a informovaným obchodným rozhodnutiam.

Pomáha podnikom či už v súkromnom sektore, alebo v prostredí verejnej správy, výskumníkom a vývojárom preskúmať a pochopiť zložité dátové štruktúry, ako sú sociálne siete, IT infraštruktúry a obchodné procesy. Importovaním množín údajov z rôznych zdrojov prostredníctvom intuitívneho rozhrania umožňuje

# Error! Reference source

Graphlytic používateľom vytvárať interaktívne, prispôsobiteľné vizualizácie ich údajov a vykonávať pokročilú analýzu údajov pomocou sofistikovaných techník grafovej analýzy.

Víziou Graphlyticu je umožniť každému používateľovi vysokohodnotnú interaktivitu grafov s intuitívnym a ľahko použiteľným používateľským rozhraním. Hoci každá grafová analýza zdieľa rovnaké základné techniky vizualizácie grafov a manipulácie, každý scenár je iný a riešenie je často potrebné vylepšiť špecificky pre dáta a oblasť záujmu klienta. To je dôvod, prečo je našou stratégiou odhodlanie pomáhať každému klientovi s nezrejmými nuansami grafovej analýzy a automatizácie implementáciou vlastných doplnkov a widgetov, ktoré dokonale zapadajú do konkrétnych prípadov použitia.

Hlavným cieľom Graphlytic je pomôcť používateľom odhaliť skryté poznatky a vzorce v ich údajoch, ktoré tradičné metódy analýzy údajov nemusia odhaliť. Výsledkom je, že Graphlytic sa stáva čoraz populárnejším v rôznych odvetviach vrátane financií, zdravotníctva, telekomunikácií a marketingu, kde analýza sieťových údajov zohráva kľúčovú úlohu pri zisťovaní podozrivých aktivít, predpovedaní trendov a prijímaní informovaných rozhodnutí.

Grafické riešenia zvyčajne pozostávajú z niekoľkých integrovaných komponentov. Na obrázku nižšie je architektúra na vysokej úrovni možných integrovaných systémov. Viac informácií nájdete v dokumentácii v kapitole Integrácie.

Jednou z hlavných funkcií Graphlyticu je jeho široká škála možností dotazovania a filtrovania. Používatelia môžu využiť výkonné možnosti vyhľadávania na identifikáciu špecifických údajových bodov alebo uzlov v grafe, rýchlo skenovať anomálie alebo podozrivú aktivitu. Okrem toho existujú rôzne možnosti filtrovania a zoskupovania, ktoré používateľom umožňujú zamerať sa na konkrétnu podmnožinu údajov grafu. Používatelia môžu napríklad zoskupovať uzly do rôznych kategórií na základe vlastností, ako je vek, poloha alebo história nákupov.

Ďalšou dôležitou vlastnosťou Graphlyticu je jeho schopnosť vykonávať pokročilé techniky grafovej analýzy. Graphlytic podporuje celý rad analytických algoritmov vrátane analýzy centrality, detekcie komunity a analýzy cesty. Analýza centrality vypočítava dôležitosť uzlov v grafe a používa sa na identifikáciu uzlov, ktoré sú kľúčové pre sieť. Komunitná detekcia identifikuje skupiny uzlov, ktoré sú husto prepojené a považujú sa za podobné. Analýza cesty umožňuje používateľom

# Error! Reference source

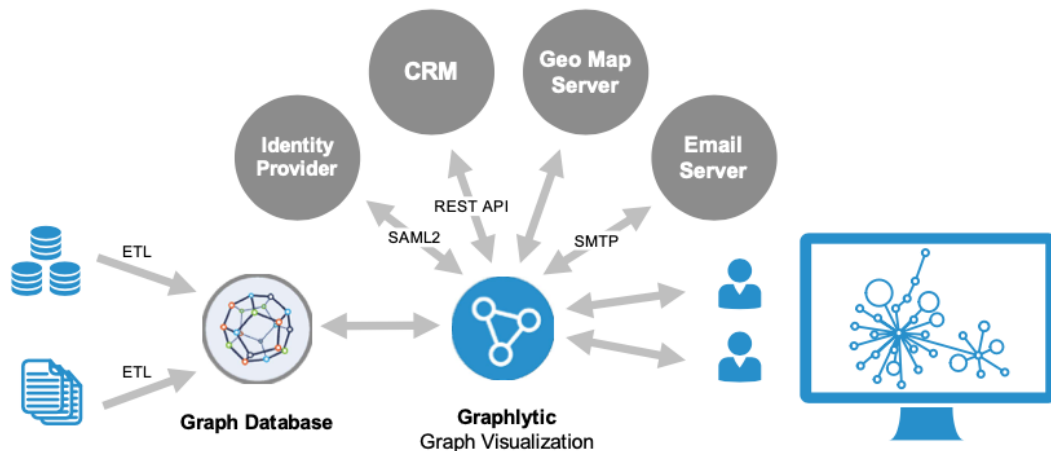
identifikovať najkratšiu cestu medzi dvoma bodmi v grafe a vyhodnotiť vlastnosti cesty.

Graphlytic tiež poskytuje množstvo prispôsobiteľných možností vizualizácie. Používatelia si môžu vybrať z rôznych rozložení a režimov zobrazenia, vrátane rôznych interaktívnych vizualizácií, tepelných máp, tabuliek a iných analytických grafov. Platforma ponúka užívateľsky prívetivú funkciu drag-and-drop, ktorá uľahčuje zmenu usporiadania uzlov a hrán. Používatelia môžu tiež prispôbiť farbu, tvar a veľkosť uzlov a hrán tak, aby odrážali charakteristiky ich údajov.

## 5.1.3 Základné komponenty Graphlyticu

1. **Príjem údajov:** Na import údajov do databázy grafov sa použijú existujúce možnosti ETL zákazníka. Modul ETL spoločnosti Graphlytic možno použiť na ďalšie úlohy, ako je vylepšenie modelu pomocou odvodených vzťahov, monitorovanie podozrivých vzorov a odosielanie upozornení.
2. **Diagram znalostí:** Databáza grafov (napr. Neo4j, Cosmos DB, Janus Graph, AWS Neptune, Apache TinkerPop atď.) sa používa na ukladanie a spracovanie údajov z grafu znalostí. Toto je primárny zdroj údajov pre analytickú vizualizáciu. Definované zdroje údajov je možné importovať do databázy grafov v preddefinovaných intervaloch (každú hodinu, deň alebo týždeň). Ad-hoc zdroje údajov sa importujú do grafu databázy grafov na požiadanie. Odvodené dáta a vlastnosti sú vypočítané pomocou algoritmov z teórie grafov.
3. **Vizualizácia:** Graphlytic poskytne pre aplikáciu možnosti vizualizácie grafov. Používatelia budú pomocou aplikácie vyhľadávať v databáze grafov vzory, vytvárať vizualizácie, a exportovať výstupy, ako sú množiny údajov a obrázky.
4. **Integrované systémy:** Grafycké riešenie je možné integrovať s niekoľkými externými systémami lepšie zapadať do existujúcej infraštruktúry. Zvyčajne ide o integráciu SMTP s existujúcim e-mailovým serverom integrácie SAML2 na poskytovanie používateľov a jednotné prihlásenie. Graphlytic prichádza s REST API, ktoré možno použiť na implementáciu akéhokoľvek vlastného integračného rozhrania.

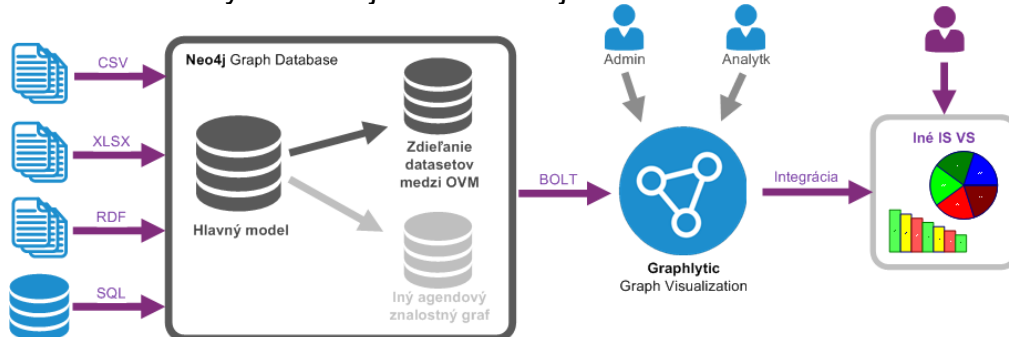
# Error! Reference source



Obrázok 1: Aplikačná architektúra Graphlyticu

## 5.2 Architektúra riešenia

Schéma architektúry riešenia je na nasledujúcom obrázku.



