

UNIVERZITET U BEOGRADU
FAKULTET ORGANIZACIONIH NAUKA

Srdan N. Lalić

**UNAPREĐENJE SISTEMA POSLOVNE
INTELIGENCIJE PROCESOM GRUPNOG
ODLUČIVANJA**

doktorska disertacija

Beograd, 2016

UNIVERZITET U BEOGRADU
FAKULTET ORGANIZACIONIH NAUKA

Srdan N. Lalić

**UNAPREĐENJE SISTEMA POSLOVNE
INTELIGENCIJE PROCESOM GRUPNOG
ODLUČIVANJA**

doktorska disertacija

Beograd, 2016

UNIVERSITY OF BELGRADE
FACULTY OF ORGANIZATIONAL SCIENCES

Srdan N. Lalić

**IMPROVING SYSTEM OF BUSINESS
INTELLIGENCE THROUGH GROUP
DECISION MAKING PROCESS**

PhD thesis

Belgrade, 2016

MENTOR:

**Prof. dr Milija Suknović, redovni profesor
Fakultet organizacionih nauka, Beograd**

ČLANOVI KOMISIJE:

**dr Aleksandar Marković, redovni profesor
Fakultet organizacionih nauka, Beograd**

dr Dragan Radojević, naučni savetnik

DATUM ODBRANE:

UNAPREĐENJE SISTEMA POSLOVNE INTELIGENCIJE PROCESOM GRUPNOG ODLUČIVANJA

Rezime

Cilj ove doktorske disertacije jeste da ukaže i objasni, kako se procesom grupnog odlučivanja može unaprediti sistem poslovne inteligencije. Sama reč inteligencija u terminu poslovna inteligencija, izvorno označava izveštavanje ili obaveštavanje namenjeno podršci poslovnom odlučivanju. Ovako koncipirana, ova inteligencija pospešuje snalaženje u novim situacijama i brzinu i tačnost rešavanja problema, pri čemu se od dominantne, psihološke upotrebe termina inteligencija razlikuje po tome što se ne odnosi na sposobnost ličnosti kojom se služi pojedinac, već na tehnološko rešenje koje koristi kompanija. Poslovna inteligencija nije ni proizvod, ni sistem. To je arhitektura i kolekcija integrisanih operativnih aplikacija i aplikacija za potporu odlučivanju. Ona predstavlja korišćenje svih potencijala podataka, samim tim i informacija i znanja u kompaniji radi donošenja boljih i ispravnih poslovnih odluka i u skladu s tim, identifikaciju novih poslovnih mogućnosti.

Podrška odlučivanju individualnom donosiocu odluke je itekako važna aktivnost za uspešno upravljanje u svakoj kompaniji. Međutim, kako rad i funkcionisanje poslovnog sistema postaju sve složeniji a samim tim i odgovornost za donetu odluku sve je veća, postoji potreba za uvažavanjem različitih ideja i mišljenja članova tima, odnosno pripremljenog kvantuma znanja, sve zarad donošenja odluka u grupi. Jer, matematički se egzaktno pokazuje da je zbir znanja članova tima po definiciji uvek veći od znanja bilo kog člana tima. Jedini preduslov jeste da postoji komunikacija i deljenje resursa među članovima tima. Sa druge strane, evidentna je potreba za poboljšanjem kvaliteta odluka donetih u grupi. Ista predstavlja stalni predmet interesovanja istraživača sa velikim praktičnim i teoretskim značajem. stoga će posebno biti tretirana problematika akvizicije, skladištenja i distribucije znanja. Jer će *znanje* provejavati kroz sva poglavlja disertacije i činiće vezivno tkivo rukopisa.

Ključne reči: Odlučivanje, odluka, poslovna inteligencija, sistem poslovne inteligencije, znanje, menadžment znanja, grupno odlučivanje.

Naučna oblast: Modeliranje poslovnih sistema i poslovno odlučivanje

Uža naučna oblast: Poslovno odlučivanje

UDK broj: 519.816

IMPROVING SYSTEM OF BUSINESS INTELLIGENCE THROUGH GROUP DECISION MAKING PROCESS

Abstract

The objective of the doctoral thesis is to indicate and explain how the system of business intelligence could be upgraded by the means of group decision making process. The term intelligence in the context of business intelligence originally denotes reporting or information envisaged to support business decision making. Conceived in this way, business intelligence facilitates orientation in new situations and contributes to accelerated and accurate problem solutions; however, it differs from the prevailing, psychological meaning of the term in the sense that here it does not relate to a personal capacity that an individual can make use of, but to a technological solution applied by a company. Business intelligence shall be regarded neither as a product nor a system. Actually it is an architecture and collection of integrated operative applications and decision supporting applications. Business intelligence represents the usage of all data potentials, consequently the corporate information and knowledge aimed at making better and appropriate business decisions, and therefrom the ability to identify new business opportunities.

The support to an individual decision maker is an activity of utmost importance for any successful corporate management. However, since business system operations and functioning gradually gain in complexity and therefore the responsibility for decisions taken becomes ever larger, there is a necessity to take account of all various ideas and attitudes shared by team members, namely the prepared quantum of knowledge, all for the purpose of group decision making. Moreover, it is a mathematically proved fact that the sum of knowledge of all team members by definition always exceeds the knowledge of any team member. The only prerequisite here is the existent communication and share of resources among team members. On the other hand, the need for quality improvement concerning group decision making is evident. This represents an ongoing topic of interest for researchers with considerable practical and theoretical knowledge. Therefore, the issues of knowledge acquisition, storage and distribution will be dedicated special attention and treatment; knowledge will underlay all chapters of the thesis and will make the connective tissue of the paper.

Key words: Decision making, decision, business intelligence, system of business intelligence, knowledge, knowledge management, group decision making.

Academic Expertise: Modelling business systems and business decision making

Major in: Business decision making

UDC: 519.816

SADRŽAJ

	strana
1.0. UVOD	1
2.0. DEFINISANJE PREDMETA I CILJA ISTRAŽIVANJA	8
2.1. Predmet istraživanja	8
2.2. Ciljevi istraživanja	9
2.3. Pregled dosadašnjih rezultata istraživanja u domenu na koji se disertacija odnosi	9
2.4. Polazne hipoteze	10
2.5. Metode istraživanja	11
2.5. Očekivani rezultati istraživanja	12
3.0. FENOMEN ODLUČIVANJA	13
3.1. Savremeno odlučivanje	13
3.2. Grupno odlučivanje	16
4.0. VAŽNOST PODATKA U POSLOVNOJ INTELIGENCIJI	24
4.1. Podatak kao osnova poslovne inteligencije	24
4.2. Važnost kvaliteta podataka u poslovnoj inteligenciji	36
4.3. Problem nedostajućih podataka	43
4.3.1. Tradicionalni pristupi za rešavanje problema nedostajućih podataka	48
4.3.2. Predviđanje nedostajućih vrednosti	49
5.0. SISTEM POSLOVNE INTELIGENCIJE (SPI)	56
5.1. Uvod u poslovnu inteligenciju	56
5.2. Sistemi poslovne inteligencije	68
5.3. Koraci projektovanja skladišta podataka	73
5.3.1. Aktuelni pristupi dizajnu skladišta podataka	76
5.3.2. Zahtevi u razvoju skladišta podataka	78
5.3.3. Faze dizajna skladišta podataka	84
5.4. ETL (Extraction, Transformation, Loading) procesi	93
5.5. Otkrivanje zakonitosti u podacima	99
6.0. IMPLEMENTACIJA SISTEMA POSLOVNE INTELIGENCIJE	106
6.1. Razvoj skladišta podataka	106
6.1.1. Izgradnja baze SQL Servera	108

6.1.2. Učitavanje nomenklatura (šifarnika) MUPa u bazu	110
6.1.3. Razdvajanje podataka iz tekstualnog fajla u tri nova tekstualna fajla (Nezgodna, Vozilo i Lice)	110
6.1.4. Definisavanje strukture tekstualnih fajlova Nezgodna, Vozilo i Lice	112
6.1.5. Definisavanje tipa podataka u tekstualnim fajlovima	115
6.1.6. Provera validnosti podataka (logička kontrola)	116
6.1.7. Upisivanje tekstualnih fajlova u odgovarajuće tabele SQL Servera	117
6.1.8. Formiranje konekcije ka SQL Server bazi podataka	118
6.1.9. Formiranje strukture analitičke baze podataka	119
6.1.10. Učitavanje i procesiranje podataka u analitičkoj bazi	119
6.1.11. Izvoz analitičke baze podataka	121
6.2. Rad sa analitičkom bazom podataka o saobraćajnim nezgodama	123
6.2.1. Rad sa bazom u Microsoft Analysis Service	123
6.2.2. Rad sa bazom u Microsoft Excel	126
6.2.3. Primeri rada sa analitičkom bazom	130
7.0. MENADŽMENT ZNANJA U ORGANIZACIJI	135
7.1. Znanje	135
7.2. Menadžment znanja	139
7.3. Koraci projektovanja sistema menadžmenta znanja	145
7.4. Tehnike modelovanja znanja	156
7.5. Studija slučaja – uvođenje menadžmenta znanja	162
8.0. GRUPE I DRUŠTVENE MREŽE	168
8.1. Formalizmi za predstavljanje društvenih mreža	168
8.2. Važne osobine društvene mreže za poslovnu inteligenciju	173
9.0. MOGUĆNOSTI EKSTREMIZACIJE ZNANJA	181
10. ZAKLJUČAK	196
11. LITERATURA	200
INDEKS SLIKA	211
INDEKS TABELA	214

1. UVOD

Postojeći uslovi funkcionisanja poslovnog sistema, karakteriše visok stepen dinamičnog i sve više neizvesnijeg okruženja, bez obzira na veličinu, počev od malih poslovnih sistema do velikih multinacionalnih kompanija. Konstantno se traži način kako da svoja poslovanja učvrste, povećaju i poboljšaju. Tu se ne misli samo na povećanje proizvodnje, osvajanje novih tržišta, povećanje asortimana novih i različitih proizvoda, već i na povećanje efikasnosti i efektivnosti samog poslovnog sistema, putem bolje organizovanosti i prilagođavanju poslovnih i proizvodnih procesa tržištu. Sasvim je sigurno da je u savremenom menadžmentu sazrelo mišljenje da je ključni resurs i komplementarna prednost pre svega posedovanje informacije i znanja.

Tako je i cilj ove disertacije da definiše pojam poslovne inteligencije i prikaže osnovne komponente i karakteristike sistema poslovne inteligencije. Povezanost Interneta i srodnih tehnologija otvara korisnicima velike mogućnosti upravljanja informacijama na način koji osigurava najveću poslovnu iskorišćenost. Istraživanja brojnih autora danas pokazuju da se u preduzećima bez sistema poslovne inteligencije analizira samo 10% podataka, a sa sistemom poslovne inteligencije procenat analize se povećava na 90%, videti [88].

Sama reč inteligencija u terminu poslovna inteligencija, izvorno označava izveštavanje ili obaveštavanje namenjeno podršci poslovnom odlučivanju. Ovako koncipirana, ova inteligencija pospešuje snalaženje u novim situacijama i brzinu i tačnost rešavanja problema, pri čemu se od dominantne, psihološke upotrebe termina inteligencija razlikuje po tome što se ne odnosi na sposobnost ličnosti kojom se služi pojedinac, već na tehnološko rešenje koje koristi kompanija, videti [127].

Prethodnoj konstataciji ide na ruku činjenica, da je jedna od osnovnih karakteristika savremenog globalizovanog sveta permanentni ubrzani protok nesagledive količine podataka. Količine podataka rastu veoma brzo, a posledica tog rasta je da je veoma mala verovatnoća da će ti podaci ikad biti pregledani od strane nekog analitičara, inženjera znanja, donosioca odluke ili menadžera. Osim potrebe za nalaženje novih načina za čuvanje i skladištenje tih podataka, potrebno je i nalaziti načine za pretvaranje tih sirovih podataka u korisne informacije, koje će igrati značajnu ulogu u daljem procesu poslovnog planiranja i donošenja odluka. Otkrivanje zakonitosti u podacima - Data mining je proces automatskog otkrivanja korisnih informacija iz velikih količina podataka i danas postoji razvijen veliki broj aplikacija koje koriste upravo ovaj pristup.

Sam pojam poslovne inteligencije (engl. *Business Intelligence*) objedinjuje metodologije, tehnologije i platforme za skladištenje podataka (*Data Warehousing*), OLAP procesiranje podataka (*On-line Analytical Processing*) i rudarenje podataka (*Data Mining*). Iste omogućavaju menadžmentu poslovnog sistema, kreiranje korisnih upravljačkih informacija iz podataka o poslovanju koji se nalaze na različitim transakcijskim sistemima i dolaze iz različitih internih i eksternih izvora.

U stvarnosti, poslovna inteligencija je s *jedne* strane način poslovnog ponašanja koji omogućava, da se poslovne odluke na svim nivoima odlučivanja donose i zasnivaju na relevantnim i ažurnim poslovnim informacijama a samim tim i znanjem, a ne na predosećaju i subjektivnom utisku. Sa *druge*, informatičke strane, poslovna inteligencija je složeni informacioni sistem koji automatizovanim procedurama prikuplja podatke iz različitih izvora, obrađuje ih, transformiše i integriše a korisnicima omogućava pristup kvalitetnim informacijama na intuitivan i lako razumljiv način.

Svaki poslovni sistem suočava se sa velikom količinom podataka koje sakuplja tokom godina poslovanja. *Sa druge strane, kompanije se guše u podacima a žedne su znanja*, videti [88]. Podaci koji se sakupljaju dolaze iz različitih geografskih područja, različitih demografskih struktura, svakodnevnog poslovanja i od samih potrošača. Svi ti podaci se analiziraju kako bi dali pravilnu sliku o trenutnoj poziciji poslovnog sistema i okruženja u kome se nalazi. Podaci mogu da budu analizirani na takav način, da obezbede menadžmentu poslovnih sistema, jasnu sliku o pravcima razvoja i odlukama koje treba da se donesu.

Tako dakle, poslovna inteligencija kao rešenje, sadrži tehnologije i proizvode čiji je cilj da obezbede informacionu podršku, ukoliko postoje poslovne nedoumice, kada treba doneti ispravne poslovne odluke, detaljnije u [128]. Koristi se na svim nivoima menadžmenta (operativnom, taktičkom ili strateškom), što je razumljivo jer je znanje imanentno kompletnom fenomenu odlučivanja.

Pojam poslovne inteligencije je takođe široko definisan. Različiti autori ovaj pojam definišu na različite načine. Ono što je svima zajedničko je, da poslovna inteligencija:

- ✓ objedinjava nekoliko vrlo važnih metodologija, koncepata i tehnologija putem kojih se može poboljšati proces odlučivanja,
- ✓ koristi sisteme za podršku odlučivanju utemeljene na poslovnim činjenicama i podacima,
- ✓ koristi informacije dobijene iz strukturiranih i nestrukturiranih podataka iz internih i eksternih izvora, koje se pomoću navedenih tehnologija transformišu u kvalitetne

informacije, primenjive u procesu odlučivanja i vizualno se predočavaju krajnjem korisniku.

Najcelishodnija definicaja poslovne inteligencije i jedna od danas najcitiranijih je definicija Hauarda Dresnera, analitičara Gartner grupe: „*Poslovna inteligencija je proces transformacije podataka u informacije i kroz proces saznanja - transformacija tih informacija u znanje*“.

Dresnerova definicija poslovnu inteligenciju određuje koristeći tri ključna pojma: *podatak, informacija i znanje*.

Podatak

Podatak je bilo koja činjenica, broj ili tekst koji može biti obrađen i ažuriran tokom poslovne obrade. Danas, organizacije nagomilavaju ogromne količine podataka u različitim formatima i različitim bazama. Ovo uključuje:

- ✓ *operativne i transakcione* podatke kao što su podaci o prodaji, troškovima, naplatama, računovodstvu,
- ✓ *neoperativni* podaci, kao što je industrijska prodaja, predviđanje podataka i makroekonomski podaci,
- ✓ *meta* podaci, podaci o podacima, kao što je logički dizajn baze ili rečnik baze podataka.

Informacija

Paterni, asocijacije ili odnosi između svih tih podataka mogu proizvesti informacije. Na primer, analiza podataka o prodaji može proizvesti informacije na osnovu kojih se mogu videti koji proizvodi se prodaju, pod kojim uslovima i kada.

Znanje

Informacija može biti pretvorena u znanje o nekim prethodnim paternima i budućim trendovima. Na primer, sumarna informacija o prodaji u nekom supermarketu može biti analizirana kako bi se došlo do saznanja o potrošačkim navikama kupaca. Tako da, proizvođač ili prodavac može utvrditi koji proizvodi su najpogodniji za promociju, itd. Kada postoji nagomilano znanje, sjedinjeno sa bihevizmom i radnim iskustvom menadžera, sledi mogućnost generisanja i otkrivanja mudrosti.

Fenomen poslovne inteligencije može se posmatrati sa dva aspekta - makro i mikro aspekta. Posmatrana sa makro aspekta, poslovna inteligencija označava složenu agregiranu kategoriju, koja se stvara sistematskim, ali ne ciljanim prikupljanjem podataka o makro ekonomskim kretanjima u određenoj geopolitičkoj sredini. Ona, takođe, podrazumeva njihovo organizovanje i strukturirano beleženje, kao i logičko-računsku obradu radi

otkrivanja trendova. Danas posebnu pažnju inženjera znanja sve više pobuđuje fenomen poslovne inteligencije posmatran sa mikro aspekta. U ovom slučaju se radi o otkrivanju prikrivenih znanja iz poslovnih podataka, koje neka organizacija prikuplja rutinski, obavljajući svoje svakodnevne poslovne transakcije.

Poslovna inteligencija nije ni proizvod, ni sistem. To je arhitektura i kolekcija integrisanih operativnih aplikacija i aplikacija za potporu odlučivanju, te baza podataka koje poslovnim korisnicima omogućuju lak pristup poslovnim podacima. Ona predstavlja korišćenje svih potencijala podataka i informacija u kompaniji radi donošenja boljih poslovnih odluka i u skladu s tim identifikaciju novih poslovnih mogućnosti. Dok neka nekonvencionalna lokalna biblioteka može do izvesne granice da ne obraća pažnju na bitne vrednosti dokumenata koje sakuplja i da ostavi korisniku da sam odvajava žito od kukolja. Informaciona osnova na kojoj počiva poslovna inteligencija to sebi nikako ne može dozvoliti. Kod nje je presudan kvalitet informacije koja je ušla u memoriju bez obzira da li je ona naučna, tehnička, socijalna, ekonomska, demografska, politička, ili neka druga svera. Informatički podržanu poslovnu inteligenciju moguće je zamisliti kao ogromnu mašinu za sortiranje, analiziranje, kritikovanje, ocenjivanje i sintezu neobrađenih podataka, koji na kraju tih procesa postaju sama suština ljudskog znanja.

Pored toga, poslovna inteligencija je način dostavljanja pravih informacija u pravom formatu u prave ruke u pravo vreme. Dobar sistem poslovne inteligencije prikuplja informacije iz svih delova organizacije, analizira ih, priprema potrebne izveštaje i dostavlja ih menadžerima koji ih koriste. Na taj način svaki pojedinac dobija informacije skrojene prema sopstvenim potrebama.

Poslovna inteligencija je izvorno bila namenjena donosiocima odluka, analitičarima u preduzeću, menadžerima. Međutim u savremenim preduzećima odlučuje se na svim nivoima. To znači da je informacija i znanje potrebno svima. Danas je poslovna inteligencija u donošenju poslovnih odluka nezamenljiv koncept.

Sa druge strane centralna tema u okviru ove doktorske disertacije biće posvećena metodologiji podrške grupnom odlučivanju. Osnovna ideja jeste, kako je konceptima grupnog odlučivanja (GO), moguće unaprediti pomenuti sistem poslovne inteligencije. Kao što je već prethodno istaknuto, osnova sistema poslovne inteligencije jeste, sled: *podatak* → *informacija* → *znanje*. Sa druge strane, evidentno je da je znanje osnova za donošenje svake ispravne upravljačke poslovne odluke.

Podrška odlučivanju individualnom donosiocu odluke (DO) je itekako važna aktivnost za uspešno upravljanje u svakoj kompaniji. Međutim, kako rad i funkcionisanje poslovnog sistema postaju sve složeniji a samim tim i odgovornost za donetu odluku sve je veća, postoji potreba za uvažavanjem različitih ideja i mišljenja članova tima, odnosno pripremljenog kvantuma znanja, sve zarad donošenja odluka u grupi. Jer, matematički se egzaktno pokazuje da je *zbir znanja članova tima po definiciji uvek veći od znanja bilo kog člana tima*. Jedini preduslov jeste da postoji komunikacija i deljenje resursa među članovima tima.

Takođe, stalne promene u neposrednom okruženju preduzeća rezultuju da njegovo funkcionisanje i opstanak na tržištu postanu sve složeniji. Samim tim i odgovornost za donetu odluku postaje veća, što rezultuje zahtevom da se ista donosi od strane specijalizovanog tima za odlučivanje. Sve sa ciljem posmatranja više alternativnih pravaca i ideja, kao i međusobnog usaglašavanja i uvažavanja različitih mišljenja, radi donošenja najprihvatljivije odluke u grupi.

Sa druge strane, evidentna je potreba za poboljšanjem kvaliteta odluka donetih u grupi. Ista predstavlja stalni predmet interesovanja istraživača sa velikim praktičnim i teoretskim značajem. Posebno se tretira problematika akvizicije, skladištenja i distribucije znanja. Sve je to razlog da se razvija savremena discipline pod nazivom Menadžment znanja. U okviru ove disertacije posebno će biti posvećena pažnja oblastima kao što su: menadžmentu znanja i mogući modeli ekstremizacije znanja.

Disertacija je planirana da se realizuje po sledećim poglavljima.

Nakon *uvoda*, u kome se daje kratak prikaz sveobuhvatnog poduhvata i svih izazova istraživanja, sledi *drugo poglavlje* u kome se definiše *predmet i cilj istraživanja*. U ovom poglavlju pored predmeta i cilja istraživanja, definiše se pregled dosadašnjih rezultata u oblasti istraživanja, zatim polazne hipoteze, metode iztraživanja, kao i očekivani naučni i stručni doprinosi.

Treće poglavlje opisuje fenomen odlučivanja, kroz određivanje mesta odlučivanja u savremenom menadžmentu, njegovog interdisciplinarnog okvira, vreste odlučivanja i prelaza u poslovnu inteligenciju. Posebno se ukazuje na objedinjavanje fenomena odlučivanja u savremenu naučnu disciplinu pod nazivom *teorija odlučivanja*. Daje se kratak istorijski pregled fenomena odlučivanja, kao i modela odlučivanja. Ukazaće se na jaku vezu trojnosti kvalitativnog, kvantitativnog i informatičkog pristupa u odlučivanju. U nastavku ovog poglavlja, imajući u vidu naslov teze, biće posebno tretirano grupno odlučivanje i njegovi najvažniji elementi, kao što su *grupa, proces grupnog odlučivanja* i rezultat, odnosno *grupna odluka*.

Sledeće, *četvoro poglavlje* tretira važnost i ulogu *podatka* u poslovnoj inteligenciji. Razlog tome jeste činjenica da je podatak nosilac svih informacija a samim tim i znanja. Poseban akcenat će biti stavljen na kvalitet podataka, jer od toga naprosto zavisi i informacija, odnosno dobijeno znanje, na osnovu koga treba doneti ispravnu upravljačku poslovnu odluku. Jer od dobre odluke, po pravilu se ostvaruju i dobri rezultati. U nastavku ovoga poglavlja biće ukazano na jedan gorući problem, a to je probleme šta raditi sa nedostajućim podacima. Za ovakav realan problem, biće predstavljena dva pristupa, a to su: jedan tradicionalni pristup i jedan pristup predviđanja nedostajućih podataka.

Peto poglavlje treba da ukaže na celovit osvrt na sisteme poslovne inteligencije. Postavljeni model poslovne inteligencije proširuje se novim alatima za podršku odlučivanju, sa posebnim osvrtom na grupu, odnosno članove tima koji u procesu odlučivanja, treba da predoče najprihvatljivije rešenje. Na početku ovog poglavlja biće dat osnovni koncept poslovne inteligencije. Zatim se isti sistemski proširuje i kompletira sistem poslovne inteligencije. Naravno osnova za svaki sistem poslovne inteligencije jeste bogata baza podataka ili skladište podataka. To je dovoljan razlog da se u ovom poglavlju posebna pažnja posvećuje osnovnim koracima za projektovanje i razvoj skladišta podataka. Kao veza u odnosu na prethodno poglavlje, posebno se ukazuje na proces ekstrakcije, transformacije i učitavanja podataka. Na samom kraju ovoga poglavlja, biće ukazano na osnove otkrivanja znanja u samim podacima.

U okviru *šestog poglavlja*, daje se prikaz aplikativnog dela sistema poslovne inteligencije. Ukazuje se kako je moguće upotrebiti sistem poslovne inteligencije ako su podaci i sama baza podataka, atributi koji opisuju saobraćajne nezgode. Konkretno, radi se delu baze podataka koja tretira saobraćajne nezgode na teritoriji grada Beograda. Podaci su transformisani i u bezličnoj su formi, tako da je akcenat na aplikativnoj formi sistema poslovne inteligencije. U prvom delu ovog poglavlja biće prikazan sam razvoj skladišta podataka, sa posebnim osvrtom na triplet, nezgoda, vozilo i lice. U drugom delu ovog poglavlja biće prikazan sam rad sa analitičkom bazom, i to kroz dva softverska alata. Na samom kraju ovoga poglavlja biće ukazano na sam primer rada sa analitičkom bazom podataka.

Sledeće *sedmo poglavlje* prestavlja logički nastavak prethodnog, sa posebnim osvrtom na menadžment znanja i to preko celina modelovanja i upravljanja znanjem. Jer će kroz disertaciju i sva poglavlja, vezivno tkivo biti upravo znanje. Zato će u ovom poglavlju na početku biti ukazano na sam fenomen znanja. Nakon toga, a sa aspekta savremenog poslovanju kompanije, biće uveden koncept menadžmenta znanja, kao i osnovni koraci za

projektovanje menadžmenta znanja. Na kraju ovoga poglavlja, biće ukazano na tehnike modelovanja znanja, kao i na samu studiju slučaja o mogućnostima uvođenja menadžmenta znanja.

Osmo poglavlje treba da stavi akcenat na grupe i društvene mreže, jer se radi o unapređenju sistema poslovne inteligencije, radom grupnog odlučivanja. Upravo zbog samog naslova disertacije, na početku ovog poglavlja prvo će biti predstavljene neki osnovni formalizmi za prikaz i predstavljanje društvene mreže, jer grupa u procesu grupnog odlučivanja, naprosto je jedan vid društvene mreže. Pored toga u ovom poglavlju, poseban akcenat će biti stavljen na prikaz važnih osobina društvene mreže sa aspekta poslovne inteligencije.

Deveto poglavlje će biti pokušaj da se modelima i metodama ekstremizacije na grafovima pokuša doći do odgovora: Koji je to najkraći put u grupi za prenos i distribuciju kvantuma znanja. Biće uveden jedan jasan matematički formalizam, kao pandam problemu trgovačkog putnika. S tim što resurs za distribuciju nije ništa drugo nego upravo znanje, a umesto lokacija ili gradova prolaska kretanje se realizuje od jednog do drugog donosioca odluke u društvenoj mreži. Biće prikazana studija slučaja sa jednim hipotetičkim pristupom sa aspekta koeficijenta efikasnosti u odnosu na znanje koje poseduje svaki donosilac odluke i potrebno vreme za donošenje odluke po svakom učesniku sesije.

Na kraju će biti predložen *zaključak* sa daljim pravcima istraživanja, kao i korištena *literatura*.

2. DEFINISANJE PREDMETA I CILJA ISTRAŽIVANJA

U ovom delu disertacije osvrt je na definisanje predmeta istraživanja, zatim ciljeva istraživanja, polazne hipoteze istraživanja, metode koje će se koristiti tokom istraživanja kao evidentiranje očekivanih naučnih i stručnih doprinosa.

2.1. Predmet istraživanja

Tendencija razvoja i napredka savremenog društva praćena je razvojem raznih oblika savremenih tehnologija u sferi globalnog poslovanja. Samo poslovanje svakog poslovnog sistema beležilo se u bazi podataka posmatranog poslovnog sistema. Ti podaci su narastali iz jedne u drugu poslovnu godinu. Odnosno, u životnom ciklusu svakog poslovnog sistema, stižu se i prikupljaju razni poslovni podaci kako o njegovom internom poslovanju tako i o integraciji sa okruženjem. Sasvim je sigurno da će uspešan razvoj posmatranog privrednog subjekta biti sigurniji i lakši ukoliko se na vreme prikupljeni podaci obrade i pretoče u korisne informacije, koje mogu poslužiti kao smernice za budući period razvoja. Time je posmatrani poslovni sistem fleksibilniji kako na unutrašnje tako i na spoljne poremećaje.

Prikupljanje podataka u adekvatnu bazu podataka (BP), kroz faze razvoja posmatranog poslovnog sistema, karakteriše velika količina podataka. Tako BP kompanije narasta po svom obimu u tzv. skladište podataka, koje prati poslovanje preduzeća. Kako se za strategijsko odlučivanje, uglavnom ne koriste svakodnevne operativne rezultati poslovanja poslovnog sistema, već zbirni i izvedeni iz predhodnog vremenskog perioda, skladište podataka predstavlja dobru osnovu za poslovno odlučivanje. Poznato je da nivo strategijskog poslovnog odlučivanja najčešće prate nestrukturirani problemi, ide na ruku da skladište podataka postane dobra osnova za savremeno poslovno odlučivanje.

Ekstrahovani relevantni podaci iz pomenutog skladišta podataka, praktično, podaci koji nose najveću količinu informisanosti za prirodu problema koja se rešava, postaju odlična osnova o kojoj grupa ili tim za odlučivanje raspravljaju, radi predloga najprihvatljivije alternative.

Samim tim predmet doktorske disertacije predstavlja mogućnost unapređenja sistema poslovne inteligencije procesom grupnog odlučivanja. Pri tome su na raspolaganju svi koncepti grupnog odlučivanja, pre svega na ovom mestu tretirane relevantne metode i modeli, sve sa ciljem sveobuhvatnog pristupa, pravljenja brojnih analiza, zarad predloga

najprihvatljivijeg rešenja. Jer pomenuto skladište podataka kao noviji tehnološki koncept, ima upravo ulogu da objedini srodne podatke poslovanja, vitalnih funkcija preduzeća u oblik pogodan za sprovođenje raznih analiza. Tako se nasuprot svakodnevnoj *operativnoj* obradi podataka javlja potreba za *analitičkom* obradom podataka.

2.2. Ciljevi istraživanja

Cilj istraživanja je sprovođenje celovitog istraživačkog poduhvata, unapređenja sistema poslovne inteligencije, procesom grupnog odlučivanja. Gradi se opšte primenljiv model sistema poslovne inteligencije, integrisanog preko skladišta podataka, ekstrahovanog podatka, koji nosi najveću informisanost i njegovo prevođenje u protumačenu informaciju, odnosno znanje. Pomenuta informacija, samim tim i znanje, postaje odlična osnova i predmet rada, tima ili grupe, koja preko menadžmenta znanja, treba da konvergira u predlog najprihvatljivijeg rešenja.

Naučni cilj rada se ogleda u definisanju i razvoju modela i metoda grupnog odlučivanja uz pomoć kojih se može unaprediti sistem poslovne inteligencije. Prikazuju se osnove odlučivanja, poseban osvrt na grupno odlučivanje i timski rad, zatim preko osnovnog modela poslovne inteligencije njegovo umreženje u sisteme poslovne inteligencije. Menadžment znanja i modeli ekstremizacije znanja treba da na najbolji mogući način naprave kopču od sistema poslovne inteligencije do grupnog odlučivanja. Od stepena simbioze pomenuta dva sistema u celini zavisi stepen i kvalitet predloženog rešenja, odnosno sam kvalitet poslovne odluke.

2.3. Pregled dosadašnjih rezultata istraživanja u domenu na koji se disertacija odnosi

Na prvi pogled pomenuti procesi mogu da izgledaju jednostavno, jer se odavno odlučivalo na svim nivoima a sa druge strane pojavom sistema poslovne inteligencije taj proces bi trebao da bude unapređen. Upravo iz toga razloga izučavanje pomenutih procesa predstavlja izuzetno atraktivnu disciplinu, sa permanentnim zahtevom za unapređenjem.

Tokom pripreme na prikupljanju pisanog naučnog i stručnog materijala iz ove oblasti, može se konsatovati da i u domaćoj a pogotovu u stranoj literaturi ponuda je jako dobra. Brojne su reference koje jako dobro prikazuju fenomen savremenog odlučivanja sa posebnim osvrtom na timski rad i grupno odlučivanje. Neke od domaćih su [29], [129], itd. Što se tiče stranih referenci, to su pre svega [37], [23], itd. Opisan je proces formiranja grupe, njenog

rada tokom sesije, zarad ostvarenja definisanog poslovnog cilja. Proces je formalizovan, dat je na početku i jedan bihevioristički pristup, da bi se na kraju sve to kvantifikovalo.

Sa druge strane, proces poslovne inteligencije i sistema poslovne inteligencije je novijeg datuma istraživanja pa su i reference u domaćoj literaturi nešto siromašnije. Naveo bih ipak jednu, koja tretira problematiku poslovne inteligencije i sistema za podršku odlučivanju, kao što je [127].

Mnogo je brojnija i raznovrsnija strana literatura iz ove oblasti. Posebno ističem, [20], [88], [90], itd, sa jasnim i konciznim aspektom procesa poslovne inteligencije.

Ono što je posebno interesantno, jeste, da nema zajedničkog pogleda na pomenute procese. Time se još jednom pokazuje opravdanost istraživanja, i to: *kako je moguće procesom grupnog odlučivanja unaprediti sistem poslovne inteligencije.*

2.4. Polazne hipoteze

Opšta hipoteza razvijena u okviru istraživanja, polazeći od postavljenih ciljeva i zadataka istraživanja glasi:

H1: Primenom procesa višekriterijumskog grupnog odlučivanja, moguće je unaprediti sistem poslovne inteligencije.

Posebne hipoteze:

H1.1: Moguće je identifikovati ključne elemente sistema poslovne inteligencije i sve veze između elemenata u sistemu.

H1.2: Primenom metodologije podrške višekriterijumskog grupnog odlučivanja, moguće je kvalitetno upotrebiti podatke ekstrahovane iz skladišta podataka.

H1.3. Moguće je sistemom menadžmenta znanja napraviti akviziciju, skladištenje i distribuciju informacija i znanja članovima tima za odlučivanje.

H1.4. Primenom algoritama ekstremizacije, moguće je izračunati najkraće puteve za prenos znanja među članovima tima.

2.5. Metode istraživanja

Osnovne metode istraživanja koje će biti korištene tokom izrade doktorske disertacije, baziraju se na postojećim teorijskim rezultatima i eksperimentalnom radu u navedenoj oblasti. Samim tim, biće realizovano sakupljanje i proučavanje dostupne literature, njena analiza i sistematizacija, a sve sa ciljem da se pokaže opravdanost i korisnost razvoja modela grupnog odlučivanja za unapređenje sistema poslovne inteligencije.

Aplikativni deo biće realizovan kroz dva pristupa. Prvi će se odnositi na projektovanje i razvoj sistema poslovne inteligencije nad podacima koji prate saobraćajne nezgode u jednom gradu. drugi aplikativni deo rada će se ogledati u razvoju i primeni algoritama za pronalaženje ekstremne vrednosti puta prenosa znanja. Pomenuta količina znanja, na raspolaganju je članovima tima radi pronalaženja najprihvatljivijeg rešenja, tokom sesije grupnog odlučivanja. Time će kompletan sistem poslovne inteligencije dobiti jednu sasvim novu komponentu sa aspekta uvećanog (nagomilanog) znanja, kao ključnog preduslova za donošenje kvalitetnih i ispravnih poslovnih odluka.

Osnovne metode istraživanja koji će se koristiti pri rešavanju postavljenog problema su sledeće:

- ✓ metoda analize će se upotrebljavati kroz postupak naučnog istraživanja, raščlanjivanjem složenih pojmova, sudova i zaključaka na njihove jednostavnije sastavne delove i elemente, odnosno kroz postupak mišljenja od posebnoga ka opštem,
- ✓ metoda deskripcije će se koristiti za opisivanje pojava i procesa od interesa, uz objašnjenja važnih obeležja opisivanih pojava i procesa, uočavanje zakonitosti i uzročnih veza i odnosa,
- ✓ komparativna metoda će se koristiti kroz postupak upoređivanja rezultata modela, a radi utvrđivanja njihove sličnosti u ponašanju i razlika među njima,
- ✓ primena metoda sinteze će se ogledati putem sinteze jednostavnih sudova u složenije i kroz proces uopštavanja, čime će se doći do sistematizovanog znanja, odnosno do izgradnje teorijskog znanja u pravcu od posebnog ka opštem,
- ✓ metoda merenja se koristi sa ciljem da se dobiju rezultati predloženih rešenja (modela), koji će služiti za poređenje kvaliteta različitih rešenja,
- ✓ metoda kompilacije će biti primenjena u smislu preuzimanja tuđih rezultata naučnoistraživačkog rada, odnosno tuđih opažanja, stavova, zaključaka i spoznaja, pri čemu će se ova metoda upotrebiti i u kombinaciji s drugim metodama u naučnoistraživačkom radu, a kako bi disertacija u najvećoj meri nosila lični pečat

autora, koji će, uz lični pristup pisanju naučnog dela korektno i na sistematičan način citirati sve ono što je od drugih preuzeo,

- ✓ matematička metoda se upotrebljava kroz sistemski postupak koji se sastoji u primeni matematičke logike, matematičkih relacija, matematičkih simbola i matematičkih operacija u naučnoistraživačkom radu,
- ✓ metoda modeliranja se sastoji u razvoju modela koji treba da predstavlja stvarnu pojavu, sa ciljem da se dobijeni rezultati i unapređenja modela mogu preneti i na realno stanje sistema,
- ✓ statistička metoda će se koristiti radi utvrđivanja statističke značajnosti dobijenih rezultata, itd.

Primena svakog pojedinačnog metoda, treba da rezultuje ostvarenjem primene opšte primenljivog modela, radi unapređenja sistema poslovne inteligencije procesom grupnog odlučivanja, za potrebe integralnog upravljanja u poslovnom sistemu.

2.6. Očekivani rezultati istraživanja

Doprinos koji se žele ostvariti ovom disertacijom jeste uvođenje i definisanje skladne procedure, uz pomoć koje je moguće uspešno primeniti kvantum znanja članova tima na rešavanje konkretnih problema sistemom poslovne inteligencije.

Shodno tome, očekivani rezultati predložene disertacije bili bi:

- ✓ Pregled savremenih modela koji mogu da se koriste za unapređenje sistema poslovne inteligencije procesom grupnog odlučivanja,
- ✓ Podizanje nivoa opšte stručne svesti o mogućnostima primene modela i tehnika grupnog odlučivanja u sisteme poslovne inteligencije,
- ✓ Jasan pregled i sistematizacija dosadašnjih istraživanja u ovoj oblasti,
- ✓ Kritički i analitički osvrt na relevantne dosadašnje pristupe,
- ✓ Projektovanje i razvoj jedinstvenog i jednoznačnog pristupa ekstremizacije protoka kvantuma znanja u sistemu poslovne inteligencije zarad rešavanja realnog problema.

Pored navedenih, sigurno je da se u doprinose može navesti povezivanje više naučnih disciplina, čij je zajednički imenilac *odlučivanje* u jedinstvenu i skladnu celinu.

3. FENOMEN ODLUČIVANJA

U ovom delu disertacije daje se pregled oblasti fenomena odlučivanja, radi odgovora na postavljenu temu i ranije definisane hipoteze istraživanja. Sledi prvo deo koji se odnosi na savremeno odlučivanje.

3.1. *Savremeno odlučivanje*

Odluka i odlučivanje se kao reči javljaju i koriste svakodnevno, ali kako često u takvim situacijama biva, te reči su u upotrebi, a da se pri tome i ne zna uvek njihovo pravo značenje. Da se radi o značajnoj ljudskoj delatnosti (iza svake odluke stoji neka delatnost), govori i činjenica da se više od osamdeset godina neguje i razvija posebna naučna disciplina zvana **teorija odlučivanja**, detaljno u [29].

Neki od pravaca istraživanja, polovinom prošlog veka, koji su doveli do razvoja pomenute discipline.

- ✓ razvoj teorije naučne organizacije rada;
- ✓ donosioci odluka su počeli da izučavaju bihevioristički aspekt odlučivanja;
- ✓ pojava teorije igara, unosi nove okvire za proveru odluka, kao i otkrivanje mera korisnosti koje predstavljaju ekonomske nivoe ljudskih želja za različite potrebe;
- ✓ sve veće uključivanje ekonomista sa svojim idejama u proces odlučivanja,
- ✓ najveća pokretačka snaga u razvoju discipline o odlučivanju svakako je razvoj niza kvantitativnih metoda, gde su se na interdisciplinarni način našla na okupu moćna znanja matematike i statistike,
- ✓ i naravno krajem prošlog i početkom ovoga veka hiper intenzivan razvoj informaciono komunikacione strukture, kao odličnog preduslova razvoja ove discipline.

Jedinstvo pomenutih pravaca istraživanja, rezultovalo je 50-tih godina prošlog veka, razvojem nauke o odlučivanju. Posebne karakteristike ovoga perioda nauke o odlučivanju, prema [29] su:

- ✓ programski pristup u izučavanju pravih vrednosti, uloge i ograničenja teorije odlučivanja;
- ✓ veliki naglasak u izučavanju i analizi okruženja, u okviru koga se vrši odlučivanje;
- ✓ pridaje se veći značaj dobijanju zadovoljavajućih rešenja za izvesne teže probleme, nego traženju optimalnih rešenja;

- ✓ ulažu se maksimalni naponi za što bolju integraciju kvantitativne analize sa analizama ponašanja i okruženja u rešavanju realnih problema;
- ✓ informaciona tehnologija se sve više koristi u procesima rešavanja polustrukturiranih i nestrukturiranih problema, i druge.

Posebno treba istaći sazrele preduslove potpunog razvoja ove discipline a to su: *kvalitativni aspekti, kvantitativni aspekti i naravno informaciono komunikacioni aspekti*. Ovaj trougao umnogome i do dana današnjeg u potpunosti zadovoljava sve koncepte razvoja savremenog odlučivanja kako na teoretskom tako i na aplikativnom nivou.

Sam fenomen odlučivanja, može da se formalizuje na sledeći način:

Kao što je istaknuto, brza tendencija razvoja prirodnih nauka uslovlila je jaku ekspanziju pre svega kvantitativnog i informatičkog pristupa u odlučivanju, i to na početku definisanjem osnovnog formalizma problema odlučivanja. Problem odlučivanja prema [129], je petorka (A, X, F, Θ, \succ) u kojoj je:

- A : konačan skup raspoloživih alternativa (akcija), koje učesnik sesije rangira u cilju izbora najprihvatljivije;
- X : skup mogućih rezultata koji slede kao posledica izbora alternative;
- Θ : skup stanja sveta, zavisi od nepoznatog stanja sveta $\theta \in \Theta$, jer se posledice izbora alternative $a \in A$ mogu razlikovati;
- $F : A \times \Theta \rightarrow X$ određuje za svako stanje sveta θ i za svaku alternativu a , rezultujuću posledicu $x = F(a, \theta)$
- \succ : relacija slabog poretka na X , tj. binarna relacija koja ispunjava sledeće uslove:
 - (i) $x \succ y$ ili $y \succ x, \forall x, y \in X$
 - (ii) \succ je tranzitivna, tj. $x \succ y, y \succ z \Rightarrow x \succ z$.