Obrázok 2: Schéma architektúry riešenia

Schéma znázorňuje postupnosť spracovania údajov (zľava doprava):

- 1. Import údajov zo zdrojových súborov a systémov do grafovej databázy.** Automatizácia importu údajov je umožnená modulom Jobs nástroja Graphlytic. Podporovaných je viacero formátov zdrojových dát, napríklad CSV, MS Excel, RDF, priame napojenie na SQL DB, príjem údajov z API systémov tretích strán. Vytvorené importné joby je možné plne automatizovať a spúšťať v pravidelných intervaloch alebo spustením na diaľku pomocou REST API rozhrania aplikácie Graphlytic. Pre zoznam konkrétnych dátových zdrojov v riešení pozri kapitolu 4
- 2. Zdroje údajov použité v grafových modeloch.**

# Error! Reference source

3. **Vytvorenie odvođených, tzv. “agendových” znalostných grafov** pre potreby špecifických analýzy alebo pohľadov na údaje.
4. **Sprístupnenie údajov vo forme analytickej vizualizácie v nástroji Graphlytic.** Aplikácia umožňuje:
  - a. analyzovať data pomocou user-friendly interaktívnej vizualizácie,
  - b. vytvárať skupiny používateľov s rôznym stupňom oprávnení a rôznou úrovňou prístupu ku dátam v znalostnom grafe,
  - c. manažovať importné joby
  - d. manažovať vytváranie agendových znalostných grafov
5. **Integrácia znalostných grafov na IS VS tretích strán** (iné prvky infraštruktúry CSRÚ, alebo iné systémy mimo CSRÚ).

## 5.2.1 Zdroje údajov použité v grafových modeloch

V aktuálnom riešení sú zapojené tieto dátové zdroje:

1. [Číselník Agiend VS](#) importovaný z Meta IS.
2. [Číselník Úsekov VS](#) importovaný z Meta IS.
3. [Číselník Životných situácií](#) importovaný z Meta IS.
4. **Integračný manuál CSRÚ** – export dát z CSRÚ databázy manuálne upravený do podoby Integračného manuálu. Obsahuje údaje o aktuálne poskytovaných datasetoch v platforme CSRÚ.
5. **Koncepčný model** - MS Excel súbory (~60 ks), ktoré sú výstupom analýzy na jednotlivých OVM a obsahujú informácie o OVM, prevádzkovaných IS VS, zdieľaných datasetoch a súvisiacich právnych predpisoch.
6. **Logický model** - MS Excel súbory s popisom štruktúry údajov jednotlivých Objektov Evidencie s rozkladom na dátové atribúty a ich mapovanie na CMU.

## 5.3 Grafové modely a ich použitie

V riešení je použité dvojvrstvé generovanie grafových modelov, pričom jeden grafový model je „hlavný“, ktorý obsahuje najpodrobnejšie dáta zo všetkých dátových zdrojov prepojených do jedného znalostného grafu. Z tohto hlavného modelu sú generované odvođené modely, ktoré obsahujú zjednodušený pohľad na konkrétnu agendu. Takéto odvođené modely sú vhodnejšie na vyhodnotenie konkrétnej agendy, pretože model špecificky pripravený na analytickú prácu v rámci danej problematiky.

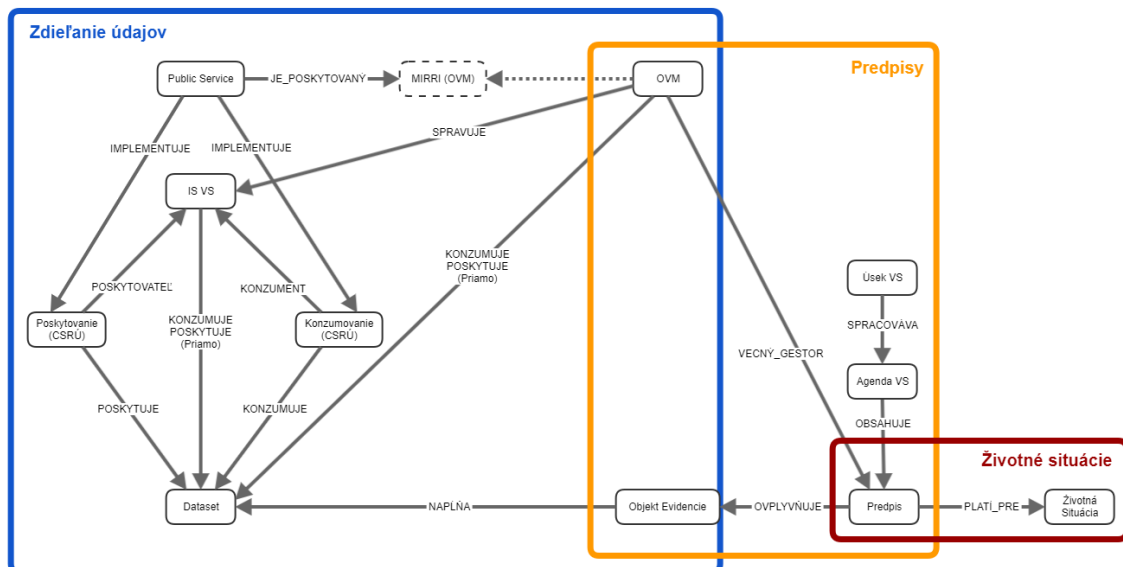
# Error! Reference source

## 5.3.1 Hlavný grafový model

Hlavný model obsahuje všetky údaje v najpodrobnejšej forme. Všetky odvodené (agendové) modely obsahujú podmnožinu údajov z tohto modelu.

Hlavný model obsahuje prienik viacerých agend:

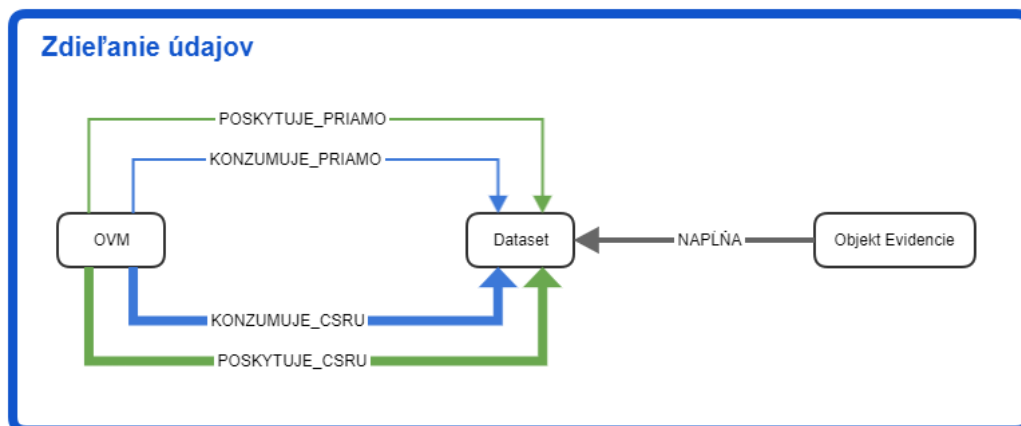
- **Zdieľanie údajov** – hlavná časť modelu, ktorá zachytáva podrobnosti o poskytovaní a konzumácii datasetov.
- **Predpisy** – časť modelu, ktorá mapuje právne predpisy v oblasti Objektov Evidencie a zodpovedností rôznych OVM v súvislosti s poskytovaním údajov.
- **Životné situácie** – časť modelu obohacujúca časť Predpisy o životné situácie.



Obrázok 3: Hlavný grafový model

# Error! Reference source

## 5.3.2 Agendový grafový model – Zdieľanie datasetov medzi OVM



Obrázok 4: Agendový grafový model

OVM

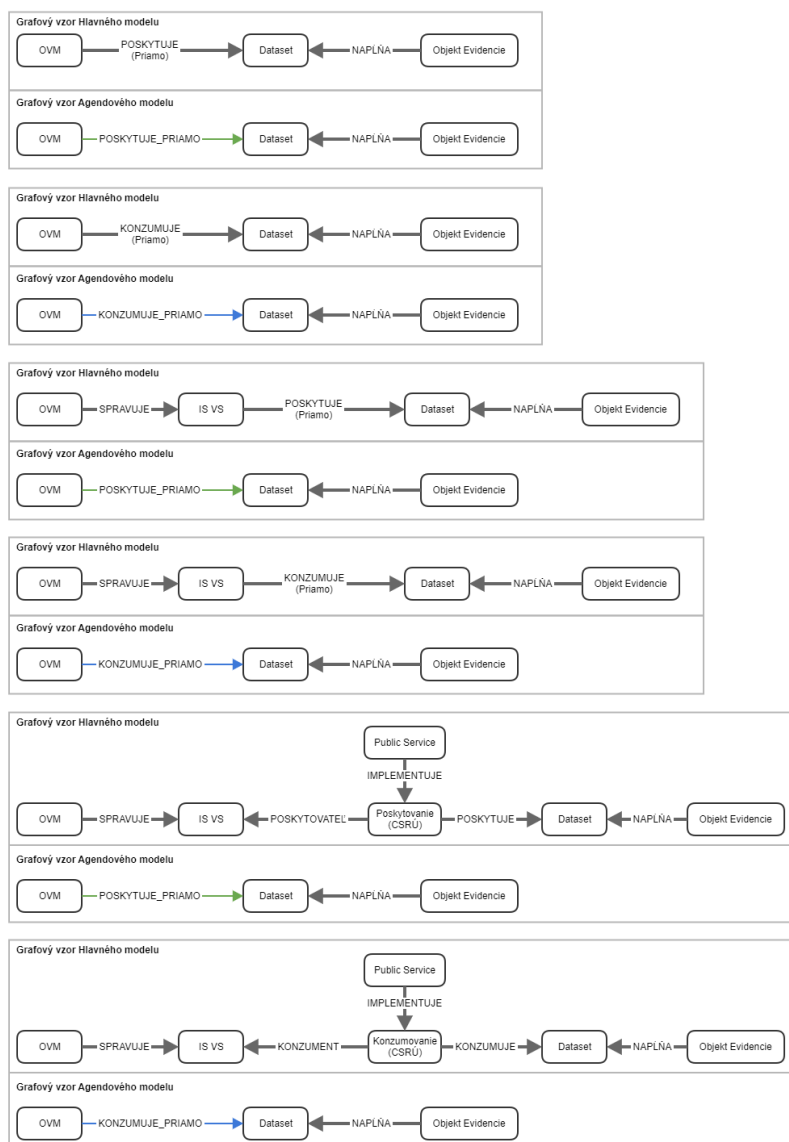
a Datasetom vyjadrujú existenciu integračnej väzby. Typy zachytených väzieb sú:

- POSKYTUJE\_PRIAMO
- KONZUMUJE\_PRIAMO
- POSKYTUJE\_CSRU
- KONZUMUJE\_CSRU

## 5.3.3 Mapovanie grafových vzorov

Nasledujúci obrázok zachytáva mapovanie grafových modelov, ktoré bolo uplatnené na dáta Hlavného modelu a ktoré vyprodukovalo dáta v agendovom modeli „Zdieľanie datasetov medzi OVM“.

# Error! Reference source



Obrázok 5: Mapovanie grafových vzorcov

## 5.4 Práca s aplikáciou Graphlytic

Aplikácia Graphlytic slúži na analýzu údajov jednotlivých znalostných grafov, riadenie prístupu ku údajom pre jednotlivé skupiny používateľov a tiež na manažment automatizácie importu údajov a generovania agendových grafov.

### 5.4.1 Prístup do aplikácie

Prístup do aplikácie je zabezpečený viacerými bezpečnostnými prvkami:

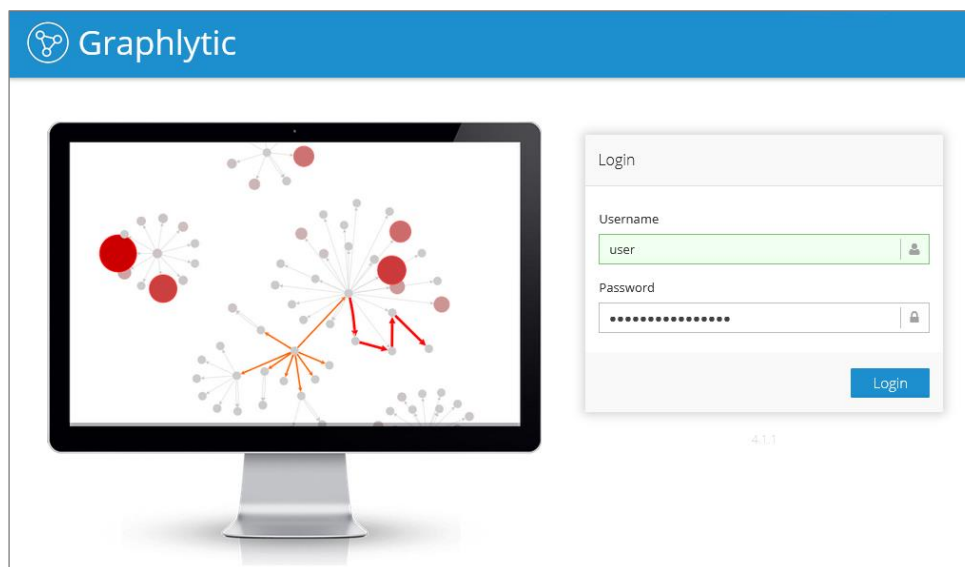
- VPN tunel vládneho cloudu
- ACMPP - server pre prístup do produkčnej infraštruktúry projektu CSRÚ



# Error! Reference source

- Proxy nastavenia prehliadača
- Login a heslo aplikácie Graphlytic

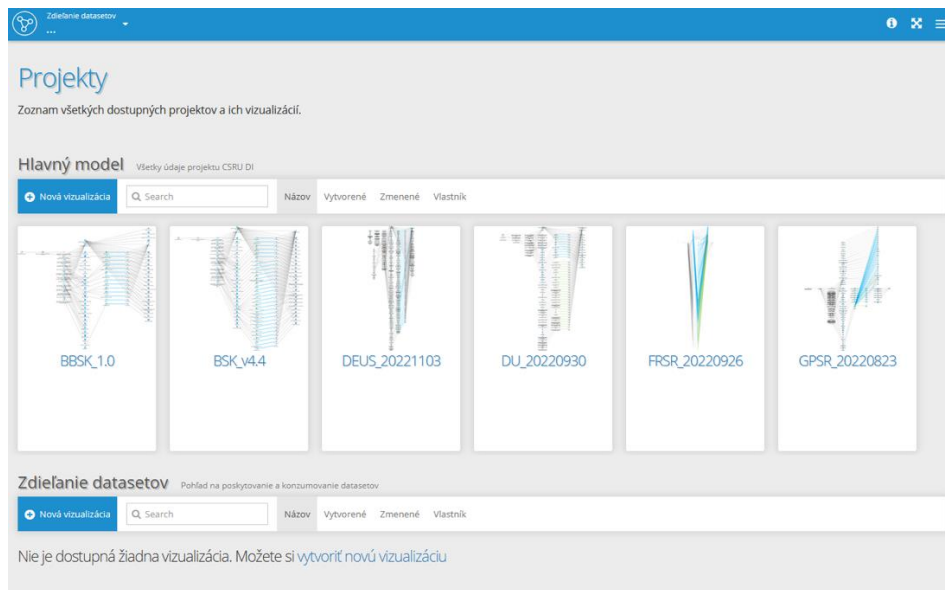
Pre získanie prístupu prosím kontaktujte svojho administrátora. Po úspešnom nastavení potrebných bezpečnostných prístupov je prihlasovacia obrazovka aplikácie Graphlytic dostupná na URL <https://app27a.csru.sk.cloud:6443/>



Obrázok 6: Prostredie pre prístup do Graphlyticu

Po úspešnom prihlásení sa zobrazí obrazovka s dostupnými projektami (znalostnými grafmi) a vizualizáciami:

# Error! Reference source



Obrázok 7: Náhľad na vizualizáciu možnosti rozdelenia projektov

## 5.4.2 Práca s hlavným modelom

Projekt „Hlavný model“ obsahuje všetky dáta importované v rámci riešenia. Používa sa najmä na komplexné hodnotenie integrity nimportovaných údajov, ich prepojení, a hľadanie nekonzistencií v rámci čistenia dát. Vid' kapitola 5.5.3.1 Hlavný grafový model pre viac informácií.

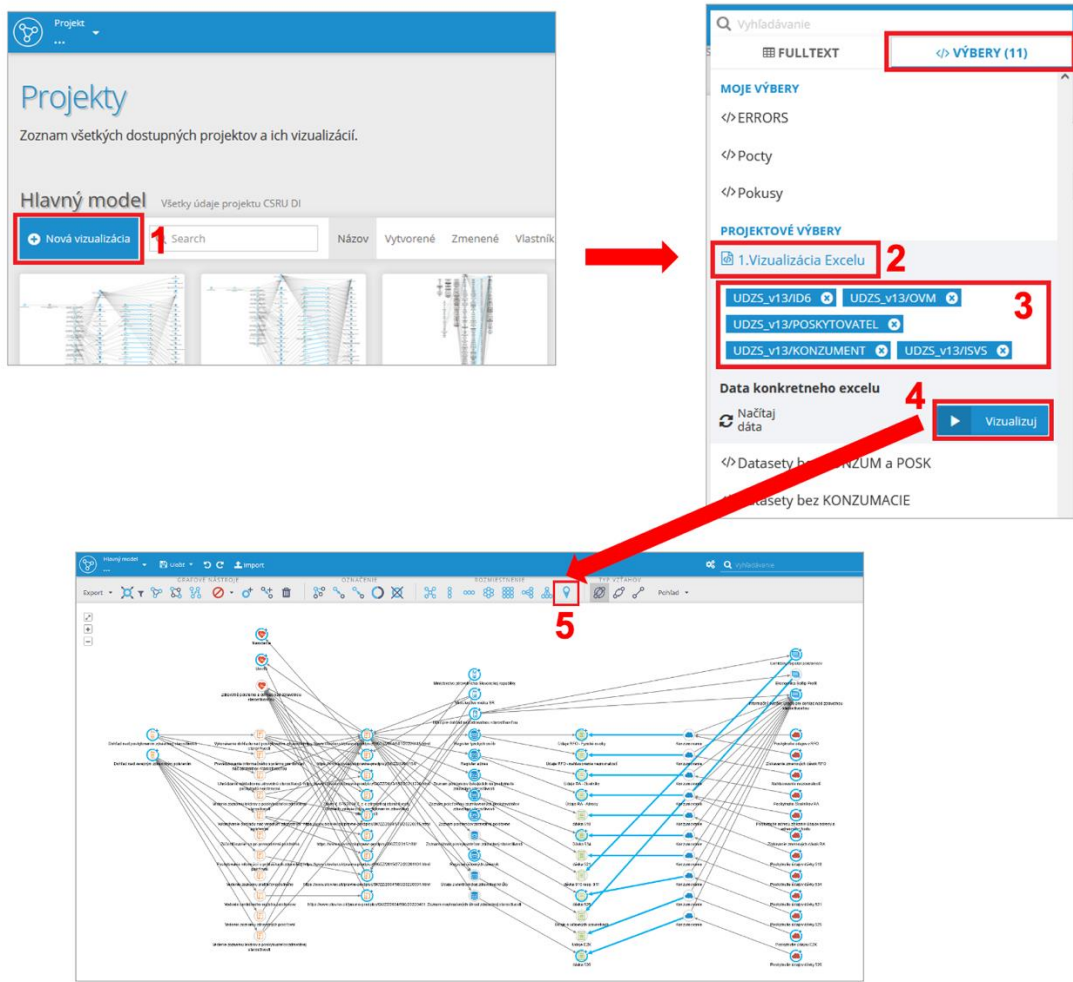
## 5.4.3 Vizualizácia dát konkrétneho OVM.

Typický príklad použitia je vizualizácia dát koncepčného modelu zozbieraných na konkrétnej OVM.

1. Otvoríme novú vizualizáciu tlačidlom “Nová vizualizácia” (vid' obr. nižšie).
2. Otvoríme vyhľadávanie v hlavičke aplikácie, zvolíme menu “Výbery” (Queries) a z projektových výberov zvolíme “1.Vizualizácia Excelu”.
3. Pridáme záložky excelov, ktoré chceme vizualizovať. Ak chceme vidieť všetky údaje danej OVM tak je potrebné pridať všetky záložky danej OVM. Napr. pre pridanie všetkých dát ÚDZS treba vybrať záložky zobrazené na obrázku nižšie (OVM, ISVS, ID6, POSKYTOVATEL, KONZUMENT).
4. Tlačidlom “Vizualizuj” dotiahneme údaje danej OVM zo znalostného grafu a zobrazíme ich vo vizualizácii.

# Error! Reference source

6. Použijeme tlačidlo “Usporiadaj podľa typu” na usporiadanie do stĺpcov podľa typu uzlu.



Obrázok 8: vizualizácia údajov konkrétneho OVM

## 5.5 Práca s agendovým modelom „Zdieľanie datasetov medzi OVM“

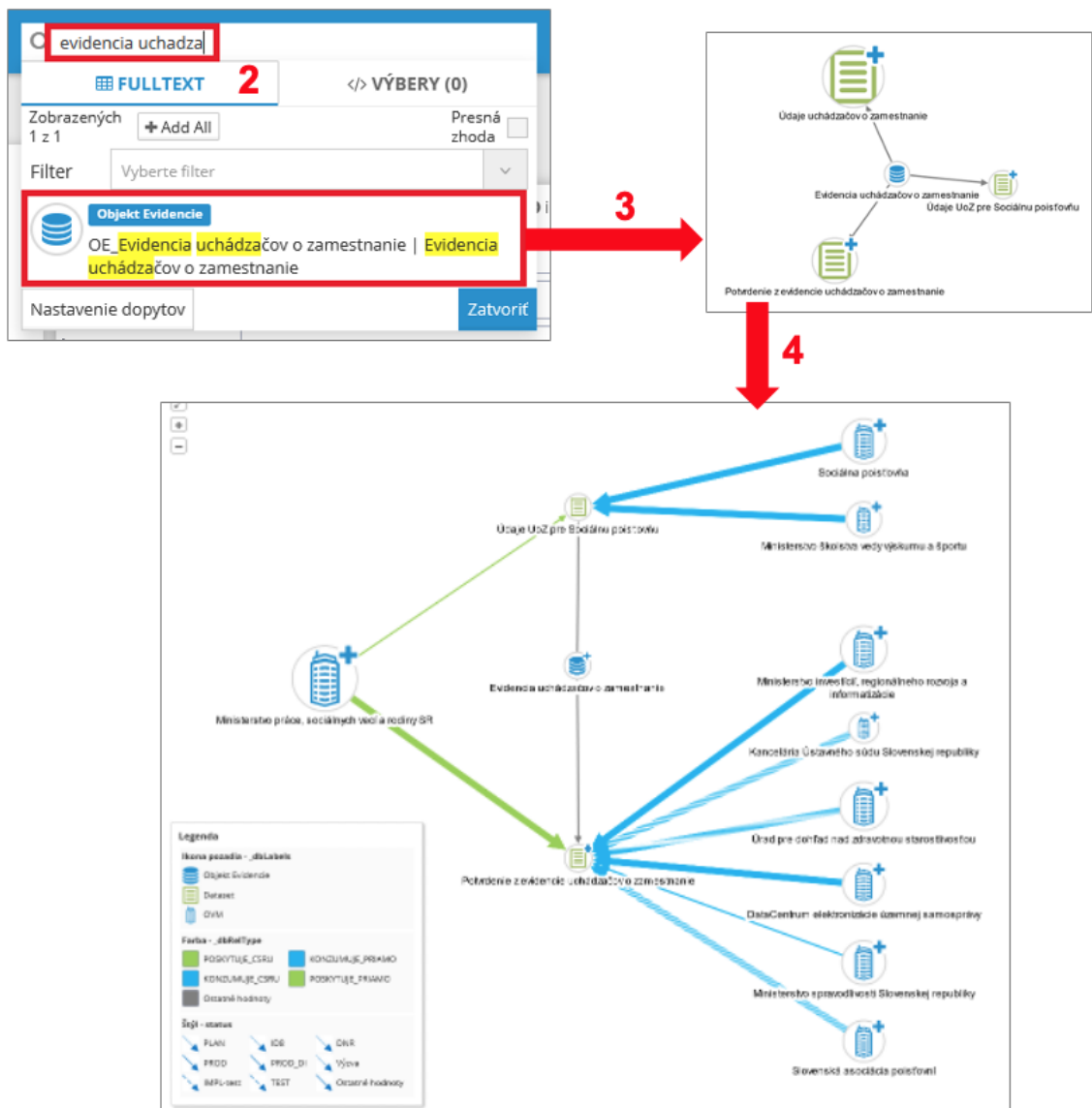
Projekt „Zdieľanie datasetov medzi OVM“ obsahuje podľa na agendu „Zdieľanie datasetov“, pričom sú zachytené len vzťahy (Objekt Evidencie) => (Dataset) <= (OVM). Vid' kapitola 5.5.3.2 Agendový grafový model – Zdieľanie datasetov medzi OVM pre viac informácií.

# Error! Reference source

## 5.5.1 Vizualizácia zdieľania konkrétneho Datasetu

Postup vizualizovania poskytovania a konzumácie konkrétneho Datasetu:

1. Otvoríme novú vizualizáciu tlačidlom “Nová vizualizácia” (vid’ obr nižšie).
2. Otvoríme vyhľadávanie v hlavičke aplikácie a napíšeme časť názvu hľadaného Datasetu alebo Objektu Evidencie.
3. Vyberieme jeden alebo viacero záznamov z výsledkov vyhľadávania – klikaním na záznamy sa pridajú uzly do vizualizácie.
4. Následne rozklikneme vzťahy uzlov po požadovanej hĺbky.



Obrázok 9: 5.5.1 Vizualizácia zdieľania konkrétneho Datasetu

# Error! Reference source

## 5.6 Synergia Graphlytic s Neo4J

Neo4j a Graphlytic sú silné nástroje pre spracovanie a analýzu grafových dát, ktoré majú široké uplatnenie v rôznych oblastiach, ako je podniková analýza, sociálne siete, biomedicína, doprava a mnoho ďalších. Ich kombinácia umožňuje efektívne pracovať s veľkými a komplexnými grafovými dátami a získať cenné informácie a vzorce zo štruktúry dát.

Majú úzku vzájomnú väzbu, pričom môžu byť použité spoločne na efektívne spracovanie, skladovanie, vizualizáciu a analýzu grafových dát.

Neo4j je grafová databáza, ktorá poskytuje silnú základňu pre ukladanie a spracovanie grafových dát. Je optimalizovaná pre rýchly prístup k dátam v grafových štruktúrach a ponúka bohatú sadu funkcií pre dotazovanie a analýzu grafových dát. S Neo4j je možné vytvárať a spravovať veľké grafové databázy a vykonávať pokročilé operácie nad nimi.