Relacija \succ karakteriše donosioca odluke i naziva se relacija preferencije. Stroga preferenca $x \succ y$, znači da važi $x \succ y$, ali ne i $y \succ x$. Relacija indiferentnosti $x \sim y$ znači da važi $x \succ y$ i $y \succ x$. Najčešći način rešavanja problema odlučivanja jeste transformacija slabog poretka \succ na X u uobičajeni poredak \geq nad realnim brojevima pomoću funkcije korisnost. U uslovima generalnog problema odlučivanja, pretpostavlja se da je stanje sveta θ poznato, X je višedimenzionalno i poznato za svaku alternativu kao skup relevantnih vrednosti atributa.

Sam proces odlučivanja predstavlja metodološki logičan skup faza i aktivnosti koje omogućavaju sistemsku analizu i rešavanje realnog problema. Skoro da nema reference i autora iz ove oblasti koji na izvestan način ne definiše algoritam sa fazama i aktivnostima procesa odlučivanja. Lee i Moore proces odlučivanja definišu sa sledeće četiri osnovne faze, prikazane u [29]:

- ✓ formulisanje problema;
- ✓ kreiranje modela;
- ✓ rešavanje modela;
- ✓ primena rešenja.

Druga pak, nešto šira procedura za definisanje faza procesa odlučivanja, prikazana u [29], sastoji se iz sledećih koraka:

- ✓ otkrivanje i formulisanje problema;
- ✓ kreiranje mogućeg rešenja;
- ✓ izgradnja modela;
- ✓ određivanje rezultata (posledica);
- ✓ izbor sistema vrednosti, i
- ✓ donošenje odluke.

Odlučivanje kao ljudski fenomen, sa aspekta nivoa i stepena složenosti problema koji se rešava, kao i broja učesnika koji ga rešavaju prema [29], se deli u sledeće kategorije:

✓ *Individualno (pojedinačno) odlučivanje.*

Ovaj oblik odlučivanja ujedno je najjednostavniji i najčešće objašnjavan i korišćen do sada. Brojne reference ukazuju na to, sa zajedničkim imeniocem, gde su aktivnosti procesa odlučivanja poverene samo jednom donosiocu odluke.

✓ *Grupno odlučivanje - timski rad.*

Karakteriše ga veći nivo faza i aktivnosti procesa odlučivanja. Učestvuje veći broj (tim) donosioca odluke (učesnika sesije) u proceduri izbora najprihvatljivije alternative. Predmet ove disertacije upravo su metode, modeli, čijom primenom u procesu odlučivanja može da se unapredi sistem poslovne inteligencije.

✓ *Organizaciono odlučivanje.*

Ovaj oblik odlučivanja karakteriše visok nivo nestrukturiranosti problema koji prate organizaciju. Iako i ovaj vid odlučivanja prate brojni istraživački eksperimenti, još uvek se nije došlo do sistematičnih znanja koja bi značajno unapredila ovu vrstu odlučivanja.

✓ *Metaorganizaciono odlučivanje.*

Ujedno predstavlja i najviši nivo primene niza sistematizovanog znanja u oblasti odlučivanja. Praktično se evidentira na nivou jedne zemlje (države), sopstvenog nacionalnog interesa, opredmećenog preko socijalnog blagostanja, kulture, dohotka, boljitka života, itd.

3.2. Grupno odlučivanje

Istaknuto je da je predmet disertacije upoznavanje sa sistemskim metodološkim pristupom grupnog odlučivanja u procesu unapređenja sistema poslovne inteligencije, što je razlog da se u nastavku ove celine stavlja akcenat upravo na fenomen grupnog odlučivanja.

Podrška odlučivanju individualnom donosiocu odluke je itekako važna aktivnost za uspešno upravljanje svakim preduzećem. Kako rad i funkcionisanje preduzeća postaju sve složeniji, a samim tim i odgovornost za donetu odluku sve je veća, postoji potreba za uvažavanjem različitih ideja i mišljenja u cilju donošenja odluka u grupi.

Takođe, stalne promene u neposrednom okruženju preduzeća, rezultuju da njegovo funkcionisanje i opstanak na tržištu postanu sve složeniji. Samim tim i odgovornost za donetu odluku postaje veća, što rezultuje zahtevom da se ista donosi od strane specijalizovanog tima za odlučivanje. Sve sa ciljem posmatranja više alternativnih pravaca i ideja, kao i međusobnog usaglašavanja i uvažavanja različitih mišljenja, radi donošenja najprihvatljivije odluke u grupi.

Sa druge strane, evidentna je potreba za poboljšanjem kvaliteta odluka donetih u grupi. Ista predstavlja stalni predmet interesovanja istraživača sa velikim praktičnim i teoretskim značajem.

U tom kontekstu, može se konstatovati da su do danas iskristalisana sledeća četiri pristupa GO:

- ✓ racionalni pristup,
- ✓ politički pristup,
- ✓ proceduralni pristup, i
- ✓ pristup „korpe" za otpatke.

Racionalni pristup u potpunosti podržava ekonomski aspekt odlučivanja. Podrazumeva rešavanje problema za koje se prethodno u potpunosti definiše cilj, sa jasno definisanim alternativama i njihovim mogućim posledicama. Pri tome se posebna pažnja poklanja praćenju svake alternative (prednostima i nedostacima) u odnosu na ostvarenje cilja.

Politički pristup razmatra unapred pripremljene predloge od strane učesnika sesije u cilju rešavanja problema. Teži se od strane svakog učesnika sesije da se postigne nivo referentne tačke, tako da svaki od njih u izvesnoj meri odstupa od početnog stava. Osnovni zahtev pri tome jeste minimizacija nesporazuma i konflikata.

Proceduralni pristup u potpunosti podržava unapred definisane procedure dolaska do grupne odluke. Sve aktivnosti GO su strukturirane. Tako je donošenje grupne odluke kod ovog pristupa zasnovano na preciznom algoritmu odlučivanja.

Pristup „*korpe*“ za *otpatke*, predstavlja oblik GO gde nisu u potpunosti definisane aktivnosti samog procesa GO. Učesnici sesije se kolebaju između protivrečnih stavova i ideja za izbor u datom vremenskom trenutku. Koristan je proces u uslovima kada je neophodno doneti odluku što pre, i pored toga što je ta procedura ponekad u potpunosti nestrukturirana. Bazira se pre svega na smanjenoj koheziji među članovima tima i aktiviranju „moždane oluje“.

Uz pretpostavku analitičkog pristupa rešavanja problema višekriterijumskog grupnog odlučivanja, moguće je uvrežiti matematički formalizam na sledeći način:

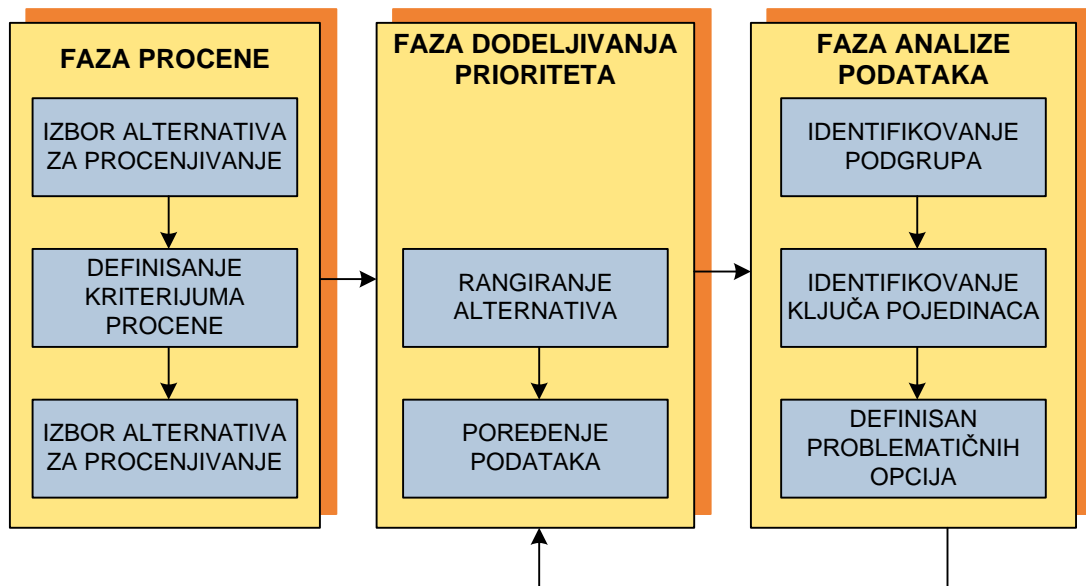
Treba pretpostaviti postojanje prostora odluka (alternativa, akcija), sa uvedenim skupom promenljivih A_x i prostora dobijanja mogućih rezultata R_q . Ukoliko su oba prostora definisana, sa dozvoljenom funkcijom preslikavanja $f : A_x \rightarrow R_q$, ostvaruje se set rešenja $q_0 = f(x_0) \in C_q$, gde je $x_0 \in A_x$, skup prihvatljivih odluka, prema [129].

Za pomenutu vrstu problema skoro da je nezamisliv adekvatan i opravdan izbor najprihvatljivije alternative, bez upotrebe adekvatne metodologije za rešavanje problema grupnog odlučivanja. S tim u vezi neophodno je dati odgovore na sledeća tri veoma važna pitanja:

- ✓ Šta je proces grupnog odlučivanja?
- ✓ Šta je grupa?
- ✓ Šta je grupna odluka?

Detaljan skup odgovora za pomenute koncepte, videti u [29].

Sasvim je sigurno da je proces donošenja odluka u grupi izuzetno složen postupak. Po jednom istraživačkom konceptu, videti u [129], grubo se mogu identifikovati faze procesa odlučivanja, prikazane na sledećoj slici.



Slika 3.1. Konceptualni okvir faza grupnog odlučivanja

Dakle, konceptualni okvir faza grupnog odlučivanja, sastoji se iz sledeće tri celine (faze):

- ✓ faza procene,
- ✓ faza dodeljivanja prioriteta, i
- ✓ faza analize podataka i zaključivanja.

Faza procene obuhvata tri osnovne aktivnosti: definisanje alternativa, definisanje kriterijuma i definisanje praga saglasnosti između učesnika sesije. Druga faza, dodeljivanje prioriteta, obuhvata procedure za određivanje redosleda važnosti alternativa uz mogućnost poređenja podataka. I na kraju, treća faza analize podataka ima za cilj da na osnovu već prikupljenih podataka identifikuje podgrupe i eventualne problematične aktivnosti, kao i da utvrdi nivo indikatora neusaglašenosti u grupi. Svaka od navedenih faza procesa grupnog donošenja odluke, može se dobro uočiti prilikom procedure primene višekriterijumskog modela grupnog odlučivanja.

Prema [23], postoji šest osnovnih načina na koji članovi grupe mogu doneti odluku:

- 1) Odluke uz nedovoljan odziv - odluke donete uz nepotpuno kritičko razmatranje i procenu.
- 2) Odluke pod uticajem autoriteta - veoma efikasna odluka u pogledu korišćenja vremena, ali da li će biti dobra zavisi od kompetentnosti i informisanosti člana tima koji ima dominantnu ulogu pri njenom donošenju, kao i od toga u kojoj meri je prihvataju ostali članovi grupe.
- 3) Odluke pod uticajem manjine - jedan, dvoje ili troje ljudi su u mogućnosti da dominiraju grupom i usmere je ka donošenju odluke koja im odgovara.

- 4) Odluke pod uticajem većine - odluka se donosi glasanjem kako bi se utvrdili stavovi većine. Ovakva odluka ima tendenciju da stvori koalicije “pobednika” i “gubitnika”, u takvoj situaciji “gubitnici” se mogu osećati izostavljenim što se dalje može odraziti na kvalitet sprovođenja odluke.
- 5) Jednoglasne odluke - predstavlja gotovo savršen metod grupnog odlučivanja, ali ga je izuzetno teško primeniti u praksi.
- 6) Odluke donete konsenzusom - uspeh u ostvarivanju konsenzusa zahteva disciplinu i podršku svih članova grupe, njihovo uključivanje u sve faze diskusije, što za rezultat ima razumevanje i prihvatanje donete odluke, kao i osećaj odgovornosti za njeno sprovođenje.

Jedan od osnovnih razloga što se grupe ponekad kreću ka donošenju odluka pod uticajem autoriteta ili manjine je to što je jednoglasnost ili konsenzus teško ostvariti. S obzirom na značaj konsenzusa kao oblika odlučivanja u poslovnim organizacijama u nastavku ovog poglavlja biće dat kratak osvrt na neke njegove osnovne karakteristike.

Konsenzus – model kome se teži

Konsenzus ne podrazumeva da svi članovi grupe uvek smatraju da je doneta najbolja moguća odluka ili da doneta odluka uopšte može rešiti razmatrani problem. Ono što konsenzus podrazumeva je da u procesu odlučivanja nijedan član tima nema osećaj da je njegovo mišljenje zanemareno ili nepravilno shvaćeno.

Donošenje odluka konsenzusom zahteva više vremena, ali zato podstiče kreativnost i među članovima tima stvara osećaj privrženosti donetoj odluci. Ovakav oblik odlučivanja može dati dobre rezultate ukoliko:

- ✓ grupa ima zajedničke vrednosti,
- ✓ članovi grupe poseduju izvesno iskustvo i veštine vezane za grupne procese i rešavanje konfliktnih situacija,
- ✓ postoji posvećenost članova grupi i osećaj odgovornosti za rad u njoj,
- ✓ svi članovi grupe mogu da izdvoje dovoljno vremena da učestvuju u procesu odlučivanja.

Donošenje odluke konsenzusom podrazumeva najpre iznošenje jednog predloga kao mogućeg rešenja razmatranog problema. Ovaj predlog se u toku diskusije prilagođava i menja u skladu sa iznetim mišljenjima članova grupe ili se od njega potpuno odustaje ukoliko se pokaže da on ne vodi rešavanju problema. Oni članovi grupe koji se zalažu za odbacivanje razmatranog predloga imaju obavezu da predlože novo, alternativno rešenje.

Osnovna pravila na kojima se konsenzus zasniva su:

- ✓ svi članovi tima za odlučivanje imaju priliku da slobodno izraze svoje mišljenje, i
- ✓ svako je obavezan da sasluša i uvaži mišljenje ostalih članova.

Kada se u postupku donošenja odluke formira klima opšteg razumevanja razmatranog predloga i nema zahteva za nekim dodatnim izmenama istog, moderator se obraća svim članovima grupe i poziva ih da ukoliko imaju bilo kakve zamerke koje se odnose na predloženo rešenje upoznaju grupu sa njima. Taj poziv je poslednja prilika da se spreči usvajanje konkretnog rešenja jer ukoliko nema primedbi sledi poziv na konsenzus. Trenutak tišine u sali nakon ovog poziva znači da je konsenzus postignut, odnosno odluka doneta. Donetu odluku moderator sastanka još jednom saopštava članovima grupe i time se postupak odlučivanja okončava.

Ukoliko se neki od članova tima za odlučivanje ne slaže sa donetom odlukom ili sa predlogom koji je pred neposrednim usvajanjem, on to može izraziti na različite načine, videti [37]:

- ✓ ne podržavanje ("Ne vidim razlog za usvajanje predloženog rešenja, ali ću ga prihvatiti");
- ✓ rezervisanost ("Mislim da bi ovo mogla biti greška, ali donošenje odluke neću sprečavati");
- ✓ stajanje po strani ("Ja, lično to ne mogu uraditi, ali druge neću odvracati ili sprečavati");
- ✓ blokiranje ("Ja ovu odluku nemogu podržati, niti ću dozvoliti grupi da je prihvati" - ukoliko je finalna odluka u suprotnosti sa fundamentalnim stavovima ili vrednostima nekog člana grupe on može blokirati konsenzus);
- ✓ povlačenje iz grupe.

Jedan od bitnih činilaca uspešnosti opisanog pristupa je izbor moderatora, koji ima ulogu rukovodioca sastanka. Taj izbor je važan iz razloga što moderator: pomaže grupi u definisanju problema odlučivanja, pomaže članovima da prođu kroz sve faze procesa odlučivanja, usmerava diskusiju na razmatrani problem, pruža priliku svim članovima tima da učestvuju u diskusiji. Njegova uloga je da upravlja tokom sastanka a ne njegovim sadržajem. On nikada ne donosi odluku umesto grupe. Sa druge strane treba odgovoriti na pitanje, *Šta je grupa?*

Grupa se obično definiše kao skup ljudi koji su u učestaloj interakciji jedni sa drugima tokom određenog vremenskog perioda i koji uviđaju međusobnu zavisnost u pogledu ostvarivanja jednog ili više zajedničkih ciljeva.

Iako se u svakoj organizaciji mogu identifikovati različite grupe, predmet disertacije će biti isključivo formalne grupe ili timovi za odlučivanje. Neke od osnovnih karakteristika grupe za odlučivanje su, videti [23]:

- ✓ čini je mali broj ljudi sa komplementarnim veštinama angažovanim u zajedničke svrhe,
- ✓ specifičan način ostvarivanja cilja,
- ✓ zajednički pristup problemu, i
- ✓ međusobna podela odgovornosti.

Očigledno je da su grupe prirodan i veoma koristan fenomen u modernim organizacijama. Njihove osnovne prednosti su u tome što, detaljno u [23]:

- ✓ podstiču inovativnost i kreativnost među članovima,
- ✓ omogućuju kombinovanje različitih stilova rešavanja problema,
- ✓ objedinjuju različita znanja, ideje i informacije,
- ✓ veće razumevanje i posvećenost rešavanju evidentiranog problema,
- ✓ veća zainteresovanost članova za sprovođenje donetih odluka,
- ✓ i kao najvažnije, odluke donete u okviru grupe su kvalitetnije od odluka koje donose pojedinci.

I pored evidentnih prednosti grupe, uočeni su i izvesni njeni nedostaci:

- ✓ dolazi do povećanja upotrebe sredstava organizacije,
- ✓ moguća dominacija pojedinaca, ako proces nije dobro definisan,
- ✓ moguća difuzija odgovornosti,
- ✓ Ringelmann-ov efekat - uočena tendencija pojedinih članova grupe da prilikom grupnog odlučivanja ulažu manje napora nego pri individualnom radu, smatrajući da će ostali članovi obaviti sav potreban posao.

Na osnovu prethodno identifikovanih prednosti i nedostataka grupe, možemo uvesti pojam efektivne grupe kao donosioca odluke. Efektivnom grupom se može smatrati grupa sa sledećim osobinama, videti [23]:

- ✓ jasna predstava o tome u čemu je problem,
- ✓ jak osećaj podeljene odgovornosti za rezultate i postupke grupe,
- ✓ članstvo bazirano na stručnosti, a ne na ličnosti,
- ✓ poštovanje želje za diskusijom svakog člana grupe,
- ✓ prevazilaženje razlika i mogućih sukoba među članovima,
- ✓ prisustvo međusobnog poverenja i otvorenosti u grupi.

Prema tome, proces grupnog odlučivanja može se definisati kao situacija u kojoj, videti [23]:

- (i) dve ili više osoba, svaka od njih karakterisana svojim posebnom percepcijom, stavovima, motivacijom i ličnošću,
- (ii) koje prepoznaju postojanje uobičajenog problema, i
- (iii) pokušavaju da dođu do jedinstvene grupne odluke.

Šta više, u grupi može da postoji istovremena interakcija (tj. paralelni međuzavisni način) ili da članovi grupe donose nezavisno individualne odluke, a onda ih kolektivno sučeljavaju i raspravljaju o rezultatima (tj. sekvencijalni međuzavisni način).

Mogu se razmatrati različiti tipovi grupnog odlučivanja: jedan donosilac odluke koji radi u grupnom okruženju, kooperativno odlučivanje, i nekooperativno odlučivanje. U nekooperativnom odlučivanju, donosioci odluka igraju uloge suprotstavljenih ili diskutanata. Konflikt i takmičenje su uobičajene forme nekooperativnog odlučivanja. Dok konflikt predstavlja situaciju u kojoj diskutanti teže da nadmaše protivnike da bi ostvarili svoje interese, takmičenje karakteriše činjenicom da je svaki takmičar kandidat za akciju, i on teži da nadmaši drugu.

U kooperativnom okruženju donosioci odluka pokušavaju da donesu zajedničku odluku na profesionalan način i uz poverenje, i da podele odgovornost. Konsenzus, pregovaranje, čak i obraćanje trećoj strani kako bi se razrešile razlike, su primeri ovog tipa grupnog odlučivanja.

Takođe, literatura koja se bavi donošenjem odluka opisuje dva tipa situacija pri odlučivanju u kome učestvuje više članova tima: paralelni međuzavisni i sekvencijalni međuzavisni. U situaciji paralelnog odlučivanja, donosioci odluka se okupljaju da bi sačinili manje-više homogenu grupu i pokušali da reše zajednički problem simultano. Sa druge strane u sekvencijalnoj međuzavisnoj situaciji, članovi grupe mogu da rešavaju zajednički problem u različitim periodima vremena, posmatrajući odluke iz različitih uglova.

Druga klasifikacija pristupa grupnom rešavanju problema opisana u literaturi, je razlikovanje između pristupa koji su orjentisani na sadržaj i onih koji su orjentisani na proces.

Prvi pristup se fokusira na sadržaj problema, pokušavajući da nađe optimalno ili zadovoljavajuće rešenje, uzimajući u obzir određena društvena ili grupna ograničenja ili ciljeve. Nasuprot tome, drugi pristup se temelji na zapažanju da grupa prolazi kroz određene faze u procesu grupnog odlučivanja, i na verovanju da se može pronaći način za efikasno suočavanje sa ovim fazama.

Kada kolektivna odluka ne uspe, neophodno je da učesnici u grupnom rešavanju problema počnu da pregovaraju dok se ne dođe do konsenzusa. Tako pregovaranje uključuje diskusiju u okviru određenog kriterijuma ili pitanja, sve zbog traganja za konsenzusom.

Danas, većina poslovnih sistema je prihvatila pojam grupa. Ipak, pregledom će se pronaći da su nekoliko organizacija i nekoliko pojedinaca u njima naročito zadovoljni sa načinom na koji njihove organizacije rade. Timovi mogu biti neophodna komponenta organizacionog uspeha, ali, naravno njihovo prisustvo ne garantuje uspeh.