Graphlytic na druhej strane je nástroj zameraný na vizualizáciu a analýzu grafových dát. Je navrhnutý tak, aby poskytoval interaktívne a vizuálne atraktívne prostredie na preskúmanie grafových štruktúr a objavovanie vzorov a informácií v grafových dátach. Graphlytic umožňuje užívateľom importovať dáta zo zdrojov, ako je napríklad Neo4j, a následne ich vizualizovať, analyzovať a dotazovať sa na ne prostredníctvom svojho rozhrania.

V praxi je možné použiť Neo4j ako základňu pre ukladanie a spracovanie grafových dát a následne tieto dáta importovať do Graphlytic pre ich vizualizáciu a pokročilú analýzu. Graphlytic poskytuje pripojenie k Neo4j cez svoje rozhranie a umožňuje importovať dáta z Neo4j databázy, pričom si udržuje ich štruktúru a atribúty. Takýmto spôsobom môžu užívatelia využiť výhody Neo4j ako robustnej grafovej databázy a zároveň využívať silu a interaktívne funkcie Graphlytic pre analýzu a vizualizáciu grafových dát.

Neo4j a Graphlytic sa vzájomne dopĺňajú v niekoľkých aspektoch:

1. **Ukladanie a správa grafových dát:** Neo4j je navrhnutý na efektívne ukladanie a spracovanie veľkých grafových dát. Jeho štruktúrovaný prístup a dotazovací jazyk Cypher umožňujú robustné operácie s grafovými dátami. Graphlytic na druhej strane poskytuje prostredie na vizualizáciu a analýzu týchto dát. Importom grafových dát z Neo4j do Graphlytic je možné vytvoriť interaktívne vizualizácie a analytické modely.

# Error! Reference source

- 2. Vizualizácia a interakcia s grafovými dátami:** Graphlytic ponúka široké možnosti vizualizácie grafových dát. Umožňuje užívateľom prispôbiť vzhľad vrcholov, hrán a ich atribútov. Interaktívne funkcie, ako približovanie, presúvanie a filtrovanie, umožňujú preskúmanie a hlbšiu analýzu grafových dát. Pri práci s veľkými a komplexnými grafmi poskytuje Graphlytic prehľadné a súvislé zobrazenie štruktúry dát.
- 3. Pokročilá analýza grafových dát:** Neo4j disponuje bohatou sadou algoritmov pre analýzu grafových dát, vrátane centrality, detekcie komúní, vyhľadávania vzorov a predikcie. Tieto algoritmy môžu byť využité pri spracovaní dát v Neo4j. Graphlytic na druhej strane poskytuje prostredie pre použitie týchto algoritmov a vizualizáciu výsledkov. Užívatelia môžu využívať pokročilú analýzu grafových dát v kombinácii s vizualizačnými nástrojmi Graphlytic.
- 4. Integrácia s ďalšími nástrojmi a systémami:** Neo4j a Graphlytic sú flexibilné a umožňujú integráciu s ďalšími nástrojmi a systémami. Neo4j poskytuje rôzne možnosti exportu a importu dát, čo umožňuje ich prepojenie s inými databázovými systémami. Graphlytic podporuje import dát z rôznych zdrojov a môže byť integrovaný s rôznymi nástrojmi na spracovanie a analýzu dát.

V závere je spojenie Neo4j a Graphlytic silnou kombináciou pre prácu s grafovými dátami. Poskytujú užívateľom možnosti ukladania, spracovania, vizualizácie a analýzy grafových dát. Ich vzájomná integrácia umožňuje efektívne a intuitívne pracovať s grafovými dátami a získavať cenné informácie a vzorce zo štruktúry dát. S vývojom a rozšírením grafových dát v rôznych oblastiach bude spojenie Neo4j a Graphlytic pravdepodobne ďalej rásť a poskytovať nové možnosti pre objavovanie a analýzu informácií z grafových dát.

# Error! Reference source

## 6 Identifikácia priestoru na zlepšenie

### 6.1 Využitie znalostných grafov

V posledných desaťročiach sme boli svedkami exponenciálneho rastu množstva údajov, ktoré ľudstvo vytvára a spracováva. Zároveň sú údaje menej štruktúrované a viac prepojené. Alternatívy k tradičným relačným databázam sú v reakcii na tieto trendy čoraz populárnejšie. Databázy grafov a nástroje na spracovanie grafov prijímajú organizácie všetkých veľkostí na ukladanie, dopytovanie a analýzu vysoko prepojených údajov v širokej škále domén.

Interaktívne grafové vizualizácie zohrávajú hlavnú úlohu pri odhaľovaní nových vzorcov v údajoch. Naše mozgy sú skvelé na rozpoznávanie vzorov, ale údaje musia byť prezentované v stráviteľnej forme a vizualizácia musí ponúkať dostatok interaktivity na vykonávanie operácií na vyššej úrovni s grafom, ako je napríklad interaktívna rozširovanie údajov grafu, pridávanie údajov na základe podmienok, filtrovanie nepotrebných údajov, zvýrazňovanie pod vzorcov, simulácia správania na základe jedinečných znalostí analytika a mnoho ďalšieho. Analytici sú tiež schopní definovať nové hypotézy založené na údajoch a potom tieto hypotézy veľmi rýchlo vyhodnotiť na skutočných údajoch, ktoré môžu potenciálne identifikovať ďalšie prípady rovnakých scenárov.

Účelom znalostného grafu poskytnúť úplnejšie a presnejšie pochopenie sveta okolo nás. Využitím najnovších technológií a algoritmov AI pomáhajú znalostné grafy preklenúť priepasť medzi ľudskou a strojovou inteligenciou. Znalostné grafy zohrávajú v moderných aplikáciách AI čoraz dôležitejšiu úlohu, od poskytovania presnejších výsledkov vyhľadávania až po umožnenie prispôsobenejších odporúčaní.

Jadrom znalostného grafu je myšlienka trojíc – subjekt, predikát a objekt. Zjednodušene povedané triplet. Triplet je výrok, ktorý spája dve entity. Napríklad „Elon Musk je generálnym riaditeľom spoločnosti Tesla“. V tomto príklade je Elon Musk subjektom, „je generálnym riaditeľom“ je predikát a Tesla je objektom. Znalostný graf je vo svojej podstate súborom vzťahov medzi entitami. Každá entita je v grafe znázornená ako uzol, ktorý je s ostatnými uzlami spojený hranami reprezentujúcimi typ vzťahu medzi nimi.

Napríklad v prípade registra fyzických osôb, ak by sme chceli reprezentovať informácie o osobe, mohli by sme vytvoriť uzol reprezentujúci danú osobu a prepojiť ho s ďalšími uzlami reprezentujúcimi jej meno, dátum narodenia, povolanie a ďalšie relevantné informácie o nej.

# Error! Reference source

Usporiadáním všetkých informácií do trojíc poskytujú znalostné grafy flexibilný spôsob kódovania vedomostí. To znamená, že používatelia môžu jednoducho pridávať alebo odstraňovať prepojenia medzi entitami a aktualizovať informácie uložené v grafe.

Jednou z kľúčových výhod znalostných grafov je, že sa dajú použiť na reprezentáciu zložitých informácií spôsobom, ktorý možno ľahko vyhľadávať a analyzovať. Keďže údaje sú uložené v grafovej štruktúre, je ľahké sa v nich orientovať a extrahovať konkrétne informácie. Použitím techník spracovania prirodzeného jazyka – analýzy sentimentu a iných foriem – môžu znalostné grafy poskytnúť odpovede na zložité otázky, aj keď odpoveď nie je explicitne uvedená. Ak sa napríklad spýtame otázku ako „Ktorých je desať krajín s najvyšším HDP na obyvateľa?“, znalostný graf dokáže efektívne navigovať v grafe a poskytnúť odpoveď.

Ďalšou výhodou znalostných grafov je, že umožňujú algoritmom strojového učenia lepšie reprezentovať a pochopiť vzťahy medzi rôznymi entitami. Zhromažďovaním údajov z viacerých zdrojov a vytváraním komplexného, hierarchického pohľadu na doménu môžu znalostné grafy pomôcť modelom, v ktorom fungujú. To im zase umožňuje robiť presnejšie predpovede a odporúčania, či už ide o identifikáciu relevantného obsahu pre používateľov alebo predpovedanie správania zákazníkov.

## 6.2 Využitie znalostných grafov v kontexte E-Government-e

Využitie znalostných grafov v e-Governmente je oblasťou s veľkým potenciálom, ktorá môže výrazne zlepšiť procesy a služby verejnej správy. Znalostné grafy umožňujú modelovať a reprezentovať vzťahy medzi rôznymi entitami a atribútmi v dátach, čo je kritické pre efektívne spravovanie a využívanie informácií. Tu je prehľad o využití znalostných grafov v e-Governmente:

1. **Správa a integrácia dát:** Jedným z hlavných prínosov znalostných grafov v e-Governmente je schopnosť spravovať a integrovať rôznorodé dáta z viacerých zdrojov. Znalostné grafy umožňujú prepojiť a integrovať dáta z rôznych systémov a databáz do jedného centrálného grafu. To umožňuje lepšiu interoperabilitu a spojitosť medzi dátami, čo je kľúčové pre poskytovanie komplexných a personalizovaných služieb verejnej správy.
2. **Personalizácia služieb:** Znalostné grafy umožňujú efektívne zisťovať preferencie, potreby a vzťahy občanov a podnikateľov prostredníctvom analýzy ich interakcií s rôznymi službami a dátami. Týmto spôsobom je



# Error! Reference source

možné personalizovať poskytované služby a ponúkať občanom a podnikateľom relevantné informácie a možnosti. Napríklad na základe historických dát o interakciách a preferenciách občana je možné mu poskytnúť personalizované odporúčania a informácie.

3. **Optimalizácia procesov:** Znalostné grafy môžu byť využité na optimalizáciu a automatizáciu procesov v e-Governmente. Na základe znalostí o vzťahoch a závislostiach medzi entitami a atribútmi v grafe je možné identifikovať efektívne postupy a zlepšiť tok informácií a rozhodovanie v rôznych oblastiach verejnej správy. Napríklad, graf môže identifikovať a odhaliť slabé miesta v procesoch a navrhnúť lepšie alternatívy alebo automatické riešenia
4. **Zlepšenie interoperability a integrácie:** Znalostné grafy v E-Governmente prispievajú aj k zlepšeniu interoperability a integrácie medzi rôznymi systémami a dátovými zdrojmi. Vďaka grafom je možné modelovať vzťahy medzi rôznymi entitami a atribútmi, čo umožňuje vytvárať jednotný pohľad na dáta a služby z rôznych systémov. Týmto spôsobom je dosiahnutá lepšia komunikácia a zdieľanie informácií medzi rôznymi verejnými orgánmi. Grafy umožňujú prepojenie rôznych typov dát a informácií, čo vedie k lepšiemu zdieľaniu a využívaniu informácií medzi rôznymi e-Government aplikáciami a službami. Týmto spôsobom sa zlepšuje koordinácia a efektívnosť procesov v rámci verejnej správy.
5. **Analýza a predikcia:** Znalostné grafy poskytujú bohatý zdroj dát pre analytické účely v e-Governmente. Analytické metódy a algoritmy môžu byť aplikované na grafovú štruktúru, aby sa získali nové poznatky, identifikovali vzory a získali informácie pre rozhodovanie. Napríklad analýza grafov môže odhaliť vzťahy medzi entitami, identifikovať kritické body alebo predpovedať trendy a udalosti.
6. **Správa znalostí a legislatívy:** Znalostné grafy môžu slúžiť ako nástroj na správu a organizáciu znalostí a legislatívy v e-Governmente. Grafická reprezentácia vzťahov medzi rôznymi právnymi dokumentmi, legislatívnymi aktmi a vzťahmi k iným entitám umožňuje lepšiu orientáciu v zložitých právnych systémoch. To podporuje rýchlejšie vyhľadávanie a porozumenie legislatívnych predpisov, čo je dôležité pre kvalitné poskytovanie služieb a zabezpečenie súladu so zákonnými požiadavkami.
7. **Transparentnosť a verejná kontrola:** Znalostné grafy môžu prispieť k väčšej transparentnosti a verejnej kontrole v e-Governmente. Grafická reprezentácia vzťahov medzi entitami a atribútmi môže pomôcť občanom a podnikateľom lepšie porozumieť, ako sa rozhodnutia a procesy v verejnej správe formujú. Týmto spôsobom sa zvyšuje dôvera verejnosti a umožňuje sa aktívnejšia účasť a spolupráca.

# Error! Reference source

Znalostné grafy poskytujú efektívny a flexibilný spôsob organizácie a správy dát v e-Governmente. Ich využitie prináša výhody v oblasti správy a integrácie dát, personalizácie služieb, optimalizácie procesov, analýzy a predikcie, správy znalostí a legislatívy, ako aj v oblasti transparentnosti a verejnej kontroly. Implementácia znalostných grafov v e-Governmente môže zlepšiť kvalitu a dostupnosť verejných služieb, zvýšiť efektívnosť rozhodovacích procesov a posilniť vzťah medzi verejnou správou a občanmi/podnikateľmi.

## 6.3 Využitie znalostných grafov v kontexte Dátovej kancelárie

Znalostné grafy majú v kontexte Dátovej kancelárie mnoho prínosov a výhod, ktoré prispievajú k lepšej organizácii, analýze a využitiu dát. Využitie znalostných grafov v kontexte Dátovej kancelárie prináša významné výhody v organizácii, správe a analýze dát. Pomáha zlepšovať interoperabilitu, efektívne vyhľadávať a dotazovať dáta, integrovať rôzne zdroje, analyzovať vzťahy a štruktúry dát, podporovať kolaboráciu a zdieľanie znalostí a vytvárať analytické modely. S týmito prínosmi môže Dátová kancelária dosiahnuť lepšie výsledky, efektívnejšiu prácu a rozhodovanie na základe výrazne zlepšenej kvality a spracovania dát. Podrobnejší opis hlavných prínosov znalostných grafov pre Dátovú kanceláriu môžeme zhrnúť do nasledovných bodov:

1. **Flexibilná reprezentácia dát:** Znalostné grafy poskytujú flexibilnú a výstižnú reprezentáciu dát. Dáta sú modelované ako uzly (entity) a hrany (vzťahy) medzi nimi. Táto grafická reprezentácia umožňuje jasné zobrazenie vzťahov medzi dátami a ich hierarchické usporiadanie. Grafy tiež umožňujú pripojenie ďalších atribútov a metadát k jednotlivým uzlom a hranám, čo pridáva ďalšiu hĺbku a komplexnosť do reprezentácie dát.
2. **Organizácia a správa dát:** Znalostné grafy poskytujú efektívny spôsob organizácie a správy dát v Dátovej kancelárii. Uzly grafu môžu reprezentovať rôzne typy dát, ako sú dokumenty, metadáta, zamestnanci, oddelenia a ďalšie relevantné informácie. Hrany medzi uzlami modelujú vzťahy a spojenia medzi entitami. Týmto spôsobom je možné jednoducho vyhľadávať, pristupovať a aktualizovať dáta v rámci Dátovej kancelárie.
3. **Vyhľadávanie a dotazovanie dát:** Znalostné grafy umožňujú vysoko efektívne vyhľadávanie a dotazovanie dát v Dátovej kancelárii. S využitím jazyka dotazovania, ako je napríklad SPARQL, je možné vykonávať zložité dotazy, ktoré zahŕňajú vzťahy a atribúty v grafoch. Napríklad je možné vyhľadávať dokumenty podľa ich obsahu, vzťahu k zamestnancom alebo príslušnosti k určitému oddeleniu. Taktiež je možné vykonávať analýzy na základe vzťahov a atribútov v grafoch.

# Error! Reference source

4. **Integrácia dát:** Znalostné grafy ponúkajú vynikajúci nástroj pre integráciu dát z rôznych zdrojov a systémov v rámci Dátovej kancelárie. Dáta z rôznych zdrojov môžu byť transformované na grafovú reprezentáciu a následne integrované do jedného centrálného grafu. Týmto spôsobom je možné prekonať problémy spojené s heterogénnymi dátovými formátmi a štruktúrami. Integrácia dát v rámci grafov umožňuje lepšiu interoperabilitu a jednoduchšie spravovanie dát v Dátovej kancelárii.
5. **Analýza vzťahov a štruktúr:** Znalostné grafy umožňujú analýzu vzťahov a štruktúr medzi dátami v Dátovej kancelárii. Algoritmy prehľadávania a analýzy grafov umožňujú odhaľovať skryté vzorce, identifikovať závislosti a získať nové poznatky zo štruktúrovaných aj neštruktúrovaných dát. Napríklad pomocou algoritmov prehľadávania grafu je možné identifikovať najdôležitejšie dokumenty v rámci určitého kontextu, objaviť skryté vzťahy medzi zamestnancami alebo identifikovať kľúčové termíny v dokumentoch.
6. **Kolaborácia a zdieľanie znalostí:** Znalostné grafy podporujú kolaboratívne prostredie, kde viacerí zamestnanci môžu prispievať a upravovať dáta v grafovej databáze. Týmto spôsobom je možné zdieľať a využívať znalosti v rámci celej organizácie, čo zlepšuje spoluprácu a efektivitu práce. Umožňuje to vytváranie a spravovanie centrálného zdroja informácií, ktorý je prístupný pre všetkých zainteresovaných.
7. **Zlepšenie interoperability a integrácie:** Znalostné grafy prispievajú k zlepšeniu interoperability a integrácie dát v Dátovej kancelárii. Grafy umožňujú reprezentovať a spracovávať viacero typov vzťahov medzi entitami, ako sú hierarchické vzťahy, asociačné vzťahy, časové vzťahy a podobne. Týmto spôsobom je možné vytvárať komplexnejšie a bohatšie dátové modely, ktoré lepšie zodpovedajú reálnemu svetu a umožňujú rôznym systémom a aplikáciám vzájomne komunikovať a využívať dáta.
8. **Vytváranie analytických modelov:** Znalostné grafy poskytujú vhodné prostredie pre vytváranie analytických modelov a predikčných analýz v rámci Dátovej kancelárie. Analytické nástroje môžu využívať vzťahy a atribúty v grafoch na identifikáciu vzorcov, predikciu správania alebo odhaľovanie anomálií. Napríklad je možné vytvoriť modely na odporúčanie optimálneho spájania dát, identifikáciu vzťahov medzi zákazníkmi alebo predikciu vývoja a trendov na základe existujúcich dát v grafoch.