Kada grupa dobro funkcioniše (bilo da je u pitanju radna grupa, sportski tim, prijateljska grupa, hor ili orkestar, religiozna grupa, dobrovoljna grupa, itd.), dinamika grupe i osećaj pripadanja i prihvaćenosti može da iznese ono što je “najbolje” u svim članovima tima. Grupe mogu da unaprede rešavanje problema i kreativnost, da stvore razumevanje, prihvaćenost, podršku i privrženost. Pored toga, grupe mogu da povećaju moral, obezbede ispust za udruženje, povećaju samopoštovanje, pomognu stvaranju konsenzusa i sigurnosti.

Nakon prikaza procesa savremenog odlučivanja, od individualnog preko grupnog u nastavku sledi opis važnosti podatka u sistemu poslovne inteligencije. Jer već je u samom uvodnom poglavlju ukazano na osnovu poslovne inteligencije a to je sled *podatak, informacija i znanje*.

4.0. VAŽNOST PODATKA U POSLOVNOJ INTELIGENCIJI

Već je u samom uvodu doktorske disertacije istaknuto da je osnova poslovne inteligencije sled: *podatak* → *informacija* → *znanje*. To je dovoljno jak razlog, da se na početku ovog dela poglavlja koje se upravo odnosi na PI, posveti sveobuhvatna pažnja podatku. To podrazumeva kompletan višekriterijumski aspekt od izvora podataka, pripreme podataka, kvaliteta podataka, skladištenja podataka i napokon samog tumačenja podataka. Sledi poglavlje posvećeno njegovoj visosti *podatku*.

4.1. Podatak kao osnova poslovne inteligencije

Cilj ovog dela disertacije jeste da ukaže na važnost i ulogu podatka u poslovanju svakog poslovnog sistema. Već je prethodno istaknuto da od podatka i njegovog tumačenja zavisi informacija i znanje. Znanje člana tima ili grupe, na kome se bazira donošenje ispravne poslovne odluke.

Ako su podaci pak netačni, nekonzistentni ili nepouzdana, isti predstavljaju opasnost za svaki poslovni sistem. Imajući u vidu da su količine podataka sa kojima kompanija raspolaže konstantno u usponu, jasno je da je kvalitet podataka postao ključni problem. Mnogi faktori utiču na sposobnost da se precizno konsoliduju podaci i pruže pouzdane informacije: sve veći broj sistema i standarda, novih nezavisnih podataka koji se ne mogu lako integrisati, kao dupliranih podataka i slično. U mnogim slučajevima, standardi nisu dostupni ili se ne poštuju dosledno u celom poslovnom sistemu.

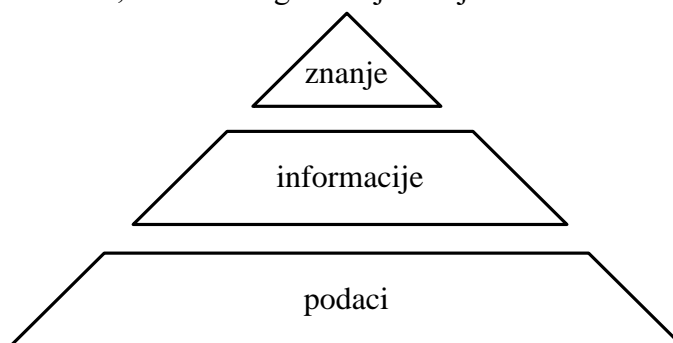
Razvoj interneta i informacionih tehnologija otvara korisnicima ogromne mogućnosti upravljanja informacijama, na način koji osigurava najveću poslovnu efikasnost. Nasuprot stvaranju nepreglednih skladišta podataka, ovo je moguće ostvariti sistemskim pristupom upotrebe najboljih znanja i iskustava za pojedine oblasti poslovanja, tako što istraživači ta znanja skrivena u podacima prepoznaju i izražavaju na određen način. Sticanje i korišćenje znanja baziranog na činjenicama u svrhu unapređivanja kako strateškog tako i taktičkog poslovanja na tržištu, može se jedino realizovati ako se poseduju ispravni poslovni podaci.

U stvarnosti, protumačeni poslovni podatak (informacija), omogućava način poslovnog razmišljanja koji obezbeđuje, da se poslovne odluke na svim nivoima odlučivanja donose tako da budu utemeljene na relevantnim i ažurnim poslovnim informacijama, a ne

utemeljene na intuiciji i subjektivnom utisku. S druge strane, u informatičkom smislu, to je složeni informacijski sistem koji automatizovanim procedurama prikuplja podatke iz raznih izvora, obrađuje ih, transformiše i integriše, i na taj način omogućava donosiocu odluke (menadžeru) kvalitetnu projekciju poslovne odluke.

Ključno pitanje savremenog poslovanja je, šta je najvažnija pretpostavka opstanka jednog poslovnog sistema na savremenim turbulentnim tržištima? Odgovor je – informacija, ili, preciznije, informacija koja omogućava preduzimanje odgovarajuće akcije, odnosno, donošenje adekvatne odluke. Navedena konstatacija podjednako važi za sve oblasti poslovanja.

Pojmovi podatak, informacija i znanje nalaze se u međusobnom odnosu logičke nadređenosti i podređenosti, odnosno logičke hijerarhije.



Slika 4.1. Logička hijerarhija podataka, informacija i znanja

Na vrhu piramide nalazi se znanje. Ono se može protumačiti kao mogućnost ostvarivanja ili posedovanja dubokog uvida u određeni fenomen. Tako se može reći da kompanija poseduje znanje o svojim klijentima kada može u potpunosti da shvati i razume njihove osobine, ponašanje i preferencije. Znanje je moguće posedovati samo ako se poseduju informacije koje odražavaju neku celinu. Informacija je temelj na kome se gradi znanje. Istovremeno, informacija se gradi na podacima. Ako se ne raspolože dobrim podacima, neće biti ni valjanih informacija, pa će svako znanje izvedeno iz njih čak biti opasno po poslovni sistem.

Važnost podataka i njihovog pravilnog interpretiranja igra ključnu ulogu u čitavom tom procesu. Upravo zato, bavljenje načinom obrade i pretraživanja podataka predstavlja svojevrsan izazov. Dinamički pristup se nameće kao najprihvatljivije rešenje, kada je priroda posmatranih podataka kompleksna. Ono što se želi postići je što viši nivo tačnosti u predviđanjima, i to kroz afirmisanje primene savremenih algoritama u radu.

Sam termin podatak potiče od latinske reči *datum*, *data* koja označava ono što je dato, podatak. Za pojam podatka postoji mnoštvo definicija, koje se za objašnjenje uglavnom

služe terminima „informacija“, „činjenica“ ili „pokazatelj“. Neke od definicija iz referenci prikazanih u literaturi su sledeće, videti [135]:

- ✓ Činjenice su događaji ili stanja, a podaci su njihova simbolička prezentacija.
- ✓ Informacije su podaci koji su obrađeni u oblik koji je smislen za primaoca i koji imaju stvarnu ili opaženu vrednost u sadašnjim ili predstojećim akcijama i odlukama.
- ✓ Podaci su deskripcije fenomena, tj. diskretne zabeležene činjenice o fenomenima, od kojih se dobijaju informacije o svetu.
- ✓ Podaci su vrednosti pokazatelja koje odgovaraju prihvaćenoj skali, itd.

Pomenute različite interpretacije šta je podatak, dovoljno govori za sebe da se o ovoj oblasti odavno ukazivalo i vodilo računa. Shodno tome, sledeći opisi mogu se uzeti kao opšte prihvaćeni:

Činjenica označava:

- ✓ relativno istinit iskaz o nekom stvarnom stanju stvari,
- ✓ samo to stanje stvari.

Podatak označava:

- ✓ neki iskaz,
- ✓ neki znak prirodnog ili veštačkog jezika o stanju stvari.

Informacija označava:

- ✓ vrstu podataka (ili obratno: podatak vrstu informacije),
- ✓ neke vrste podataka, a neki podaci neke vrste informacija (logičko preklapanje),
- ✓ upotrebljivi, obrađeni oblik sirove građe podataka.

Pokazatelj je svojstvo, promenljive koja može da uzima različite vrednosti iz nekog određenog skupa vrednosti.

Odnos podatka i informacije je jako relativan, u smislu da ono što za jednu osobu može da bude informacija za drugu može biti tek sirovi podatak. Malo detaljnija definicija je da je podatak iskaz ili neki znak (verbalni, brojčani, ikonički ili drugačiji). nekog jezika (prirodnog ili veštačkog), o nekom stanju stvari, kojim se:

- ✓ tvrdi ili poriče postojanje ili prisutnost nekog objekta, događaja, pojave,
- ✓ pripisuje ili odriče neko svojstvo objektu ili neka relacija objekta sa drugim objektima,
- ✓ tvrde ili poriču ili upoređuju vrednosti tih svojstava.

Relevantni kriterijumi za razvrstavanje podataka, prema [139] su:

- ✓ izvor (npr. primarni i sekundarni),
- ✓ priroda (npr. kvantitativni podaci su oni kod kojih se razlikuju kategorije u koje su svrstani objekti, a kod kvalitativnih podataka se razlikuju i kategorije i relacije uređenosti kategorija),
- ✓ forma u kojoj su izraženi (npr. verbalni i neverbalni), itd.

Najznačajnija svojstva podataka su: *pouzdanost, verodostojnost, tačnost i značajnost*. Podaci ne nastaju ni iz čega, nego su oni plod neke namere, oni se kreiraju. I jos jedna bitna karakteristika je da ne mogu da budu sasvim nepristrasni, neutralni i nezavisni od vremenskog, kulturnog, teorijskog, metodološkog konteksta, itd., prema, [68].

Nakon identifikacije da se “rodio” podatak, neophodno je uraditi obradu istoga. Svrha obrade i analize podataka je crpljenje informacija neophodnih za sticanje odgovora na pitanje zbog kojeg se istraživanje vrši, odnosno odluka donosi. Kao odgovor na zahteve koje postavlja, razvijen je veći broj statističkih paketa za analizu podataka. Neki od njih su SPSS (*Statistical Packages for the Social Sciences*), SAS (*Statistical Analysis System*), CSS (*Complite Statistical System*), BDMP (*BioMeDical Computer Programs*), itd. U analizi podataka je pored izbora jednog od ovakvih paketa potrebno izabrati i odgovarajuće postupke statističke analize.

Značajni kriterijumi za izbor odgovarajuće procedure analize podataka su:

- ✓ broj promenljivih koje se istovremeno analiziraju,
- ✓ svrha analize (deskripcija uzroka ili statističko zaključivanje o osnovnom skupu iz kojeg je uzorak uzet i nivo merenja),
- ✓ vrsta pretpostavki o rasporedu osnovnog skupa kojem uzorak pripada, itd.

Prema broju promenljivih uključenih u statistički model, biraju se postupci univarijatne, bivarijatne ili multivarijatne analize. U grupi postupaka multivarijatne analize razlikuju se dva podskupa:

- ✓ saznavanje međusobne povezanosti promenljivih unutar jednog skupa, pri čemu nijedna promenljiva nije označena kao zavisna od ostalih, to su: faktorska analiza, klusterska analiza i multidimenzionalno skaliranje,
- ✓ saznavanje međusobne povezanosti promenljivih unutar jednog skupa, pri čemu je najmanje jedna iz skupa promenljivih određena kao zavisna od drugih; tu opet, postoje dve varijante – slučaj kada postoji samo jedna zavisna promenljiva (tada se mogu primeniti regresija, analiza varijanse i kovarijanse, diskriminaciona analiza

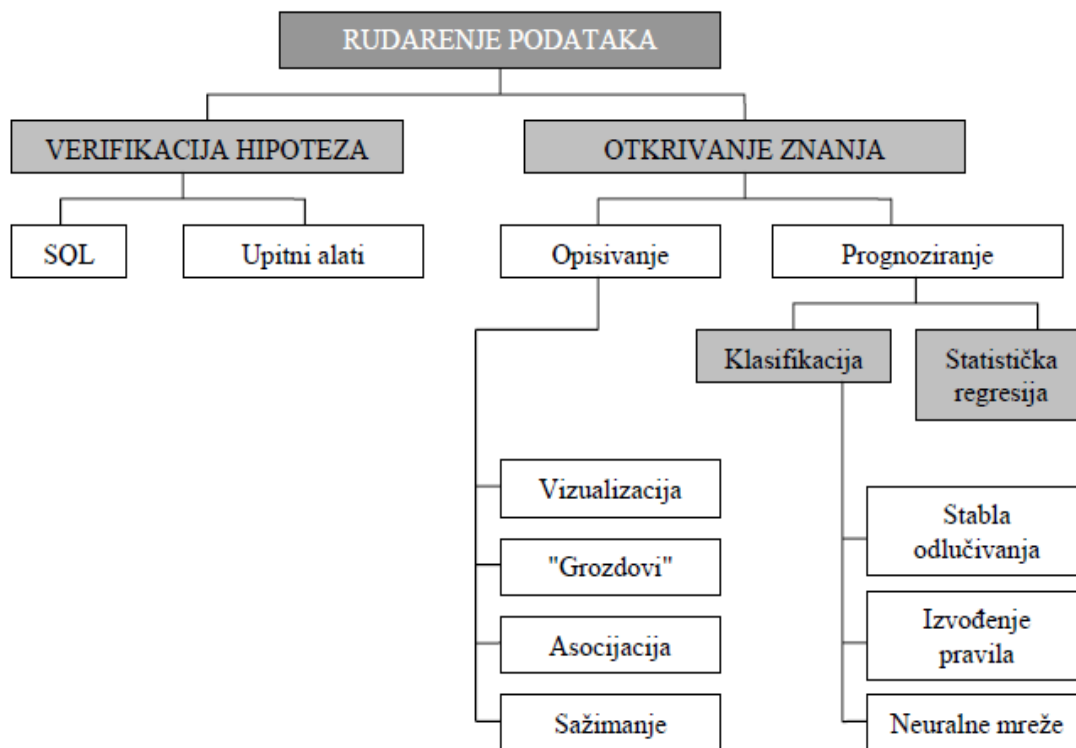
itd.) i slučaj kada postoji više od jedne zavisne promenljive (obično se koriste multivarijatna analiza varijanse ili kanonička korelacija).

U narednom poglavlju posebna pažnja biće posvećena pojmu pronalaženje zakonitosti u podacima, dok će u okviru ovog dela disertacije biti date neke osnovne smernice za pretraživanje podataka, pre svega sa aspekta njegove veličine „podatak“. Pojam pretraživanje ili rudarenje podataka može se definisati kao pronalaženje zakonitosti među podacima. Podaci kao predmet proučavanja mogu biti sadržani u bazama podataka, mogu biti razni tekstualni podaci, nestrukturirani podaci ili podaci organizovani u vremenske serije. Primenom metoda i alata za rudarenje podataka mogu se otkriti često iznenađujući obrasci ponašanja o pojavama koje se posmatraju ili neočekivani odnosi među pojavama.

Jedna od definicija rudarenja podataka glasi: rudarenje podataka je sistemski, interaktivan i iterativan (ponavljajući) proces izvođenja i prikazivanja korisnog, implicitnog i inovativnog znanja iz podataka [35]. Važno je uz definiciju spomenuti da uspešnost primene metoda i alata ove namene zavisi prvenstveno od stručnosti i poslovne kompetencije onih koji tumače dobijene rezultate, pre svega donosioci odluke ili menadžeri. Upravo donosioci odluke, svojim znanjem i iskustvom mogu biti sposobni da neki naizgled besmislen podatak interpretiraju na poslovno korektan i smislen način i pretvore ga u vrednu informaciju.

Dva su osnovna tipa pretraživanja podataka:

- ✓ verifikacija hipoteze – cilj je proveriti da li je neka ideja ili pretpostavka o važnosti odnosa među određenim podacima osnovana ili ne, i
- ✓ otkrivanje novih znanja – među nekim pojavama mogu postojati neki još nepoznati, a statistički važni odnosi koje čovek ni iskustvom, ni svojim intelektualnim sposobnostima ne može otkrije.



Slika 4.2. Taksonomija rudarenja podataka, [35]

Rudarenje podataka je izrazito multidisciplinarno područje. Ono obuhvata područja baza podataka, ekspertnih sistema, teorije informacija, statistike, matematike, logike i čitav niz pridruženih područja. Područja u kojima se rudarenje podataka može uspješno primjenjivati su raznolika, kao na primer, poslovanje preduzeća, ekonomija, mehanika, medicina, genetika, itd. Uopšte gledano, rudarenje podataka primenjivo je u svim onim područjima gde se raspoložuje velikom količinom podataka, na osnovu kojih je potrebno otkriti određene pravilnosti, veze i zakonitosti.

Obzirom na to da postoji čitav niz faktora koji mogu uticati na ishod nekog događaja, zadatak pretraživanja podataka je da otkrije najznačajnije među njima i njihove karakteristike. Bez obzira na područje primene, dobro iskorišćene metode pretraživanja podataka mogu otkriti zakonitosti iz velike količine podataka, pri čemu je naglasak na samim podacima, a ne na području sprovođenja analiza, što ustvari predstavlja jednu od najvećih prednosti primene ovih metoda.

Izraz pretraživanje/rudarenje podataka odnosi se ili na čitav proces otkrivanja znanja ili uže, na specifičnu fazu "mašinskog učenja". Razni dobavljači DM tehnologija daju razne interpretacije termina. SAS Institut, vodeći u području tehnologija za rudarenje podataka (posebno statističkih), određuje ga kao skup naprednih metoda za istraživanje i modeliranje

veza u velikim količinama podataka. Među osnovnim tehnologijama za rudarenje podataka su statistika, sistemi bazirani na pravilima, neuronske mreže, mašinsko učenje, veštačka inteligencija, upravljanje bazama podataka, itd. Ipak, svi se slažu da je analitička statistika, jezgro tih novih procesa za otkrivanje znanja. Iz statističke perspektive, otkrivanje zakonitosti u podacima se može opisati kao računarski automatizovana istraživačka analiza podataka iz obično velikih i složenih baza podataka s različitim platformi, lokacija, operacionih sistema i softvera.

U savremenom okruženju, poslovni procesi i naučni instrumenti mogu lako generisati količine podataka koje se mere u terabajtima. Jako brzo se širi jaz između mogućnosti prikupljanja podataka i sposobnosti analize podataka kako u biznisu, tako i u nauci. Rudarenje podataka najveću primenu nalazi tamo gde je stalno prisutna pretnja “poplave” informacijama. Slede opisi nekih metoda i alata za pronalaženje zakonitosti u podacima.

Metode i alati korišćeni u raznim fazama razvoja koncepta pretraživanja podataka uglavnom su bili izvedeni iz metoda statističke analize. Nova znanja su se otkrivala uglavnom analizom:

- ✓ raspodele vrednosti podataka, ili
- ✓ odnosa među podacima.

Analiza raspodele vrednosti podataka je jedna od osnovnih statističkih metoda. Poznata su dva pravca te analize, u zavisnosti od toga da li se koriste opisni (deskriptivni) ili broječni (numerički) podaci. Jednostavna analiza raspodele na osnovu opisnih podataka deliće podatke prema npr. polu, geografskom području, opredeljenju, itd. Analiza opisnih podataka najčešće kao rezultat ima raspodelu zasnovanu na apsolutnoj vrednosti podataka. Sa druge strane, analiza raspodele zasnovane na numeričkim podacima najčešće se naziva jednostavnom statistikom, jer dobijeni rezultati predstavljaju neke dobro poznate statističke pokazatelje (aritmetičke sredine, standardne devijacije, varijanse, itd).

Veliki broj tehnika za rudarenje podataka koje se preporučuju u literaturi, može se svrstati u jednu od sledećih grupa, videti [35]:

- ✓ klasične metode za multivarijantnu analizu podataka i statistiku (klaster analiza, diskriminativna analiza, linearna i nelinearna regresija, metoda glavnih komponenta, analiza vremenskih serija, itd.),
- ✓ metode grananja (*tree-based methods*) klasifikacijska i regresiona stable (CART), hi-kvadrat automatska detekcija interakcija (CHAID),
- ✓ neuronske mreže.

S obzirom na funkciju, alati za rudarenje podataka mogu se podeliti na sledeći način, detaljno u [35]:

- ✓ klasifikacija: klasifikovanje podataka (entiteta) u jednu od nekoliko prethodno definisanih klasa,
- ✓ regresija: uspostavljanje relacija uz pomoć prediktorskih varijabli,
- ✓ klasterovanje: klasifikovanje podataka (entiteta) u jednu od nekoliko klasa (klastera), pri čemu se klase moraju odrediti iz podataka - za razliku od klasifikacije, kod koje su klase unapred definisane,
- ✓ sažimanje, uključujući vizualizaciju i istraživačku analizu podataka,
- ✓ modelovanje zavisnosti,
- ✓ asocijacije,
- ✓ sekvencijalna analiza, itd.

Da li će proces pretraživanja podataka biti uspešan ili ne zavisi od više činilaca, kao što su prethodno znanje o području, zatim raspoloživost i odabir alata za pretraživanje podataka, kvalitet podataka, dobra interakcija između članova projektnog tima, kvalitet skladišta podataka, jasnoća misije i cilja pretraživanja podataka, način prikaza rezultata, itd. Svakako, kvalitetno istraživanje podataka metodama rudarenja podrazumeva i širok raspon znanja analitičara i njihovo poznavanje pojedinosti svake od metoda koje se primenjuju, iz razloga što je neophodno pripremiti podatke za svaku pojedinačnu analizu, a da bi se priprema pravilno uradila poznavanje metode je od ključne važnosti.

Kao što je već pomenuto, kvalitet podataka je presudni faktor od kog zavise rezultati pretraživanja podataka i statističkih analiza podataka. Podrobnije o kvalitetu podataka posvećeno je sledeće poglavlje ove disertacije.

Pored samog izvora, jako značajnu ulogu imaju i postupci čišćenja i pretprocesiranja podataka. Praksa pokazuje da analitičari u proseku najviše vremena provode upravo u čišćenju i pretprocesiranju podataka, i to 80% vremena se utroši na analize, a samo 20% vremena odlazi na primene metoda pretraživanja nad pripremljenim podacima. Uvođenje skladišta podataka u poslovanje donosi prednosti njihovog korišćenja kao relativno „čistih“ izvora podataka. Za analize, između ostalog i uštedu vremena, ali ipak postoje situacije kada se mora pristupiti izvornoj bazi podataka, npr. ukoliko se neki podatak ne unosi u skladište podataka ili se ne pojavljuje na odgovarajućem nivou granulacije. Slično je i sa procesiranjem podataka koji su rezultat raznih anketa i ispitivanja tržišta koji se svakako obrađuju sirovi. U ovakvim slučajevima, korišćenje izvornih podataka je neizbežno.