## 6.4 Využitie znalostných grafov v kontexte dátovej integrácie

V kontexte dátovej kancelárie môžu znalostné grafy poskytnúť množstvo výhod, od zlepšenej správy dát až po efektívnejšiu analýzu dát. V tejto kapitole sa budeme zaoberať možnými spôsobmi využitia, akými môže dátová kancelária využívať znalostné grafy.

# Error! Reference source

1. **Integrácia viacerých zdrojov dát:** Znalostné grafy umožňujú efektívne integrovať dáta z rôznych zdrojov. Grafová štruktúra umožňuje modelovať vzťahy medzi jednotlivými entitami a ich atribútmi. Týmto spôsobom je možné jednoduchšie identifikovať a mapovať vzťahy medzi rôznymi dátovými zdrojmi, čo vedie k lepšej integrácii a spojeniu informácií.

Možno ich použiť na katalogizáciu údajov a riadenie. Znalostný graf môže poskytnúť komplexný, vzájomne prepojený pohľad na všetky dátové aktíva v rámci organizácie a z pohľadu dátovej kancelárie MIRRI prehľad na integrované objekty evidencie na IS CSRÚ. Reprezentáciou zdrojov údajov, entít a vzťahov v rámci jedného grafu môže dátová kancelária ľahšie pochopiť štruktúru a relevantnosť svojich údajov. Okrem toho môžu znalostné grafy pomôcť pri dočisťovaní prípadných konzistentnosti v názvosloví údajov, čím pomáhajú správcovi údajov udržiavať kvalitu údajov a zabezpečiť súlad s predpismi.

Určite netreba zabudnúť ani aj sémantickú integráciu dát, čo znamená, že sa zohľadňuje aj význam a vzťahy medzi dátami. Ako už bolo viackrát spomenuté grafová štruktúra umožňuje modelovať a reprezentovať sémantické vzťahy medzi entitami a atribútmi. Týmto spôsobom je možné identifikovať a využívať sémantické vzťahy medzi dátami, čo zlepšuje presnosť a relevantnosť integrovaných informácií.

2. **Zjednotenie dátových modelov:** Znalostné grafy poskytujú spoločný a jednotný dátový model, ktorý je nezávislý od konkrétnej implementácie zdrojov dát. To umožňuje zjednotenie dátových modelov a ich následnú integráciu prostredníctvom grafu. Grafová štruktúra uľahčuje mapovanie a prepojenie rôznych dátových modelov, čím sa dosiahne jednotný pohľad na dáta a zlepší sa konzistencia a spoľahlivosť informácií. Skúmaním vzťahov medzi rôznymi zdrojmi údajov môže dátová kancelária získať prehľad o tom, ako sú rôzne množiny údajov prepojené, čo môže potenciálne odhaliť nové príležitosti na analýzu k nevyriešeným problémom s kvalitou údajov. Znalostné grafy možno použiť aj na podporu sémantického vyhľadávania, čo umožňuje intuitívnejšie skúmanie veľkých súborov údajov.

Reprezentáciou údajov v grafovej štruktúre môže dátová kancelária do budúcnosti použiť algoritmy založené na grafoch na analýzu údajov, ako sú miery centrality, zoskupovanie alebo predikcia prepojenia. Tieto algoritmy môžu byť užitočné najmä pri analýze komplexných súborov údajov s

# Error! Reference source

nelineárnymi vzťahmi. Okrem toho môžu znalostné grafy poskytnúť základ pre vytváranie modelov strojového učenia tým, že poskytujú štruktúrovaný vstupný formát pre školiace sady a umožňujú začlenenie externých zdrojov vedomostí.

Napokon, znalostné grafy možno použiť na podporu vizualizácie údajov a vykazovania. Reprezentáciou údajov v grafovej štruktúre môže dátová kancelária vytvárať intuitívnejšie a vizuálne prítazlivejšie dashboardy alebo zostavy. Okrem toho môže použitie znalostného grafu umožniť dynamickejšie vizualizácie, ako sú interaktívne grafy alebo sieťové diagramy, ktoré môžu poskytnúť pútavejší a informatívnejší spôsob prezentácie údajov.

Na záver, použitie znalostných grafov môže poskytnúť množstvo výhod pre dátovú kanceláriu, vrátane zlepšeného riadenia údajov, lepšieho zisťovania a analýzy údajov, medzifunkčnej spolupráce a zdieľania znalostí a vylepšenej vizualizácie údajov a vykazovania. Keďže stále viac organizácií sa snaží lepšie využívať svoje údaje, znalostné grafy sa stávajú čoraz dôležitejším nástrojom na správu a analýzu rozsiahlych súborov údajov.

## 6.5 Integrácia Neo4j s nadstavbou Graphlytic na IS CSRÚ

V súčasnosti neexistuje prepojenie medzi Neo4j v podobe priamej integrácii na IS CSRÚ. Dáta sú importované z excelov alebo CSV súborov, ktoré sú exportované z IS CSRÚ, ale priame napojenie na niektorú databázu v IS CSRÚ zatiaľ nie je realizovaná.

V kapitole 5 sme si detailne popísali grafovú databázu Neo4j, Graphlytic ako aj výhody

Graphlytic je nástroj, ktorý môže byť vhodnou nadstavbou pre grafovú databázu Neo4j v kontexte IS CSRÚ a e-governmentu. Graphlytic poskytuje robustné funkcie pre vizualizáciu a analýzu grafových dát, čo umožňuje lepšiu interpretáciu a porozumenie vzťahov medzi entitami.

Integrácia Graphlytic s Neo4j na IS CSRÚ prináša niekoľko výhod:

1. **Vizuálna analytika:** Graphlytic ponúka interaktívne grafické rozhranie, ktoré umožňuje vizualizáciu grafových dát z Neo4j. Týmto spôsobom je možné lepšie porozumieť komplexným vzťahom a štruktúram dát v IS CSRÚ. Grafické zobrazenie umožňuje odhaľovať vzory, identifikovať závislosti a objaviť nové poznatky z dát.

# Error! Reference source

2. **Prispôsobiteľnosť a konfigurovateľnosť:** Graphlytic umožňuje prispôsobenie a konfiguráciu vizualizácií podľa konkrétnych potrieb IS CSRÚ. Môžete nastaviť rôzne typy uzlov, hrán, farieb, štýlov a ďalších vlastností grafu, aby ste vizuálne reprezentovali dáta v súlade s kontextom a požiadavkami.
3. **Vyhľadávanie a dotazovanie:** Graphlytic poskytuje nástroje na vyhľadávanie a dotazovanie dát v grafovom kontexte. Môžete vytvárať komplexné dotazy na grafovú databázu Neo4j a zobrazovať relevantné výsledky vo vizualizácii Graphlytic. Tento proces umožňuje rýchle a efektívne vyhľadávanie informácií v IS CSRÚ.
4. **Kollaborácia a zdieľanie:** Graphlytic podporuje kollaboratívne prostredie, v ktorom viacerí používatelia môžu spolupracovať a zdieľať informácie založené na grafových dátach. Týmto spôsobom sa zlepšuje komunikácia a spolupráca medzi rôznymi oddeleniami alebo subjektmi v rámci IS CSRÚ.

Integrácia Graphlytic s Neo4j a IS CSRÚ prináša prínosy v oblasti vizualizácie, analýzy, vyhľadávania a kollaborácie. Táto kombinácia nástrojov poskytuje robustné a flexibilné riešenie pre správu a využitie grafových dát v rámci Dátovej kancelárie a e-governmentu.

Spojenie Neo4j a Graphlytic predstavuje silný tandem pre spracovanie, správu, vizualizáciu a analýzu grafových dát. Neo4j poskytuje výkonný databázový systém založený na grafoch, ktorý umožňuje efektívne ukladanie, spracovanie a dotazovanie veľkých grafových dát. Graphlytic je nástroj zameraný na vizualizáciu a analýzu grafových dát, ktorý ponúka interaktívne prostredie pre objavovanie vzorov, analýzu vzťahov a vizuálne prezentácie.

Ich integrácia umožňuje užívateľom preskúmať, analyzovať a porozumieť komplexným grafovým štruktúram. Import grafových dát z Neo4j do Graphlytic poskytuje možnosť vytvárania prehľadných vizualizácií, prispôsobovania vzhľadu grafu a interaktívneho skúmania dát. Využitie dotazovacieho jazyka Cypher v Neo4j umožňuje vykonávať pokročilé dotazy a analýzy nad grafovými dátami, ktoré následne môžu byť vizualizované a ďalej analyzované v Graphlytic.