Podaci u izvornom obliku su često nekompletni, atributi imaju nedostajuće vrednosti, ili se čak javlja nedostatak atributa. Takođe, može se pojaviti nekonzistentnost u okviru samih

podataka, kao na primer nedoslednost u označavanju pojedinih kategorija ili grupa. Često se unutar podataka pojavljuju ekstremne vrednosti (*outliers*) koje su ponekad greške, ali često i nisu. Da bi se dijagnostikovao uzrok pojavljivanja ekstremnih vrednosti potrebno je izvršiti dodatne analize, i tek nakon njih odlučiti da li će se u dalju analizu ući sa eksternim vrednostima ili bez njih. Neretko ekstremne vrednosti u podacima ukazuju na vrlo poučnu devijaciju, i nikako ih ne bi trebalo isključiti bez prethodnih razmatranja o uzrocima pojavljivanja takvih vrednosti.

Problem nedostajućih podataka je od posebnog značaja. Detaljan opis ove problematike nedostajućih vrednosti, biće prikazan u Poglavlju 4.3., ove disertacije. Dešava se da se ponekad nekompletni slogovi jednostavno zanemare, međutim u svakoj ozbiljnijoj obradi, kada se u populaciji posmatranih podataka naiđe na nedostajuće vrednosti, tada se u procesu analize podataka koriste različite metode predviđanja nedostajućih vrednosti. U slučajevima kada postoje skladišta podataka, često su ove metode implementirane u sam ETL (*extract-transform-load*) proces, što u znatnoj meri olakšava dalju obradu. Tu se još na početku teži ka transparentnosti i uglavnom su unapred definisani atributi nad kojima će se vršiti obrada, pa su sva moguća čišćenja podataka locirana u ETL-u. Kod pretraživanja podataka, situacija je dosta komplikovanija.

Problemi pretprocesiranja posebno dolaze do izražaja kada se vrši prikupljanje podataka iz različitih izvora, gde često ne postoji jednoznačna definicija ključeva, različiti sistemi kategorizacije unutar podataka i sl. Primer je kad se uz standardne izvore podataka unutar nekog preduzeća, koristi i niz eksternih izvora od kojih se podaci dobijaju u različitim formatima. Kompleksnost obrade podataka u ovakvim slučajevima dobija na veličini, jer uvođenje novih izvora podataka može inicirati i promenu metodologija pretprocesiranja i čišćenja podataka i uvođenje pondera važnosti pojedinih atributa ili grupe atributa na ciljnu varijablu.

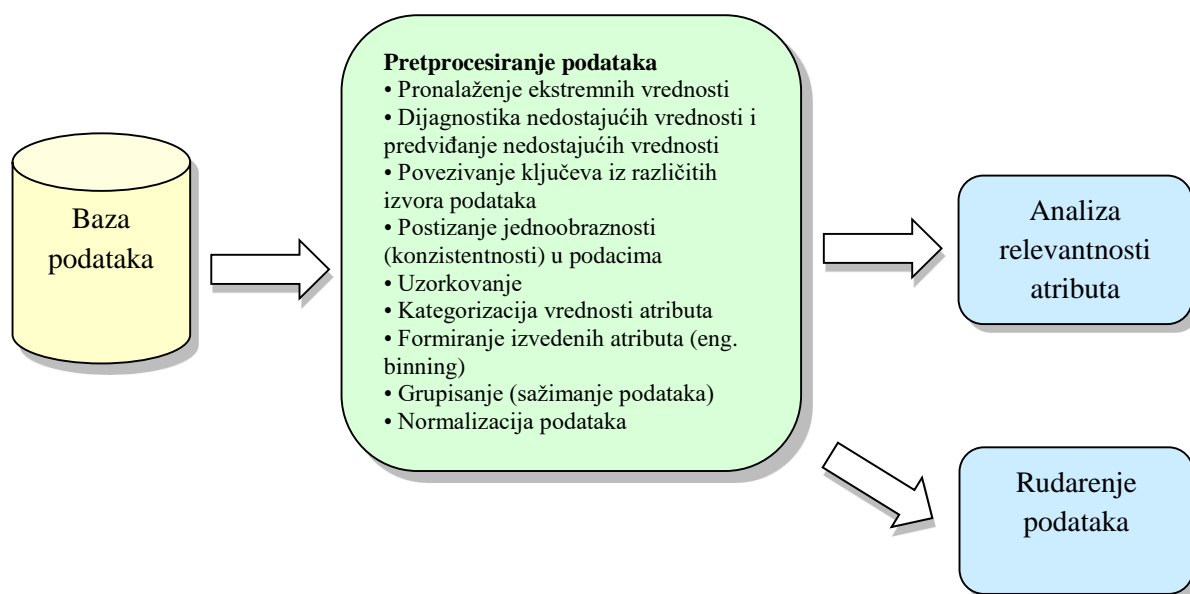
Kao jedno od rešenja koje može biti od koristi kod pretprocesiranja podataka često je zgodno iskoristiti upitni jezik SQL, koji se široko primenjuje u sistemima za upravljanje bazama podataka. Njegove prednosti posebno dolaze do izražaja kroz mogućnosti optimizacije upita koji pristupaju velikim količinama podataka, s ciljem modifikacije i kreiranja izvedenih tabela za potrebe pretprocesiranja podataka.

Postoje i programski paketi koji se koriste u pretraživanju podataka i oni imaju već unapred definisane module za pretprocesiranje podataka. Programski paketi kao što je SPSS, u sebi sadrži skriptni jezik koji pruža mogućnost brzog i efikasnog programiranja modula za pretprocesiranje podataka.

Zbog kompleksnosti područja nije redak slučaj da se pretprocesiranje vrši uz pomoć modula isprogramiranih u nekim od viših programskih jezika, pa se tek završni procesi pretprocesiranja obavljaju u okviru skriptnih jezika.

U najznačajnije metodološke postupke pretprocesiranja ubrajaju se, videti [35]:

- ✓ pronalaženje ekstremnih vrednosti,
- ✓ dijagnostika nedostajućih vrednosti i predviđanje nedostajućih vrednosti,
- ✓ povezivanje relacionih ključeva iz različitih izvora podataka,
- ✓ postizanje jednoobraznosti (konzistentnosti) u podacima,
- ✓ uzorkovanje,
- ✓ kategorizacija vrednosti atributa,
- ✓ grupisanje (sažimanje podataka),
- ✓ normalizovanje podataka, itd.



Slika 4.3. Proces rudarenja podataka

Pošto će u petom poglavlju disertacije biti detaljno opisan aspekt rudarenja podataka ili pronalaženje znanja u podacima, ovde će biti akcent na analizu relevantnih vrednosti. Nakon pretprocesiranja podataka, analitičar uvek postavlja pitanje da li je izabrao pravi skup

atributa, koji jednoznačno opisuju problematiku koju treba razrešiti, da li su vrednosti tih atributa adekvatno grupisane, i da li je opravdano vršiti analizu nad odabranim nivoom granulacije podataka.

Jedna od vrlo jednostavnih mera koja može pružiti informacije o međusobnoj zavisnosti dve varijable je koeficijent korelacije koji se računa sledećom relacijom:

$$r_{x,y} = \frac{\sum(x - \bar{x})(y - \bar{y})}{(n - 1)\sigma_x\sigma_y}$$

pri čemu je:

- $r_{x,y}$ - koeficijent korelacije,
- n - broj posmatranih slučajeva u populaciji,
- x, y - nizovi vrednosti,
- \bar{x}, \bar{y} - srednje vrednosti x i y
- σ_x, σ_y - standardne devijacije nizova vrednosti x i y

Rezultati se tumače na sledeći način:

- ✓ ako je vrednost koeficijenta korelacije veća od 0, tada su nizovi vrednosti pozitivno korelisani,
- ✓ ako je vrednost koeficijenta korelacije manja od nule, tada su vrednosti nizova negativno korelisani,
- ✓ ako je vrednost koeficijenta korelacije jednaka 0, tada ne postoji nikakva korelacijska međuzavisnost.

Ova mera se koristi kada je potrebno videti da li neka varijabla ima pozitivan, negativan ili neutralan smer korelacije u odnosu na ciljnu varijablu. Analiza relevantnosti atributa analitičaru sugeriše koji atributi imaju jak, a koji slab uticaj na ciljnu varijablu, što omogućava da se selekcija atributa koriguje da bi se izbeglo rasplinjavanje, odnosno takozvana kombinatorna eksplozija.

Pronađene zakonitosti u podacima, kao pronadeno znanje, treba prikazati na odgovarajući grafički način. Zato je važna vizualizacija podataka. Tehnikama klaster analize npr. dobijaju se brojevi podaci korelacije ili mera udaljenosti. Međutim, ljudsko oko će mnogo brže i preciznije prepoznati klastere ako su rezultati prikazani grafički umesto brojevano, a razumevanje uočene pojave će biti bolje, jer poznato je da "jedna slika daje hiljadu reči". Zato se u nastavku stavlja akcenat na vizuelizaciju analiziranih podataka.

Jedna od neizostavnih tehnika koja prati sve etape pretraživanja podataka od čišćenja, preko preprocesiranja i pretraživanja podataka je vizualizacija podataka. Vizualizacija podataka može na jednostavan i efikasan način da ukaže na osnovne smernice za dalju analize u bilo kojoj etapi istraživanja. Pojam vizualizacije se najčešće objašnjava kao “mentalna slika”. U tehničkom smislu ona se bavi prikazivanjem ponašanja odnosno stanja, posebno pri kompleksnim uslovima, na način razumljiv ljudskom oku. Ona predstavlja jednu od vrlo popularnih metoda upoznavanja s podacima pa i jednostavnijih analiza.

Vizualizaciju podataka se može posmatrati iz nekoliko osnovnih aspekta:

- ✓ vizualizacija segmenta neočišćenih podataka,
- ✓ vizualizacija očišćenih podataka,
- ✓ vizualizacija preprocesiranih podataka.

U osnovi, njen cilj je uočiti pravilnosti ili anomalije, zavisno od toga šta se želi vizualizacijom da postigne, a u svrhu povećanja kvaliteta istraživanja. Tehnike vizualizacije se mogu klasifikovati na razne načine. Najopštija je podela na:

- ✓ pretraživanje podataka,
- ✓ potvrđivanje određene hipoteze,
- ✓ manipulaciju pogledom na podatke.

Naprimera, SPSS programski paket sadrži vrlo moćne alate za vizualizaciju podataka, pri čemu prednjači modul interaktivne grafike koji omogućava kreiranje 3D vizuelnih modela podataka, koji se mogu rotirati u prostoru. Korišćenje ovih tehnika omogućava temeljno upoznavanje populacije i pruža podršku u svim fazama istraživanja.

Količina podataka koji nastaju odvijanjem poslovnih događaja i realizacijom poslovnih procesa u preduzeću proporcionalna je brojnosti tih događaja i složenosti procesa. Međutim, između količine podataka i njihovog kvaliteta ne mora uvek postojati direktna proporcionalnost. Velika količina podataka ne mora ujedno značiti visok kvalitet, vrednost i upotrebljivost tih podataka u procesima odlučivanja. Ovo navodi na potrebu za kontrolisanjem kvaliteta podataka, odnosno upravljanje sistemom u cilju prikupljanja što kvalitetnijih podataka.

Iako problem kvaliteta podataka u značajnoj meri zavisi od karakteristika firme i njenog poslovanja, moguće je uočiti četiri osnovne kategorije kvaliteta podataka, koje imaju direktni uticaj na donošenje poslovnih odluka u preduzeću:

- ✓ standardizovanost,
- ✓ podudarnost,
- ✓ verifikovanost i
- ✓ proširivost.

Standardizovanost podrazumeva ujednačenost u izražavanju i standardizaciju dokumenata koji su nosioci poslovnih informacija. Podudarnost se odnosi na identifikovanje podudarnih podataka iz brojnih izvora i njihovu eliminaciju. Verifikovanost uključuje postupak poređenja nekog podatka sa poznatim izvorom kako bi se utvrdila relevantnost podatka. Proširivost podrazumeva mogućnost dodavanja novih podataka postojećem skupu.

Prilikom osmišljavanja dobre strategije kompletnog nadzora, kontrole i upravljanja kvalitetom podataka, potrebno je uzeti u obzir set sledećih zahteva:

- ✓ jednostavnost upotrebe,
- ✓ prilagodljivost,
- ✓ efikasnost,
- ✓ nezavisnost od računarske platforme, i
- ✓ ekonomičnost.

Izračunavanje i tumačenje vrednosti poslovnih informacija vrlo je teško uraditi, između ostalog i zbog toga što informacija nije materijalna, pa se pri utvrđivanju vrednosti ne mogu primeniti konvencionalne metode. Kao odgovor na potrebe ozbiljnog pristupa kvalitetu informacija pri planiranju složenijih poslovnih sistema i aktivnosti, u srodnoj literaturi se može naći nekoliko pristupa koji se bave ovom problematikom. Sledi poglavlje posvećeno kvalitetu podataka.

4.2. Važnost kvaliteta podataka u poslovnoj inteligenciji

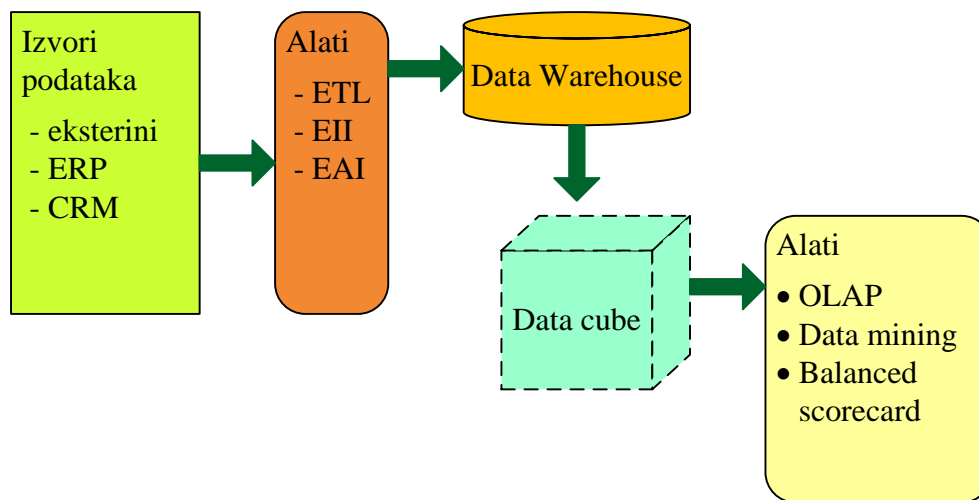
Mnoge kompanije ne shvataju ozbiljnost problema kvaliteta podataka, što često za posledicu ima to da kada se se desi katastrofalan događaj za poslovne korisnike – tek onda shvate razmere pitanja *kvaliteta* podataka. Cilj ovog dela disertacije je da predstavi ulogu upravljanja kvalitetom podataka u poslovnoj inteligenciji, kao i neke od ponuđenih rešenja kao što su servisi za upravljanje kvalitetom podataka.

Prvi korak u razvoju skladišta podataka je *analiza izvora podataka*, ključni proces koji predstavlja više od osamdeset procenata vremena u procesu izgradnje skladišta podataka. U ovom procesu potrebno je definisati pravila za preuzimanje potrebnih podataka iz izvora

podataka, izvornih sistema, jer su potrebna znanja najšee u glavama korisnika skladišta podataka. Naime, tipični problemi izvora podataka jesu sledeći, detaljno opisani u [40].

- ✓ *neKonzistentnost primarnih ključeva* – često se primarni ključevi izvornih zapisa podataka ne poklapaju. Na primer, može postojati pet fajlova o klijentima, gde svaki od njih ima različit atribut kao primarni ključ klijenta. Ovi različiti ključevi klijenata moraju se konsolidovati ili transformisati u jedan standardizovani ključ klijenta.
- ✓ *neKonzistentnost vrednosti podataka* – mnoge organizacije dupliraju svoje podatke, stvaraju kopije originalnog podatka. Tokom vremena, usled anomalija ažuriranja, prvi duplirani podaci imaju totalno različite vrednosti.
- ✓ *različiti formati podataka* – elementi podataka kao što su na primer datumi, mogu biti uskladišteni u totalno različitim formatima.
- ✓ *netačne vrednosti podataka* – da bi se korigovale netačne vrednosti podataka, mora se definisati logičko čišćenje. Algoritmi čišćenja podataka treba da se aktiviraju svaki put kada se podatak puni.
- ✓ *sinonimi i homonimi* – redundantne podatke nije uvek lako prepoznati usled toga što isti elementi podataka imaju različite nazive. S obzirom na to da sinonimi i homonimi ne smeju postojati u okruženju, neophdno je preimenovati date elemente podataka.
- ✓ *ugrađena logika procesa* – neki operacioni sistemi su ekstremno stari. Oni često sadrže nedokumentovane i arhaične relacije između pojedinih elemenata podataka. Specifikacije procesa transformacije moraju da reflektuju ovu logiku.

Posle uspešne analize izvora podataka počinje drugi korak, a to je *priprema podataka*, najbitniji proces u razvoju i izgradnji skladišta podataka, jer ukoliko je priprema podataka uspešna, biće uspešna i izgradnja skladišta podataka. Priprema podataka vrši se na osnovu već određenog izvora podataka, definisanih pravila za preuzimanje tih podataka, definisanih procedura za pripremu podataka i postavljenih zahteva korisnika. Priprema se vrši uz pomoć određenih alata za ekstrakciju, čišćenje i transformaciju podataka. Ekstrakcija i čišćenje podataka sastoji se od razvoja procedura za ekstrakciju i za čišćenje podataka, i ovo je najkomplikovaniji proces u čitavom projektu. Koristi izvore podataka iz različitih platformi koje su upravljane različitim operacionim sistemima i aplikacijama. Svrha *ETL* procesa je da spoji podatke iz heterogenih platformi u standardni format, videti [110].



Slika 4.4. Priprema podataka za proces izgradnje skladišta podataka - *ETL* proces

ETL proces počinje sa preformatiranjem podataka koji treba da unifikuje formate podataka iz različitih izvora. U drugom koraku se rešava problem konzistentnosti koji se javlja usled redundantnosti podataka. Na kraju se pristupa čišćenju onih podataka koji narušavaju poslovna pravila. Rezultati ovih aktivnosti su podaci koji će omogućiti generisanje metapodataka na osnovu kojih se dizajnira skladište podataka. *ETL* algoritmi čišćenja podataka treba da se aktiviraju svaki put kada se podatak puni. Stoga, programi transformacije ne smeju biti slabi na brzinu, već se moraju razviti na jedan strukturiran način.

EII, alat je proces integracije informacija korišćenjem apstrakcije podataka, sa ciljem obezbeđenja jedinstvenog interfejsa, poznatog kao uniformni pristup podacima, za pristup podacima unutar organizacije, videi [110]. Cilj je dobiti veliki set heterogenih izvora podataka kojeg korisnik vidi kao jednu heterogenu bazu izvora podataka. *EAI* je softver koji integriše računarske aplikacije. To je metoda, plan i alat za modernizaciju, konsolidaciju i koordinaciju aplikacija u preduzeću. *EAI* može pomoći u razvijanju potpuno novog pogleda na poslovanje preduzeća. Ekstrakcija i čišćenje podataka je prva faza u procesu pripreme podataka i sastoji se iz:

- ✓ razvoja procedura za ekstrakciju podataka, i
- ✓ razvoja procedura za čišćenje podataka.

Proces ekstrakcije podataka je izvlačenje podataka iz različitih izvornih sistema koji najčešće imaju različitu organizaciju podataka, različite formate. Programi ekstrakcije podataka treba da vrše sortiranje, filtriranje, čišćenje i da agregiraju sve zahtevane podatke, moraju da prepoznaju koji od redundantnih izvornih datoteka ili baza podataka su zapisi sistema. Na primer, isti izvorni element podatka kao što je naziv klijenta može da postoji u nekoliko izvornih datoteka i baza podataka. Ova redundantnost treba da se sortira i

konsoliduje, što uključuje korake sortiranja i spajanja, preko određenih ključeva i vrednosti podataka.

Podaci koji će biti u skladištu podataka moraju se ekstrahovati, izvući iz transakcionih sistema u kojima se željeni podaci nalaze. Nakon inicijalne ekstrakcije, na osnovu određenih procedura vrši se dodavanje novih podataka u skladište podataka. Ekstrakcija podataka je vrlo jednostavna operacija ukoliko se potrebni podaci nalaze u jednoj relacionoj bazi, ali ukoliko se podaci nalaze u višestrukim heterogenim transakcionim sistemima, izvlačenje podataka postaje veoma kompleksna operacija.

Područje prečišćavanja podataka uključuje skup procesa koji čiste, transformišu, dedupliciraju, pridružuju, arhiviraju i pripremaju podatke iz različitih izvora za skladište podataka. Često se za ove aktivnosti koriste već pomenute procedure *ETL*, *EAI*, *EII*, ili *ETT*, opisane u [27], procedura koja obuhvata sledeće procese:

- ✓ otkrivanje promena u izvornim podacima potrebnim za skladište podataka;
- ✓ izdvajanje podataka iz izvornih sistema;
- ✓ čišćenje i transformisanje podataka;
- ✓ restrukturisanje ključeva podataka;
- ✓ indeksiranje podataka;
- ✓ sabiranje podataka;
- ✓ održavanje metapodataka i
- ✓ učitavanje podataka u skladište podataka.

Navedeni procesi osiguravaju da su podaci pohranjeni u skladištu podataka:

- ✓ potrebni i korisni za poslovne korisnike;
- ✓ kvalitetni da omoguće dobre informacije;
- ✓ tačni kako bi osigurali tačne informacije i
- ✓ jednostavni za pristup kako bi se skladište podataka moglo u potpunosti efikasno koristiti od strane krajnjih korisnika.

Pojam *upravljanja podacima*, opisuje aktivnosti administratora skladišta podataka kako bi se omogućio uspešan rad skladišta: brz i siguran korisnički pristup podacima. Povećanjem obima podataka u skladištu tokom vremena ne sme se dozvoliti narušavanje performansi sistema. U slučaju da se u skladištu podataka pojave netačni ili neispravni podaci (npr. zbog ranijeg propusta u transakcionoj bazi nalazi se neistinit, semantički neispravan podatak) mora postojati brz i efikasan način da se oni uklone iz skladišta i na njihovo mesto postave ispravne vrednosti. Imajući u vidu povećan obim i kompleksnost poslovanja, sve više različitih poslovnih sistema, različito viđenje istih podataka u različitim

sistemima/odelenjima, postojanje više verzija istine i nekonzistentnosti podataka, nedefinisane odgovornosti za podatke, nesinhronizovanost različitih poslovnih rešenja i tako dalje - jasno je da upravljanje podacima ne samo neophodno, već i ključno za ostvarivanje što veće efikasnosti. Upravljanje podacima upravo rešava mnoge probleme koji se odnose na relevantnost, dostupnost, legalnost, sigurnost, vrednost, itd.

Upravljanje podacima obuhvata sve discipline koje se odnose na upravljanje podacima kao vrednim resursom. Službena definicija Međunarodne Asocijacije za upravljanje podacima glasi: "Upravljanje podacima (kao resursom) je razvoj i izvođenje arhitekture, politike, prakse i procedura koje na odgovarajući način upravljaju potpunim ciklusom podataka neophodnih u preduzeću." Ova definicija je poprilično široka i obuhvata značajan broj zanimanja koja možda nemaju direktan tehnički dodir s nižim tehničkim stepenom upravljanja podacima, gde spada npr. i upravljanje bazom podataka. Da bi se bolje razumelo šta je upravljanje podacima, treba sagledati šta ono zapravo radi, videti [40]:

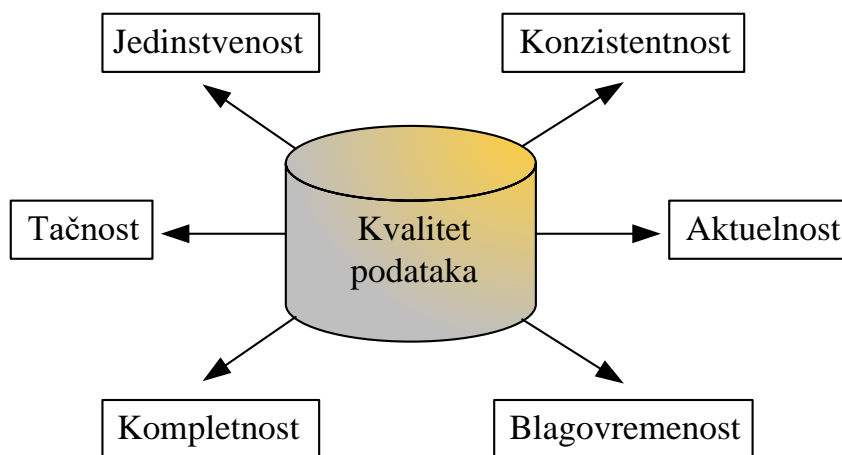
- ✓ Utvrđuje koje interesne grupe su vlasnici podataka, (značenje podataka, sadržaj i pristup reflektuju poslovnu politiku);
- ✓ Pozicionira podatke kroz odeljenja kompanije, (politikama i promenama podataka upravlja organizacioni tim);
- ✓ Osigurava da kvalitet podataka bude u skladu sa poslovnim merama i ciljevima prihvatljivosti, (prihvatljivost i tačnost podataka su opisani u skladu sa poslovnom politikom);
- ✓ Omogućava sklad i složnost IT-a od mogućih konflikata oko vlasništva podataka (IT su inženjeri podataka, a ne oni koji rešavaju probleme vlasništva nad podacima);
- ✓ Osigurava da se podacima upravlja nezavisno od aplikativnih sistema;
- ✓ Uključuje podatke u upravljanje promenama, (aplikativni sistemi reflektuju kako poslovne procese, tako i zahteve za podacima);
- ✓ Uspostavlja monitoring progressa i merenje efekata poboljšanja.

Upravljanje kvalitetom podataka jeste deo jedne šire discipline koja se zove *upravljanje podacima*. To je disciplina koja, u cilju jedinstvenog upravljanja poslovnim podacima, objedinjuje, detaljno u [139]:

- ✓ kvalitet podataka (data quality);
- ✓ upravljanje podacima (data management);
- ✓ politike podataka (data policies);
- ✓ bezbednost podataka (data security);
- ✓ upravljanje poslovnim procesima (business process management) i
- ✓ upravljanje rizicima (risk management).

Kvalitet podataka zavisi od niza događaja od kojih je većina van kontrole tima za razvoj i implementaciju skladišta podataka. Da bi se u skladištu podataka dobili kvalitetni podaci, proces prikupljanja podataka treba biti dobro dizajniran. Dobar kvalitet podataka znači da su svi osnovni podaci potpuni, konzistentni, tačni, pravovremeni i zasnovani na tehničkim standardima. Poboljšavajući kvalitet podataka, trgovinski partneri smanjuju troškove, povećavaju produktivnost i ubrzavaju poslovne procese usmerene ka tržištu. Dakle, karakteristike kvalitetnog podatka su:

- ✓ tačnost – podaci u skladištu podataka se podudaraju sa sistemom slogova;
- ✓ kompletnost – podaci u skladištu podataka predstavljaju zaokružen skup bitnih podataka;
- ✓ konzistentnost – podaci u skladištu podataka nemaju kontradikcija;
- ✓ jedinstvenost – dva ista elementa trebaju pozivati istu stvar i imati istu ključnu vrednost;
- ✓ vreme – podaci moraju biti ažurirani prema odgovarajućem rasporedu koji treba biti javan, sva odstupanja se trebaju beležiti.



Slika 4.5. Aspekti kvaliteta podataka, videti [139]

U tipičnom okruženju poslovne inteligencije, podaci iz različitih izvora su ekstrahovani, transformisani i učitani u skladište podataka poslovnog sistema, odakle se dalje koriste za izveštavanje. Kvalitet podataka igra ključnu ulogu kada je reč o uspešnosti poslovne inteligencije, jer loš kvalitet podataka može vrlo negativno uticati na poslovne odluke na svim nivoima kompanije. Kvalitet određuje korisnost podataka kao i kvalitet odluka koje se zasnivaju na podacima. Pored složnosti kao jedne od najpoznatijih karakteristika problema kvaliteta podataka, u isto vreme to je i pitanje (problem) koje jako puno košta. Jedan od razloga za to jeste činjenica da se podaci koji se koriste u okruženju poslovne inteligencije ujedno koriste i za donošenje odluka; štaviše, određeni set podataka može biti upotrebljen u više procesa odlučivanja. Da bi se bolje razumelo šta je kvalitet podataka, treba sagledati dimenzije koje karakterišu kvalitet podataka:

- ✓ Suštinski kvalitet podataka – odnosi se na tačnost, objektivnost, verodostojnost i reputacija;
- ✓ Kontekstualni kvalitet podataka – odnosi se na relevantnost, dodatnu vrednost, vremensku neograničenost, potpunost, količinu podataka;
- ✓ Prikazivanje kvaliteta podataka – odnosi se na mogućnost tumačenja, lakoću razumevanja, sažetost predstavljanja i dosledno prikazivanje;
- ✓ Dostupnost – odnosi se na dostupnost i sigurnost pristupa.

Kvalitet podataka zavisi od stepena razvijenosti informacionog sistema poslovnog sistema i razvijenosti sistema skladišta podataka, koji obuhvata sve relevantne attribute koji se planiraju koristiti u daljim analizama. U njima su približno podaci već čisti, integrisani i potpuni.

Činjenica je da je do skoro relativno malo pažnje posvećivano problemu kvaliteta podataka u poslovnoj inteligenciji podataka na osnovu kojih se kasnije donose bitne poslovne odluke. Izvorni sistemi „snabdevaju“ podacima sisteme poslovne inteligencije koja ih na određeni način agregira i prezentuje na određeni način. Ako ulazni podaci ne ispunjavaju određene nivoe kvaliteta, nerealno je očekivati da će korist od projekata i aplikacija biti dobra i zadovoljavajuća, čak i ako su ti projekti i aplikacije tehnički urađeni savršeno.

Bez obzira na to da li se radi o projektima koji se odnose na skladišta podataka, planiranja i sl., kvalitet postojećih podataka će direktno uticati na rezultate tih projekata. Često, imajući u vidu da su dostupni podaci lošeg kvaliteta, korisnici odustaju od projekata poslovne inteligencije, umesto da rade na poboljšanju kvaliteta podataka kao ključnom faktoru za uspešnost ne samo poslovne inteligencije, nego generalno cele kompanije.

Štaviše, loš kvalitet ima jak negativan uticaj i na proizvodne sisteme. Posledica svega toga može biti slabija produktivnost i mnogo više grešaka pri obavljanju raznih operacija u poslovanju. Na kraju, iz ovoga proizilazi nemogućnost za pravilno praćenje poslovnih aktivnosti i donošenje dobrih i ispravnih poslovnih odluka, a ujedno zbog svega toga postoje ogromni zahtevi za ljudskim i IT resursima. Činjenica je da danas još uvek veoma malo kompanija ima predstavu o tome koliko je kvalitet podataka bitan i u većini slučajeva postoji barem jedan projekat koji je propao ili je neefikasan zbog lošeg kvaliteta podataka. Bitno je napomenuti da problem ne leži u projektima poslovne inteligencije, već u samim podacima sa kojima se raspolaže.

U nastavku ovog poglavlja, akcentat se stavlja na slučaj nedostajućih podataka.

4.3. Problem nedostajućih podataka

Već je istaknuto da je uloga analitičara poslovanja, pre svega otkrivanje odgovarajućih međuzavisnosti između podataka. Da bi se postigao ovaj zahtev, uzima se reprezentativni uzorak iz populacije podataka koja je od interesa i sprovodi se istraživanje nad tim skupom podataka u cilju pronalaženja pomenutih međuzavisnosti. Ukoliko reprezentativni uzorak sadrži *nedostajuće vrednosti*, to predstavlja veliku prepreku za dalji rad, jer otežava izvođenje pouzdanih i tačnih zaključaka. Upravo iz tog razloga, analitičari se uvek trude da izbegnu nedostajuće vrednosti.

Ključna pitanja za analitičara koji se susreće sa nedostajućim vrednostima u podacima su:

- ✓ Koliko je problem nedostajućih vrednosti podataka ozbiljan?
- ✓ Da li model istraživanja može da se prilagodi tome?

- ✓ Da li je ispravno pretpostaviti da je raspodela nedostajućih vrednosti u potpunosti slučajna?
- ✓ Da li se mogu primeniti tehnike koje dozvoljavaju nedostajuće vrednosti?

Sa aspekta klijenta za kog se vrše analize, tj. onoga koji dostavlja podatke, pitanja bi okvirno mogla biti u sledećoj formi:

- ✓ Da li je moguće na neki način izbeći, izbrisati opservacije sa nedostajućim vrednostima, a da to ne šteti modelu?
- ✓ Da li se vrednosti koje nedostaju mogu naknadno dopuniti?
- ✓ Zna li se razlog što vrednosti nedostaju? Ako razlog postoji, da li se može definisati neka pravilnost?

Tehnike rešavanja problema nedostajućih vrednosti su dosta raznolike i variraju od potpunog ignorisanja nekompletnih slogova, do primene kompleksnih algoritama. Primena pomenutih algoritama omogućava da se nedostajuće vrednosti zamene procenjenim vrednostima u cilju prikaza što realnije slike posmatrane pojave.

Budući da prisustvo nedostajućih vrednosti dosta smanjuje kvalitet ulaznih podataka i navodi na donošenje nepouzdanih zaključaka, samim tim i odluka koje predlaže tim (grupa), idealno bi bilo izbeći ih, koliko god je to moguće. U svakom slučaju, pre sprovođenja istraživanja, istraživač je taj koji treba da razmisli o tome kolika je verovatnoća da se pojave nedostajuće vrednosti. Zatim je potrebno da odabere metodu analize podataka, koja je takva da se može primeniti na podatke kod kojih kršenje polaznih pretpostavki (hipoteza) vezanih za nedostajuće vrednosti, nema negativni uticaj na izvođenje konačnih zaključaka i odgovora na pitanja koja istraživanje postavlja. Ukoliko se ispostavi da ne postoji tehnika analize podataka koja bi odgovorila na pomenute zahteve, potrebno je razmotriti opciju predviđanja nedostajućih vrednosti.

Predviđanje nedostajućih vrednosti može se realizovati uz primenu jedne od nabrojanih tehnika:

- 1) Delimična analiza – koristi metode koje podrazumevaju isključivanje nedostajućih vrednosti:
 - ✓ Uklanjanje celih nekompletnih slogova (*listwise deletion*),
 - ✓ Uklanjanje nedostajućih vrednosti iz nekompletnih slogova (*pairwise deletion*).

- 2) Komleptna analiza – primenjuje metode koje uzimaju u obzir sve raspoložive slogove, s tim što se umesto nedostajućih vrednosti koriste vrednosti dobijene predviđanjem:

- ✓ Algoritam maksimizacija-očekivanje (*expectation-maximization algorithm*),
- ✓ Ocena maksimalne verodostojnosti (*maximum likelihood estimation*),
- ✓ Višestruko predviđanje (*multiple imputation*).

Prilikom analize seta podataka koji sadrži nedostajuće vrednosti, prvo od pitanja koje se postavlja jeste, kako prepoznati da li je uzrok što neke vrednosti nedostaju slučajnost ili ne. Za početak, ako postoji jako mali broj nedostajućih vrednosti, vrlo malo je verovatno da to nije slučajnost. U praksi, svaki skup podataka sadrži bar nekoliko nedostajućih vrednosti. Što više atributa je uključeno u studiju, veća je verovatnoća da postoji neki propust. Što je više izvora iz kojih se ekstrahuju podaci, veća je verovatnoća da u nekom od njih postoji greška. U svakom slučaju, nepisano pravilo je da ako nedostajućih vrednosti ima do 5% od ukupnog skupa, to se smatra slučajnim.

Dalje, primer za situaciju kada se mali broj nedostajućih podataka ne smatra slučajnim je kada sve nedostajuće vrednosti dolaze iz istog izvora. Ukoliko je tako, to ukazuje da postoji određeni problem, npr. površno bavljenje istraživanjem, nekorektno odnošenje prema koracima istraživanja, nepodržavanje procedure za sprovođenje istraživanja, itd.

Čak i ako se nedostajuće vrednosti javljaju u većem broju, to ne znači obavezno da su one ne-slučajne. Na neka pitanja često izostaju odgovori zato što pitanje previše zadire u privatnost ili je postavljeno na određen način koji negativno utiče na ispitanike.

Kao primer može se navesti da je u programskom paketu SPSS-u razvijen modul pod nazivom Analiza nedostajućih vrednosti (*Missing Value Analysis*) kojim se utvrđuje da li su nedostajuće vrednosti slučajne ili ne. Ovo se postiže korišćenjem metode očekivanje-maksimizacija koja proverava da li se slogovi koji imaju nedostajuće vrednosti bitno razlikuju od onih koji ih nemaju.

Nekompletni podaci se, u praksi, sreću kao sastavni deo svakog istraživanja i nakon njihove identifikacije potrebno je analizirati još nekoliko bitnih faktora. Jedan od njih je definisanje uzroka nekompletnosti podataka, budući da to igra jako važnu ulogu u postavljanju polaznih pretpostavki koje su osnova za sledeći korak, a to je odabir tehnika i metoda koje će se koristiti za predviđanje.

Mogući uzrok nekompletnosti podataka

Postoji više razloga što se podaci u istraživanjima javljaju kao nekompletni. Na primer, mogu da variraju od toga da je oprema koja se koristi u istraživanju neispravna, preko lošeg vremena u trenutku sprovođenja istraživanja, do greške pri samom unosu podataka. Tada se smatra da je *uzrok nekompletnosti u potpunosti slučajan* (*Missing Completely at Random* -

MCAR). To znači da verovatnoća da neki podatak X_i jedne opservacije nedostaje nije vezana za samu vrednost X_i , niti za vrednosti ostalih varijabli. Na primer, nedostatak podatka o prihodu porodice ne smatra se slučajnim, jer je poznato da se ljudi sa nižim prihodom ređe rešavaju da prijave koliki im je prihod od onih sa višim primanjima. Slična je situacija i sa podelom na belce i crnce, zna se da su crnci pre spremni da kažu kolika su im primanja. U tom slučaju, etnička pripadnost ima uticaja na verovatnoću da ovaj podatak izostane. Nasuprot tome, ako ispitanik usled saobraćajnog prekršaja bude sprečen da prisustvuje istraživanju, tada se nedostatak ovih podataka može svrstati u grupu *MCAR*. Drugim rečima, u ovoj gupi se nalaze podaci kod kojih je verovatnoća nekompletnosti za bilo koju od opservacija koje su obuhvaćene ista.

Ono što je važno je sama vrednost opservacije, a ne da li je ona kompletna ili ne. Ukoliko, na primer, neki ispitanici ne žele da prijave visinu prihoda, ni ličnih, ni prihoda cele porodice, to i dalje ne mora značiti da ovi podaci ne mogu biti *MCAR*, sve dok odgovor nije u vezi sa samim iznosom. Dobra strana ove vrste podataka je što ne narušava rezultate analiza. Iako ima nedostataka, izlazni parametri nisu pogođeni ovim problemom.

Sledeći nivo klasifikacije nedostajućih podataka je *grupa slučajne nekompletnosti (Missing at Random - MAR)*. Za podatke koji su u *MCAR* grupi, je istaknuto da verovatnoća da X_i nedostaje nije povezana sa vrednošću X_i , ni sa drugim varijablama u analizi. Međutim, može se smatrati da podaci nedostaju slučajno ukoliko to nije vezano za samu vrednost X_i , ali ima veze sa nekom drugom promenljivom. Na primer, ljudi koji su depresivni mogu pokazivati manje tendencija da prijave kolika su im primanja, tako da na sam podatak nema uticaja visina prihoda, već depresija. Depresivni ljudi takođe mogu imati niži prihod generalno, tako da ukoliko imamo visok procenat nedostajućih podataka vezano za ovakve osobe, realni prosečni prihod bi bio manji kad ne bi bilo nedostajućih podataka. U svakom slučaju, smatra se da ovaj tip podataka kvalifikuje da bude u grupi *MAR*.

Važno je naglasiti da to što se ovo na određen način smatra slučajnim, nažalost, ne znači da nema posledica. Ovakvi podaci ipak narušavaju realnu sliku posmatranog problema. Međutim, zahvaljući razvijenim naučnim metodama u ovoj oblasti moguće je proizvesti smislene i relativno realne statističke ocene.

Pored grupa sasvim slučajano nekompletnih podataka i slučajano nekompletnih podataka, postoje i *podaci čiji uzrok nekompletnosti nije slučajan (Missing Not at Random - MNAR)*. Na primer, ako je predmet proučavanja mentalno zdravlje i zna se da su ljudi kojima je dijagnostikovana depresija manje skloni da prijave svoj mentalni status, u tom slučaju nedostajući podaci ne smatraju se slučajnim. Jasno je da prosečni mentalni status za

raspoložive podatke neće biti realna ocena proseka koji bi se dobio da su svi podaci bili raspoloživi. Isti je slučaj i sa istraživanjem prihoda, gde se zna će osobe sa nižim prihodima u manjem procentu odgovoriti na postavljena pitanja.

Ova grupa podataka zaista predstavlja problem. Jedini način da se postignu dobre statističke ocene je da se za problem nedostajućih podataka napravi odgovarajući model koji bi ih obradio na određen način, po mogućstvu korišćenjem neke od metoda za predviđanje nedostajućih vrednosti.

Slede neka moguća rešenja problema nedostajućih vrednosti. Jedan od najčešćih pristupa zasniva se na pravilu da se jednostavno zaobiđu slogovi sa nedostajućim podacima i analize se zatim vrše nad onim što je preostalo. U tom slučaju, ako se pet ispitanika ne pojavi u zakazano vreme kad se sprovodi istraživanje, ta grupa će imati pet opservacija manje. Takođe, ako pet osoba ne da odgovore za jednu ili više varijabli, ti slogovi se jednostavno isključuju iz analize. Ovaj pristup se obično naziva uklanjanje slogova (*listwise deletion*), ali se često referencira i kao analiza samo kompletnih slogova.

Iako ovaj pristup rezultira u značajnom smanjenju veličine uzorka koji je predmet analize, on ima i svojih prednosti. Konkretno, pod pretpostavkom da je uzrok nekompletnosti podataka slučajna, statističke ocene neće biti ugrožene. Međutim, čak i kada su podaci u grupi MCAR, ipak je evidentno da su mogućnosti kod korišćenja ovog pristupa značajno slabije. U situacijama kada nekompletnost podataka nije slučajna, posebno se vide slabosti ove metode, jer je tada sasvim evidentan negativan uticaj na rezultate analiza. Primer je kada osobe sa nižim prihodima oklevaju da prijave visinu prihoda, tada se analizom dobija da su prosečna primanja grupe koja je predmet istraživanja viša nego što je to zaista slučaj. Prethodno je istaknuto i postojanje drugog pristupa. Čak i veliki broj računarskih paketa nudi opciju korišćenja pristupa poznatog kao delimično uklanjanje vrednosti (*pairwise deletion*), a u krugovima statističara i kao 'ne bas pametno' uklanjanje vrednosti (*'unwise' deletion*). Kod ovog pristupa, ocena svakog elementa matrice interkorelacije se vrši tako što se uzimaju u obzir svi raspoloživi podaci. Ukoliko neki ispitanik odgovori na pitanje kolika su mu mesečna primanja i koliko je zadovoljan svojim životom, ali ne i na pitanje koliko ima godina, ta opservacija je uključena u analizu prihoda i mišljenja o sopstvenom životu, ali ne i u deo koji se odnosi na godine. Najveći nedostatak ovog pristupa je što su parametri koji se dobijaju iz modela zasnovani na različitim skupovima podataka, različite su veličine uzorka i standardne greške. Čak je veoma verovatno da će se izgenerisati matrica koja nije sasvim tačna i pouzdana, što je velika prepreka za samu analizu.