V budúcnosti je možné očakávať ďalší rozvoj a inovácie v oblasti grafových databázových systémov a nástrojov na vizualizáciu a analýzu grafových dát. Tieto technológie budú pravdepodobne pokračovať vo svojom raste a poskytovať ešte výkonnejšie a sofistikovanejšie nástroje na spracovanie a interpretáciu grafových dát.

konečnom dôsledku je spojenie Neo4j a Graphlytic vynikajúcou voľbou pre tých, ktorí sa zaujímajú o grafové dáta a chcú získať cenné informácie z týchto štruktúr. Ich kombinácia poskytuje užívateľom nástroje potrebné na skúmanie, analýzu a vizualizáciu grafových dát, čím podporuje lepšie rozhodovanie a objavovanie nových znalostí.

## 6.6 Odporúčane odborné role pre prácu so znalostnými grafmi

Práca so znalostnými grafmi vyžaduje určité odborné znalosti a schopnosti. Nasledujú niektoré odporúčané odborné role pre prácu so znalostnými grafmi:

1. **Dátový analytik:** Dátový analytik je odborník v analýze a interpretácii dát. Má schopnosť spracovávať veľké množstvá dát a identifikovať vzory, trendy a závislosti. V kontexte znalostných grafov môže dátový analytik využívať grafové algoritmy a techniky na analýzu vzťahov medzi entitami a atribútmi.
2. **Dátový architekt:** Dátový architekt je zodpovedný za návrh a správu dátových modelov a infraštruktúry. V prípade znalostných grafov je dátový architekt zodpovedný za vytvorenie efektívneho grafového modelu, ktorý zodpovedá potrebám dátovej kancelárie. Má schopnosť navrhovať a implementovať grafové databázy a zabezpečiť ich optimálne fungovanie.
3. **Expert na ontológie:** Expert na vytváranie ontológií, čo sú formálne reprezentácie znalostí a vzťahov medzi entitami. Expert na ontológie môže byť zapojený do tvorby a rozšírenia ontológií pre znalostné grafy v dátovej kancelárii. Jeho úlohou je definovať vzťahy, atribúty a zmysel pre entitami v grafe a zabezpečiť, aby bola zachovaná kvalita a konzistencia dát.
4. **Grafové databázový inžinier / Vývojár grafových databáz:** Vývojár grafových databáz je odborník v implementácii a vývoji grafových databáz. Má skúsenosti s konkrétnymi technológiami a nástrojmi pre grafové databázy a je schopný navrhovať a implementovať efektívne grafové databázové riešenia pre dátovú kanceláriu.
5. **Biznis analytik:** Biznis analytik je odborník na pochopenie potrieb a cieľov organizácie a prevod týchto požiadaviek na analytické projekty. V kontexte znalostných grafov môže biznis analytik identifikovať prípady použitia, definovať dotazy a analýzy a interpretovať výsledky, aby podporil rozhodovacie procesy v dátovej kancelárii.
6. **Informačný architekt:** Informačný architekt je zodpovedný za návrh informačnej štruktúry a organizácie v organizácii. V prípade znalostných grafov môže informačný architekt navrhovať a implementovať štruktúru

# Error! Reference source

grafových dát, definovať pravidlá a štandardy pre zachovanie integrovaných a konzistentných dát v dátovej kancelárii.

Tieto odborné role vnímame ako podstatné v kapacitnom zastúpení centrálnej kancelárie na MIRRI.

Pre rezortné dátové kancelárie prípadne menšie inštitúcie sa môžu role prekrývať alebo kombinovať v závislosti od veľkosti a štruktúry organizácie. Je dôležité mať tím odborníkov so znalosťami v týchto oblastiach, aby bolo možné úspešne implementovať a spravovať znalostné grafy v kontexte dátovej kancelárie.

## 6.7 **Publikácia výsledkov meraní pre oblasti využitia grafových nástrojov vo Verejnej správe**

Využitie grafových nástrojov vo verejnej správe prináša množstvo výhod a prínosov. Tieto nástroje umožňujú efektívnu vizualizáciu, analýzu a manipuláciu s grafovými dátami, čo je zvlášť užitočné v prípade komplexných a vzájomne prepojených informačných štruktúr. V tomto odbornom texte sa zameriame na významné oblasti využitia grafových nástrojov vo verejnej správe.

1. **Analýza sociálnej siete:** Verejná správa je zložitým systémom vzťahov a interakcií medzi rôznymi osobami, organizáciami a skupinami. Grafové nástroje sa využívajú na analýzu sociálnych sietí, čo umožňuje identifikovať vplyvné osobnosti, skupiny s podobnými záujmami a vzťahy medzi nimi. Tieto informácie môžu pomôcť pri plánovaní politík, rozhodovaní o rozpočtoch a posilňovaní vzájomnej spolupráce.
2. **Správa znalostí a informácií:** Vo verejnej správe je dôležité efektívne spravovať a organizovať veľké množstvo znalostí a informácií. Grafové nástroje umožňujú vytvárať prepojenia medzi rôznymi typmi dát, ako sú zákony, politiky, dokumenty a ďalšie relevantné informácie. Týmto spôsobom sa zlepšuje schopnosť identifikovať a pochopiť vzťahy a závislosti medzi rôznymi aspektmi verejnej správy.
3. **Riadenie a monitorovanie projektov:** Verejná správa často zahŕňa riadenie a monitorovanie projektov. Grafové nástroje poskytujú vizuálny prehľad o vzťahoch medzi rôznymi projektovými aktivitami, úlohami, zainteresovanými stranami a zdrojmi. Tieto nástroje umožňujú sledovať pokrok projektov, identifikovať riziká a problémy a zlepšovať efektívnosť a plánovanie projektov vo verejnej správe.



# Error! Reference source

4. **Identifikácia korupcie a nezrovnalostí:** Verejná správa musí čeliť problému korupcie a nezrovnalostí. Grafové nástroje môžu byť použité na identifikáciu potenciálnych korupčných vzorov a nezrovnalostí. Analyzujú vzťahy medzi osobami, organizáciami a transakciami, čo umožňuje odhaliť nezrovnalosti a identifikovať možné korupčné praktiky. Tieto nástroje prispievajú k transparentnosti a integrite vo verejnej správe.
5. **Sledovanie a riadenie verejnej dopravy:** Grafové nástroje majú tiež využitie pri sledovaní a riadení verejnej dopravy v mestských oblastiach. Pomocou grafov je možné modelovať a vizualizovať siete dopravných trás, zastávok, spojov a cestujúcich. Tieto nástroje umožňujú analýzu dopravných tokov, identifikáciu prekážok a optimalizáciu dopravných systémov vo verejnej správe.
6. **Analýza demografických dát a územného plánovania:** Grafové nástroje môžu byť využité na analýzu demografických dát a územného plánovania v rámci verejnej správy. Pomocou týchto nástrojov je možné vizualizovať a analyzovať geografické vzťahy medzi obyvateľstvom, infraštruktúrou, miestnymi službami a ďalšími faktormi. Tieto informácie umožňujú lepšie plánovanie rozvoja miest a regiónov, a to vrátane infraštruktúry, správy územného plánu a poskytovania verejných služieb.
7. **Analýza bezpečnosti a spravodlivosti:** Grafové nástroje majú potenciál byť využité pri analýze bezpečnostných a spravodlivostných aspektov v rámci verejnej správy. Tieto nástroje môžu vizualizovať a analyzovať vzťahy medzi rôznymi kriminálnymi činnosťami, zločineckými sieťami, právnymi subjektmi a ďalšími faktormi. Týmto spôsobom je možné identifikovať vzorce a trendy v kriminalite, identifikovať rizikové oblasti a prispieť k lepšiemu plánovaniu a implementácii bezpečnostných opatrení.
8. **Správa a analýza dát v zdravotníctve:** Grafové nástroje môžu byť využité aj v oblasti správy a analýzy dát v zdravotníctve. Pomocou týchto nástrojov je možné vizualizovať a analyzovať vzťahy medzi pacientmi, zdravotníckymi záznamami, liečebnými postupmi a ďalšími údajmi. To umožňuje identifikovať vzorce chorôb, monitorovať šírenie epidémií, sledovať účinnosť liečby a zlepšovať zdravotnícke služby a rozhodovanie vo verejnej správe.
9. **Manažment informácií a znalostí:** Grafové nástroje majú významné využitie aj pri manažmente informácií a znalostí vo verejnej správe. Tieto nástroje

# Error! Reference source

umožňujú vytváranie grafových štruktúr, ktoré prepájajú rôzne informácie, dokumenty, znalosti a zdroje v rámci organizácie. Týmto spôsobom je možné v kontexte verejnej správy.

# Error! Reference source

## Contact us

**Rudolf Sedmina**

partner

[rsedmina@kpmg.sk](mailto:rsedmina@kpmg.sk)

Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates or related entities.

[www.kpmg.com](http://www.kpmg.com)

© 2023 Copyright owned by one or more of the KPMG International entities. KPMG International entities provide no services to clients. All rights reserved.

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavour to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organization.