Opšte mišljenje je da, ukoliko postoji vrlo mali broj nekompletnih opservacija, upotreba ovog pristupa neće izazvati veliku štetu. Opet, u tom slučaju, ne bi bilo loše ni isključiti cele te slogove koji su problematični, jer ih nema mnogo. Međutim, ako nekomletnih opservacija ima u većem broju, kvalitet rezultata analize bi bio znatno lošiji.

4.3.1. Tradicionalni pristupi za rešavanje problema nedostajućih podataka

Nedostatak podataka o pripadnosti grupi

U eksperimentalnim istraživanjima, obično se zna kojoj grupi ispitanik pripada, jer se istraživanje i vrši po grupama. Osim ako se desi da se podaci izmešaju, pripadnost grupi se lako određuje. Ipak, postoje i istraživanja kod kojih ovo nije poznato. Primer je sprovođenje istraživanja o nivou optimizma u različitim religijama, sa različitim antropološkim pripadnostima. Jedan od načina za ovo je sprovesti istraživanja u verskim institucijama (crkvama, sinagogama, ...) i tada je jasna klasifikacija subjekata po grupama, budući da se zna gde je koji ispitanik učestvovao u istraživanju. Drugi način bi bio da se istraživanje sprovede među ljudima na ulici, pri čemu se dodaje i pitanje koje veroispovesti je ispitanik. Neki ispitanici mogu odgovoriti da su ateisti, što je jedna od očekivanih opcija, ali isto tako odgovor može izostati jer neko ne želi da se deklariše, iako je vernik. U ovom slučaju, teško bi bilo dokazati da su ovakvi odgovori jednako zastupljeni u svim religijama. Svakako, nedostatak podatka o veroispovesti, ne može se poistovetiti sa pretpostavkom da ispitanik nije vernik.

Najjednostavnije bi bilo ignorisati sve opservacije gde pripadnost verskoj zajednici nije poznata, i pretpostaviti da je nekompletnost podataka slučajna. Međutim, malo je činjenica koje bi išle u prilog ovome. Zato je bolje uzeti i ove opservacije u obzir, tako što ih treba svrstati u posebnu grupu. Zatim bi mogao da se sprovede test hipoteze, da se grupa onih koji nisu odgovorili na pitanje o religiji ne razlikuje od ostalih grupa po posmatranoj varijabli, u ovom primeru po optimizmu. Ukoliko se ova hipoteza ispostavi kao tačna, može se nastaviti sa daljim analizama na osnovu ostalih grupa. U drugu ruku, ako se pokaže da se grupa neizjašnjenih dosta razlikuje u optimizmu od ostalih grupa, te razlike se moraju uzeti u obzir. Detaljno prikazano u [33].

Nedostajući podaci za zavisnu promenljivu

Priroda problema je sasvim drugačija kada nedostaju podaci za neki zavisni atribut i u tom slučaju su rezultati samog istraživanja mnogo teži za interpretiranje. Ukoliko se sprovodi jednostrana analiza varijanse i može se pretpostaviti da je nedostatak podataka u potpunosti slučajan, situacija i nije toliko loša. Postoji gubitak u smislu smanjenja veličine uzorka i ocene za veće grupe da bi imale manje greške, ali generalno ocene bi bile u dovoljnoj meri

realne. Ipak, ono što uvek predstavlja problem je da u praksi podaci uglavnom ne nedostaju slučajno.

Primeru radi, ako je predmet istraživanja poređenje dva načina lečenja hipertenzije. Idealno bi bilo kad bi postojale dve grupe ispitanika koji bi uzimali terapiju koja im je propisana, pri čemu bi se na kraju testnog perioda sprovelo merenje krvnog pritiska za sve njih. Međutim, medicinska istraživanja su posebno poznata po problemu da se tokom sprovođenja eksperimenta, eksperimentalna grupa često osipa. One osobe kojima terapija u određenom roku ne pomogne, uglavnom odustaju ili se čak dešava težak ishod. Tako, ako je jedna terapija jako uspešna, a druga ne pokazuje skoro nikakve pozitivne efekte, veličina uzorka za drugi tip lečenja će biti dosta manja. Pored toga, oni koji izdrže do kraja su mahom u grupi pacijenata koji su imali bar neke koristi od primene propisane terapije. Na taj način, ukoliko se pokaže da je matematičko očekivanje sredine za dve posmatrane grupe približno isto, to može navesti na zaključak da je efektivnost terapija na približnom nivou. A realno, druga terapija ima daleko lošije rezultate, ali podaci o tome su izostali. Kod ovog tipa nedostajućih podataka, interpretiranje matematičkih očekivanja svakako nije pouzdan pokazatelj i bilo bi znatno bolje poslužiti se nekim drugom statističkom ocenom kao što je, na primer, procenat uzorka koji je eliminisan iz polazne grupe. Ipak, detaljnije proučavanje ove vrste problema po kompleksnosti prevazilazi okvire ove disertacije.

4.3.2. Predviđanje nedostajućih vrednosti

Slede dva pristupa predviđanja nedostajućih podataka, tradicionalni i savremeni. Kratak opis i jednog i drugog.

Tradicionalni načini predviđanja

Postoji nekoliko pristupa za predviđanje koji su se ranije u velikoj meri koristili. I pored toga što imaju dosta nedostataka, bitno je istaći koje su njihove osnovne karakteristike i zašto treba biti oprezan u njihovom korišćenju. U svakom slučaju, oni su predstavljali logičku osnovu za razvoj modernijih metoda i tehnika.

Zamena nekompletnih slogova sličnim slogovima

Ova metoda se u literaturi sreće pod nazivom '*hot deck imputation*', i u suštini predstavlja zamenu nekompletnih slogova vrednostima sličnih slogova iz istog skupa podataka. Sam naziv potiče još iz vremena kada su se kao medijum za čuvanje podataka koristile IBM-ove bušene kartice. Termin '*deck*' je iskorišćen kao aluzija na špil karata, tj. jednu grupu kartica ili skup srodnih podataka, od kojih će jedna biti izvučena; a '*hot*' potiče od toga što je taj skup podataka bio predmet trenutne obrade.

Ovaj pristup je korišćen pre više od pola veka, od strane Biroa za popis stanovništva i srodnih institucija u Sjedinjenim američkim državama. Prednost je što je, četrdesetih i pedesetih godina dvadesetog veka, većina građana ispunjavanje upitnika u popisu shvatala kao svoju dužnost i zahvaljujući tome, nedostajući podaci su se javljali u jako malom obimu. A sa nekompletnim podacima se postupalo na sledeći način: recimo da u popisu 1950. za jednog mladog muškarca, koji je bio crnac i pripadao određenoj geografskoj popisnoj jedinici, nisu bili raspoloživi podaci ili da je jednostavno odbio da učestvuje. Postupak je podrazumevao da se uzme skup podataka (skup kartica) za sve mlađe muškarce koji su crnci i pripadaju toj istoj geografskoj jedinici, i nasumično izvuče jedna kartica. Nadalje, ona je predstavljala zamenu za nedostajuću opservaciju, i analiza se nastavljala. Iako izgleda prilično površno, ovo ipak nije bila toliko loša procedura, budući da je, kao prvo, procenat nedostajućih podataka bio zaista mali, kao drugo, zamena se vršila uzimanjem sloga iz skupa sličnih slogova, i kao treće, dobro se znalo kakve su bile implikacije na statističke rezultate. Važno je da je svojevremeno poslužila svrsi, a danas je njena primena uveliko prevaziđena.

Zamena srednjom vrednošću za posmatrani atribut

Reč je o široko primenjivanoj proceduri, najpre zbog jednostavnosti primene. Ideja je u zameni nedostajućih vrednosti srednjom vrednošću dobijenom za posmatrani skup podataka. Na primer, ako za jednu osobu nedostaje podatak o visini krvnog pritiska, kao zamena uzima se srednja vrednost visine krvnog pritiska za ceo skup podataka. Prvi od nedostataka ovog pristupa je to što pomenuta zamena ne donosi nikakve nove informacije. Matematičko očekivanje sredine će biti isto, i pre i posle sprovođenja ovog postupka, ali će ocena greške biti manja. Na primeru podataka sa jednog univerziteta koji se odnosi na zaradu koju donose pojedine knjige i učestalost njihovog citiranja u radovima iz iste naučne oblasti, došlo se do nekih zaključaka. U primeru ima 62 sloga sa kompletnim podacima i 7 slogova koji imaju nedostajuće vrednosti za učestalost citiranja. Rezultati analize su dati u sledećoj tabeli.

Analiza	N	r	b	St. Err. b
Kompletni slogovi	62	.55	310.747	60.95
Nakon zamene sr. vrednošću	69	.54	310.747	59.13

Tabela 4.1. Mere varijacije pre i posle zamene srednjom vrednošću

Da se primetiti da postoji minorna promena za koeficijent korelacije r i da je koeficijent regresije b ostao isti. Međutim, standardna greška je vidno manja nakon zamene srednjom vrednošću. To je i logično, obzirom na to da nije dodata nikakva nova informacija, a veličina uzorka je povećana. Efekat koji se dobija povećanjem uzorka je povećanje delioca u formuli za računanje standardne greške, drugim rečima, standardna greška se smanjuje. To, naravno, ne znači da su rezultati druge analize pouzdaniji.

Regresija

Jedna od opcija za predviđanje nedostajućih vrednosti je korišćenje linearne regresije na osnovu kompletnih podataka. Ovaj pristup ima široku primenu, kao i zamena srednjom vrednošću, ali ima bar jednu prednost u odnosu na nju. To je da vrednost koja se dobija predviđanjem bar na neki način zavisi od ostalih varijabli. Kod zamene srednjom vrednošću, ukoliko nedostaje težina za jednu osobu, na primer, muškarca, samo bi se uzela prosečna težina za ceo posmatrani skup. Regresijom bi se uzela prosečna težina muškaraca sličnih godina. Problem greške varijanse ipak ostaje. Zamenom vrednošću koja je predvidiva na osnovu ostalih atributa, nije dodata nikakva nova informacija u analizu, a povećana je veličina uzorka i smanjena standardna greška. Postoji način kojim se ovo rešava, naime, u SPSS-u se kroz opciju Analiza nedostajućih vrednosti svakom predviđanju, dodaje neka mala greška nasumično. Ovo ne rešava problem u potpunosti, ali ga bar delimično ublažava.

Savremeni načini predviđanja

Predviđanje na osnovu najveće verodostojnosti

Princip na kome se zasniva metoda najveće verodostojnosti je u osnovi jednostavan, ali je rešenje računski složeno. Suština se može pokazati na primeru računanja srednje vrednosti za neki atribut populacije.

Pretpostavimo da postoji reprezentativni uzorak od četiri vrednosti i to su 1,4,7,9 i treba da se izračuna srednja vrednost za celu populaciju. Pretpostavka je da vrednosti populacije imaju normalnu raspodelu i da se označi srednja vrednost populacije sa μ . Može se izračunati verovatnoća da se pojave vrednosti 1,4,7 i 9 za bilo koju određenu vrednost μ . To bi bio proizvod:

$$p(1) * p(4) * p(7) * p(9)$$

Verovatnoća koja bi se dobila, bila bi veoma mala u slučaju kad bi stvarna vrednost μ bila jednaka 10; takođe bila bi velika ukoliko bi stvarna vrednost bila između 4 i 5, a maksimalna

vrednost bi se postigla za $\mu=5.25$. Za svaku različitu vrednost μ računa se verovatnoća da se pojavi 1, $p(1)$, 4, $p(4)$, itd. I računa se njihov proizvod. Za jednu od vrednosti μ ispostaviće se da daje veći proizvod nego bilo koji drugi. Ta vrednost se zove ocena najveće verodostojnosti za μ . U ovom slučaju, evidentno je da je ocena najveće verodostojnosti za srednju vrednost populacije ustvari srednja vrednost uzorka, zato što se najveća verovatnoća da će se pojaviti vrednosti 1,4,7 i 9 dobija kada je μ jednako srednjoj vrednosti uzorka.

Na sličnim principima se zasniva i regresija, mada je ona znatno komplikovanija. Ako se pretpostavi da postoji normalna raspodela posmatranih vrednosti, mogu se izračunati ocene maksimalne verodostojnosti za srednju vrednost, varijansu i kovarijansu na osnovu datog uzorka. Kad se izračunaju ove ocene, mogu se iskoristiti za izvođenje jednačine regresije. Kod uobičajenih statističkih modela ova metoda se često koristi za ocene parametara zato što je odlikuju efikasnost, dovoljnost i konzistentnost. Međutim, ako je statistički model kompleksan, sa većim brojem nepoznatih parametara, jednačine funkcije verodostojnosti se izuzetno usložnjavaju, pa se u tim slučajevima pribegava iterativnim metodama za ocenu parametara modela.

Algoritam Očekivanje-Maksimizacija

Metoda očekivanje-maksimizacija predstavlja najčešće primenjivani pristup za iterativno izračunavanje ocena metodom najveće verodostojnosti. U literaturi se sreće pod imenom EM algoritam (*Expectation-Maximization algorithm*).

Osnovna ideja metode očekivanje-maksimizacija je povezivanje datih problema nekompletnih podataka sa kompletnim podacima, kod kojih je ocenjivanje metodom najveće verodostojnosti računski mnogo jednostavnije. Na primer, izabrani problem kompletnih podataka može biti u algebarski nerešivom obliku, za ocenjivanje metodom najveće verodostojnosti ili možda neće biti moguće izračunati ocene najveće verodostojnosti pomoću standardnih metoda. Postupak metode očekivanje-maksimizacija se onda sastoji u tome da se problem preformuliše tako da se ocena najveće verodostojnosti može izračunati, uspostavljajući vezu između funkcija verodostojnosti ova dva problema, i korišćenjem jednostavnijeg izračunavanja, ocena najveće verodostojnosti u koraku maksimizacije metode očekivanje-maksimizacija.

Sušтина je da, ako bi se znalo koje vrednosti treba da se pojave umesto nedostajućih, bilo bi lako izračunati ocene parametara modela. Slično, ako bi se znali parametri modela, bilo bi lako uraditi predviđanje za nedostajuće vrednosti. U ovom algoritmu se radi i jedno i drugo. Prvo se vrše ocene parametara na osnovu raspoloživih podataka. Zatim se vrši predviđanje

nedostajućih podataka na osnovu dobijenih parametara. Onda se ponovo vrše ocene parametara, na osnovu dopunjenog skupa podataka.

Preciznije, postupa se po sledećim koracima: računaju se ocene varijanse, kovarijanse i srednje vrednosti, na primer, korišćenjem pristupa uklanjanja nekompletnih slogova (*listwise deletion*); zatim se to koristi da se izračunaju koeficijenti regresije, i na kraju se nedostajuće vrednosti računaju na osnovu dobijene relacije regresije. Na primer, na osnovu prvobitno raspoloživih podataka vrši se ocena regresije $\hat{Y}=bX+a$, a onda na osnovu X , računamo Y gde god nedostaje. Opisani koraci su deo algoritma koji se odnosi na očekivanje. Nakon što su nedostajuće vrednosti zamenjene predviđanjem, nad novodobijenim skupom kompletnih podataka se ponovo računaju koeficijenti regresije. U prethodnom tekstu je već pomenut problem da se povećanjem uzorka na ovaj način dobija da je standardna greška manja nego što stvarno jeste. U EM algoritmu se ovaj problem rešava tako što se dodaje mala greška pri oceni varijanse, i onda se ta ocena koristi da se izračunaju nedostajuće vrednosti, dok se rešenje ne stabilizuje. Tada se može reći da za parametre modela postoje ocene najveće verodostojnosti i one se mogu iskoristiti za proračun konačnih ocena maksimalne verodostojnosti za koeficijente regresije.

Moguće je dobiti i bolje ocene maksimalne verodostojnosti od ovih koje se dobijaju pomoću EM algoritma, ali za to je potrebno dobro poznavati model raspodele atributa sa nedostajućim vrednostima.

Višestruko predviđanje

Ova metoda predstavlja alternativu za metod maksimalne verodostojnosti. Prethodna rešenja koja su razmatrana podrazumevaju predviđanje nedostajućih vrednosti i zatim korišćenje i njih i kompletnih slogova u daljoj analizi. U najprostijim slučajevima nedostajuće vrednosti se zamenjuju konstantom (najčešće srednjom vrednošću), a kod EM algoritma vrši se zamena vrednostima dobijenim na osnovu parametara izvedenih iz kompletnih slogova. Kod višestrukog predviđanja se svaki nedostajući ili nekompletni podatak zamenjuje sa dve ili više prihvatljivih vrednosti koje su izvučene iz raspodele podataka. Znači da je svaka izgubljena vrednost zamenjena sa nekoliko umetnutih vrednosti u obliku vektora.

Predviđanje se u ovoj metodi oslanja na postojeće podatke, kao i kod EM algoritma, ali postoji bitna razlika. Uz pretpostavku da treba izračunati Y na osnovu X . Za svako $X=5$, Y će dobiti istu vrednost. Ovo vodi ka smanjenju standardne greške za regresione koeficijente, jer je smanjen varijabilitet za vrednosti dobijene predviđanjem u odnosu na situaciju kad ne bi bilo nedostajućih vrednosti. Jedno rešenje je ono koje je primenjeno u EM metodi, gde se

formula za računanje delimično menjala dodavanjem određene greške u kalkulaciju. Kod višestrukog predviđanja, uzimaju se izračunate vrednosti Y i dodaje greška nasumično dobijenu iz raspodele $Y-\hat{Y}$. Ovo se zove „nasumično predviđanje“. Ni ovim rešenjem ne eliminiše se problem smanjenja standardne greške. Zato se predviđanje ponavlja nekoliko puta, pri čemu se svaki put generiše novi skup podataka sa odgovarajućim koeficijentima, koji su za svaki set različiti. Poenta je u uzimanju u obzir ovog varijabiliteta prilikom računanja konačnih ocena.

Postupak koji se primenjuje teorijski je relativno jednostavan. U prvom koraku se vrši predviđanje za ceo skup podataka, na osnovu ocena parametara dobijenih iz početnog raspoloživog skupa podataka. Tada se ocene parametara mogu dobiti tako što se prvo primeni potpuno ili delimično uklanjanje nekompletnih slogova ili upotrebom EM algoritma.

Kod multivariabilnog modela sa normalnom raspodelom, predviđanje se zasniva na regresiji atributa u kojima se pojavljuju nedostajuće vrednosti po ostalim atributima u skupu. Pojednostavljeno, ako se pretpostavi da se radi regresija atributa X po Z i da je standardna greška regresije $s_{x,z}$, predviđanje za X se može predstaviti formulom:

$$\hat{X}_i = b_0 + b_1 Z_i$$

Zbog povećanja veličine uzorka, u predviđanje se uvodi i slučajna greška:

$$\hat{X}_i = b_0 + b_1 Z_i + u_i s_{x,z}$$

gde je u_i slučajna vrednost standardne normalne raspodele. Ovim se u vrednost koja se predviđa dodaje neophodna doza neizvesnosti. Sledeći opisane korake predviđanja, uvek se dobija da vrednost koja se predviđa sadrži neku slučajnu grešku. I u svakom predviđanju će se dobiti pomalo različit rezultat.

Nasumičnost se koristi još u jednom koraku ovog procesa. Naime, koeficijenti regresije i standardna greška regresije se inače tretiraju kao parametri, a ustvari, oni predstavljaju ocene za posmatrani uzorak. Ocene parametara takođe imaju neku svoju raspodelu; na

primer, ako bi se posmatrali različiti skupovi podataka iz iste populacije, svaka od analiza bi dala drugu vrednost b_1 i mogla bi se odrediti raspodela za nju. Tako da će drugi korak biti da se uzme slučajni uzorak ovih ocena iz njihovih raspodela aposteriornih verovatnoća, dobijenih na osnovu podataka sa kojima se raspolaže.

Nakon što se izračunaju predviđanja za nedostajuće vrednosti, metoda višestrukog predviđanja se ponavlja u iteracijama, predviđaju se nove vrednosti, izvode se korigovane ocene za parametre, pa se ponovo predviđaju nedostajuće vrednosti, itd., dok se rešenje ne stabilizuje. Na kraju, pomoću konačnih ocena parametara, generiše se krajnji skup podataka sa vrednostima dobijenim predviđanjem.

Proces se ne završava ni nakon ovoga. Ide se dalje, generiše se još nekoliko skupova podataka. Praksa je pokazala da je u većini slučajeva dovoljno napraviti tri do pet skupova. Budući da se u algoritmu dodaje faktor nasumičnosti, ovi skupovi će se u određenoj meri međusobno razlikovati, a samim tim rezultati analiza koje se nad njima sprovode. Ocene parametara za konačnu jednačinu regresije se dobijaju uzimanjem proseka ocena dobijenih iz svih generisanih skupova.

U vreme ekspanzije informacija i njihovog sve većeg značaja, kvalitet podataka svakako igra jako važnu ulogu u savremenom poslovanju u svim oblastima privrede. Posledično i predviđanje nedostajućih vrednosti u podacima je domen koji se razvija. Kao odgovor na sve veće zahteve koji se javljaju, očekuje se da će budući rad na ovom polju doneti još bolje algoritme za predviđanje koji bi ujedno sa razvojem informacionih tehnologija mogli da funkcionišu i dovoljno precizno i brzo. U svakom slučaju, mogućnosti za dalji napredak ove oblasti su velike, a primena u praksi je sasvim izvesna.

U nastavku disertacije nakon ukaza na važnost podataka, sledi opis sistema poslovne inteligencije.

5. SISTEMI POSLOVNE INTELIGENCIJE

Kao što je prethodno istaknuto, Poslovna inteligencija (PI) predstavlja skup alata, metoda, modela i aplikacija koje se koriste za prikupljanje, skladištenje i analizu podataka, sve zarad kvalitetnog poslovnog zaključivanja i predloga jedinstvenog rešenja. Imajući u vidu činjenicu, da je želja autora da prikaže sistemski pristup u odlučivanju i poslovnoj inteligenciji, u nastavku ovog dela disertacije, prvo se ukazuje na kratak osvrt i istorijski razvoj PI i SPI.

5.1. Uvod u poslovnu inteligenciju

Najranija preteča poslovnoj inteligenciji su Management Information System-i (MIS), koji su jaku ekspanziju razvoja imali tokom 1960-tih i 1970-tih sa relativno lošim uspehom. Osim toga, otvorili su prostor za razvoj sistema za podršku odlučivanju i sistema poslovne inteligencije, pre svega prvih koncepata kao što su: data warehousing (DW), data mining (DM), izveštavanje, OLAP. Kao što je i logično, na početku su razvijeni kao teorijski koncepti koji su tražili svoje mesto u praksi.

Koncept sistema za podršku odlučivanju (DSS) evoluirao je iz dva glavna izvora istraživanja: teorijskog izučavanja organizacionog odlučivanja i tehničkih prodora u oblasti interaktivnih računarskih sistema. Smatra se da je DSS koncept iznikao kao posebna istraživačka oblast početkom 80-tih godina, prošlog veka. Tada se pojavljuju menadžerski informacioni sistemi (Executive Information Systems – EIS), grupni sistemi za podršku odlučivanju (Group Decision Support Systems – GDSS) i organizacioni sistemi za podršku odlučivanju (Organizational Decision Support Systems – ODSS). Smatra se da je značajno bolji komercijalni uspeh koji su postigli sistemi za podršku odlučivanju, a kasnije i sistemi poslovne inteligencije (u odnosu na koncept MIS-a), nastao pre svega kao rezultat shvatanja da je u centru procesa odlučivanja čovek, te da se podrška odlučivanju mora fokusirati na individualizaciju procesa donošenja nestrukturiranih odluka. Takođe, veliki problem MIS-a predstavljalo je kompleksno okruženje za dobavljanje podataka neophodnih za analize, koje su pretežno morali da koriste donosioci odluka. Kao alternativa u rešavanju dobro strukturiranih problema odlučivanja javlja se i koncept ekspertnih sistema (Expert System – ES), koji se oslanja na primenu veštačke inteligencije. Pojednostavljeno, za ekspertne sisteme se kaže da donose odluku "za" korisnika, dok sistemi za podršku odlučivanju donose odluku "sa" korisnikom, videti u [127].

Početak 1990-ih, data warehousing i on-line analytical processing, proširuju domen sistema za podršku odlučivanju. Naziv poslovna inteligencija je još 1978. godine Howard Dresner, upotrebio u svom članku "A Business Intelligence System", izdatom za IBM Journal. Pomenuti, g-din, Howard Dresner, istraživač u Gartner Group-i, je 1989. godine popularizovao termin poslovna inteligencija, kako bi opisao skup koncepata i metoda zasnovanih na kvantitativnim-informatičkim činjenicama. Iste treba da poboljšavaju proces poslovnog donošenja odluka, tj. iskoristio ga, po nekim mišljenjima, da opiše ono što je već bilo poznato kao sistem za podršku odlučivanju. Softver za poslovnu inteligenciju uključuje alate za "rudarenje" podataka (data mining), multidimenzionu analizu, predefinisano izveštavanje, itd., videti [128].

Za donošenje odgovarajućih poslovnih odluka zasnovanih na preciznim i ažurnim informacijama, jasno je da treba mnogo više od intuicije. Alati namenjeni analizi podataka, kreiranju izveštaja i upita, pomažu korisnicima da u moru podataka sakupe vredne informacije - danas ovi alati pripadaju kategoriji sa zajedničkim nazivom poslovna inteligencija.

Iz ove perspektive se čini da je sve teklo prirodnim tokom, ali sve većom brzinom. U nastavku je dat kratak pregled značajnih momenata u razvoju - poslovne inteligencije, detaljno u [118].:

- ✓ Prvi jezici - korisnički usmereni, pojavili su se kao most između krajnjih korisnika i IT okruženja, uspostavljajući koncept analize od strane korisnika.

- ✓ Informacioni centri kompetencije su kreirani sa ciljem unapređivanja sredstava namenjenih poboljšanju efikasnosti krajnjih korisnika. Potreba za uspostavljanjem korporacijskih standarda za analitičke alate je najznačajnija korist dobijena od informatičkih centara.

- ✓ Sa pojavom klijent/server sistema dolazi se do zaključka, da držanje podataka *u mestu* nije pogodno za analizu. Zato je reinžinjerung podataka u pomoćne PI forme i formate označen kao cilj kome se teži. Relacione baze su bile najprihvatljiviji oblik koji podržava SQL.

- ✓ *Information Warehouse* su dokazale da sakupljanje *meta podataka* o postojećim informacijama ima savršenog smisla. Pre nego što se izvrši transformaciju, neophodno je saznati što više o njihovom sadržaju i oblicima.

- ✓ *Data warehouse* donosi revolucionarni napredak i pruža nove mogućnosti. Dizajniran je da neprekidno zadovoljava potrebe poslovnih korisnika, a ne samo dnevne i operativne potrebe. Informacija postaje čista, jasna i konzistentna, bez potrebe da je tehničko

osoblje interpretira u obliku koji korisnik može da razume. Osim toga sadrži i istorijske informacije a ne samo informacije o tekućem poslovanju.

✓ Dolazi se do perioda kada se zahtevaju ugrađena, "real-time" PI rešenja. Ovde pokretačku silu predstavlja potreba za obezbeđivanjem sofisticiranih metrika i analiza za top menadžment.

Očigledno je da su skladišta podataka osnova PI. Sama po sebi, nameće se potreba za pronalaženjem znanja u tim velikim bazama podataka ili skladištima podataka. To je razlog da se u osnovi za poslovnu inteligenciju smatra da su njeni osnovni koncepti, pre svega *velike količine baza (skladišta) podataka i algoritmi za pronalaženje znanja u skladištima podataka.*

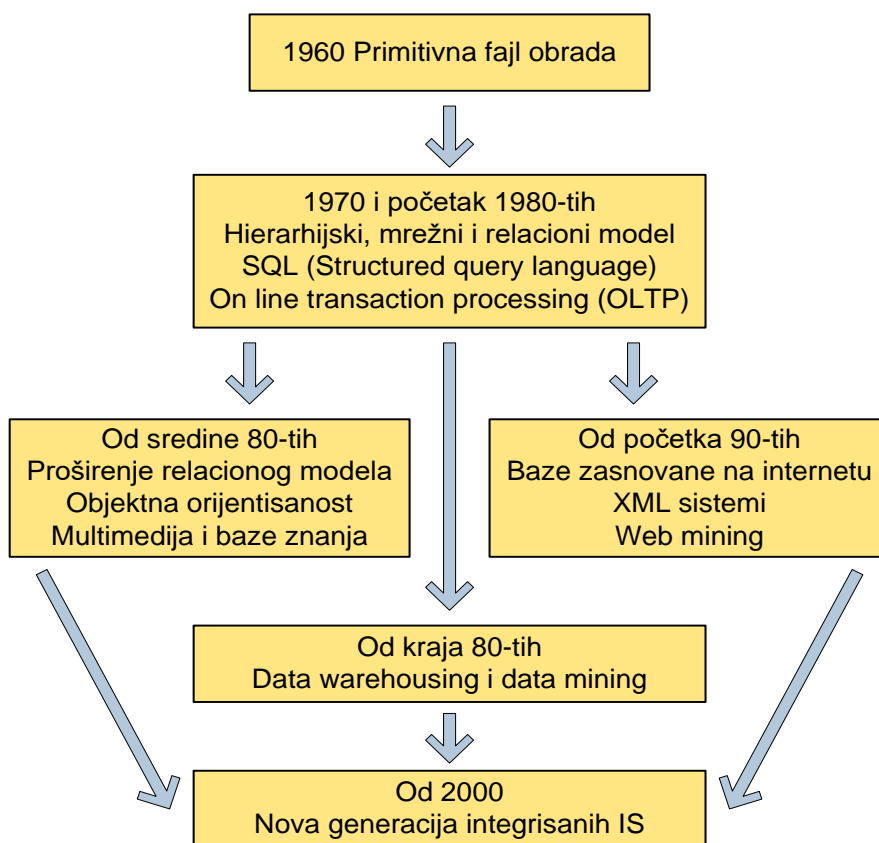
Danas se sa sigurnošću može reći da je prethodno stoleće obeležila pojava interneta i informacionih tehnologija. Sam razvoj je bio veoma dinamičan, ali kao i svi počeci imao je svoje poteškoće i probleme.

U pomenutoj informatičkoj eri, može se konstatovati da je koncept relacionih baza nadživeo mrežni i hijerarhijski model baza podataka. Još pre mnogo godina i relacionom modelu su se počele stavljati zamerke. Zamerke su bile uglavnom vezane za nemogućnost izvođenja sofisticiranih (složenih) analiza nad podacima. Kako je rasla upotreba informacionih tehnologija u poslovnom svetu sve više se razmišljalo o kreiranju specijalnih alata za analizu podataka. Tokom godina iskristalisala su se dva osnovna pravca: *data warehousing* i *data mining*. Novi koncepti su uspeli da daju zadovoljavajuće rezultate u praksi, ali je pitanje da li će njihov razvoj moći da odgovori na sve složenije zahteve korisnika.

Ukratko hronološki posmatrano, početkom pedesetih godina prošlog stoleća počela je primena računara. Računari su u to vreme upotrebljavani u naučne svrhe. Početkom šezdesetih godina počela je poslovna primena računara. To je vreme u kome su baze podataka bile u "povoju", uglavnom skladištene u tekstualne datoteke, ili jedostavne binarne fajlove. Takav način obrade podataka nije bio zadovoljavajući, ali je bio revolucionaran u odnosu na "olovku i hartiju" (Slika 5.1.). Mnogo veći problem je bilo nepostojanje odgovarajuće metodologije i manjak standarda u radu sa bazama podataka. Sedamdesete godine donele su mnogo toga, i mogu se smatrati kao prekretnica u razvoju baza podataka. Razvijeni su mrežni, hijerarhijski i relacioni model baza podataka. Na osnovu prethodnih grešaka i loših iskustava nagomilano iskustvo u radu sa bazama podataka pomoglo je da se uspostave standardi i metodološki koncepti u vezi načina projektovanja baza. Jedni od najpoznatijih su dijagrami objekti i veze, kao i stuktorna sistemska analiza.

Tokom osamdesetih godina koncept relacionih baza podataka je preovladao tako da su drugi koncepti “pali u zaborav”. Razvoj se dalje nastavio u pravcu objektnih i objektno-relacionih baza podataka, uvođenja multimedija i novih alata za razvoj i prezentaciju baza na Webu.

Već od osamdesetih godina, korisnici i projektanti su postali svesni da baza podataka ne treba da bude samo baza podataka nego i baza informacija. Ključni problem sa bazama podataka nije više optimalan dizajn baze, bekapovanje podataka, sigurnost podataka, nego kako upotrebiti te nagomilane podatke i izvući iz njih dragocene informacije i znanje, koje će poboljšati poslovne rezultate.



Slika 5.1. Evolucija baza podataka

Tokom godina razvile su se ideje pomenutih koncepata DW i DM. U početku su razvijeni kao teorijski koncepti koji su tražili svoje mesto u praksi. Tako je problem izvlačenja informacija iz baza podataka postajao sve značajniji, da su kompanije koje su lideri u razvoju baza podataka počele da isporučuju i alate za analizu podataka. Najpoznatiji primeri takvih baza podataka su DB2, ORACLE i MS SQL Server.

U novije vreme govori se o integralnim informacionim sistemima koji će sadržavati implementirane sve funkcionalosti, ali praksa će pokazati da li će te optimističke težnje biti potvrđene u praksi.

Karakteristike transakcione obrade podataka

Iz evolucije baza podataka može se zaključiti da su prvo razvijeni OLTP sistemi. Zbog rasprostranjenosti relacionih baza podataka u daljnjem izlaganju pod OLTP sistemima podrazumevaćemo upravo relacione sisteme.

Osnovne karakteristike OLTP sistema su:

- ✓ Visok stepen normalizacije,
- ✓ Konkurentna obrada podataka,
- ✓ Obezbeđuje integritet podataka,
- ✓ Visoka konzistentnost podataka,
- ✓ Namenjen je ažuriranju podataka, itd.

OLTP sistemi su dizajnirani u skladu sa pravilima normalizacije, da bi se izbegle anomalije u ažuriranju podataka. OLTP sisteme koristi veoma mnogo korisnika istovremeno, tako da ti sistemi manje ili više podržavaju konkurentnost obrade pod kojom se podrazumeva da više korisnika može istovremeno da radi sa istim podacim kao i da ih ažurira. Prilikom konkurentnog rada više korisnika ponekad dolazi i do neželjenih situacija, kao što su “dead lockovi”.

OLTP sistem kao visoko normalizovan, vrlo je pogodan za brza ažuriranja podataka, dok odabiranje i prebacivanje podataka za kompleksne izveštaje može da uzme mnogo procesorskog vremena i da naruši performanse baze. Upravo zbog navedenog razloga slede nedostaci.

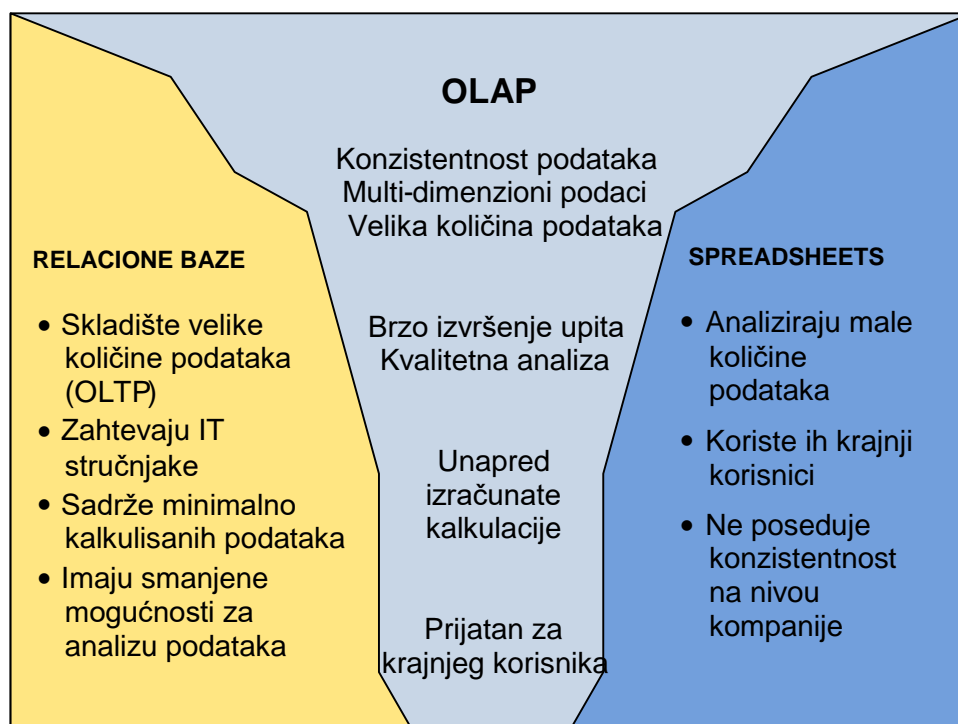
OLTP nedostaci:

- ✓ Nemogućnost kreiranja kompleksnih upita,
- ✓ Korisnici moraju da poseduju odgovarajuća predznanja kako bi mogli da formiraju proizvoljne korisnički definisane upite,
- ✓ Pri izvršavanju kompleksnih upita konkurentna obrada podataka je ugrožena,
- ✓ Upiti se sporo izvršavaju jer je potrebno mnogo spajanja tabela da bi se dobio rezultat, itd.

OLTP baze podataka uglavnom koriste SQL (Structured Query Language) jezik upita koji je namenjen za relacione baze podataka. Iako je tokom godina isti razvijen i unapređivan, ipak nije dovoljno jednostavan da bi se koristio u kompleksnim analizama.

Dodatni problem je što implementatori IS nikad ne mogu da predvide sve upite koji će biti potrebni, tako da korisnicima uvek nedostaju kvalitetni izveštaji. Korisnik OLTP sistema da bi mogao da postavlja korisnički definisane upite ako ima odgovarajuća predznanja iz relacionih baza podataka i poznaje SQL jezik, a to tražiti od običnog korisnika nije realno. Osim toga, izvršenje upita je sporo zato što u upitima učestvuju mnogo tabela, a njihovo spajanje (join) zahteva određeno procesorsko vreme, a naročito je zahtevno kada je reč o tabelama sa mnogo slogova, videti [128].

Jedna moguća varijanta rešenja pomenutih problema, jeste upotreba OLAP rešenja, koje će biti objašnjeno u nastavku.



Slika 5.2. Rešenje je u OLAPu

Karakteristike analitičke obrade podataka

Pored stalne (transakcione) obrade podataka, sa vremenskim trendom nagomilavanja podataka, javlja se potreba za analizom podataka. Osnovne karakteristike ove analize su sledeće:

- ✓ Visok stepen denormalizacije podataka,
- ✓ Podaci su namenjeni za čitanje,
- ✓ Obezbeđena je konzistentnost podataka,
- ✓ Podaci su namenjeni za analizu, itd.

U OLAP sistemima se teži visokom stepenu denormalizacije, *naprosto* da bi se lakše formirali upiti. U OLAP bazama podataka akcenat se stavlja na čitanje i njihovo pregrupisanje da bi se iz podataka protumačile informacije. Konzistentnost podataka je i ovde veoma bitna karakteristika jer, ako podaci nisu konzistentni bilo koja informacija proistekla iz njih je pogrešna. Osim toga ovi sistemi se ne bave obradom podataka nego samo tumačenjem i analizom.

Relacione baze podataka su istovremeno i OLTP i OLAP sistemi, jer ne postoji relacioni sistem koji istovremeno ne služi u obe svrhe, dok u svakom sistemu prevlađuje jedna od "važnijih" uloga pomenuta dva sistema.

Prednosti OLAPa su:

- ✓ Izuzetne performanse u izvršenju kompleksnih upita,
- ✓ Ne ugrožava konkurentnost obrade,
- ✓ Ne zavisi od DBMS,
- ✓ Može da objedinjuje podatke iz različitih DBMS,
- ✓ Posедуje jezik specijalno dizajniran za analizu podataka, itd.

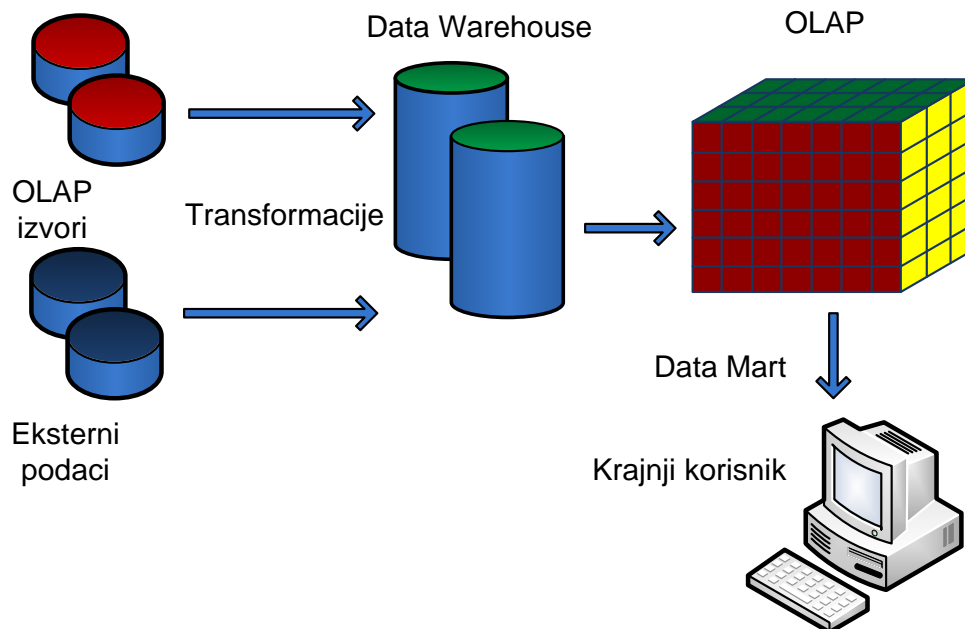
Pravi OLAP sistemi sadrže kopije podataka iz OLTP sistema, tako da ne ugrožavaju konkurentnost obrade, jer se nad istim podacima ne vrše operacije ažuriranja i kreiranja izveštaja. OLAP sistem može da preuzme podatke iz bilo kog relacionog ili ne relacionog izvora podataka jer on podatke skladišti na "svoj" način, i nije preterano bitna vrsta izvora podataka. Ova osobina omogućena je zahvaljujući činjenici da je moguće uzimanje podataka iz različitih DBMS i objedinjavanje u zajednički OLAP sistem. Način skladištenja podataka je optimizovan u cilju kreiranja veoma kompleksnih izveštaja. OLAP poseduje i veoma moćan jezik upita za kreiranje multi-dimenzionih upita, koji poseduje mnogo više mogućnosti nego SQL jezik.

Poređenje ova dva sistema:

OLTP	OLAP
Obrada podataka,	Čitanje podataka,
Mnogo kratkih transakcija,	Dugački i kompleksni upiti,
Mb-Gb podataka,	Gb-Tb podataka,
Sirovi podaci,	Sumarni i prerađeni podaci.
Mnoštvo različitih korisnika,	Korisnici su DO i analitičari,
Sadrži ažurne podatke, itd.	Sadrži istorijske podatke, itd.

Najčešći korisnici OLTP sistema su operateri, kontrolori, programeri, itd., dok su korisnici OLAP sistema uvek analitičari ili članovi menadžmenta na najvišem nivou. U svakom slučaju reč je o osobama koje nisu informatički u potpunosti obrazovane, pa je sama ideja OLAPa prilagođena njima. Korisnici OLAPa ne moraju i (uglavnom) nemaju visoka znanja iz informacionih tehnologija nego su stručnjaci isključivo u oblasti u kojoj se bave.

Na Slici 5.3., prikazan je modelski koncept projektovanja skladišta podataka, koji se detaljno objašnjava na realnom primeru.



Slika 5.3. Razvoj skladišta podataka

Drugo, OLTP baze uvek su orijentisane na prošlost i sadašnjost i one u svakom trenutku imaju ažurno stanje svih podataka u bazi, dok OLAP baze sadrže istorijske podatke i povremeno po unapred ustaljenoj proceduri preuzimaju automatski podatke iz OLTP baze

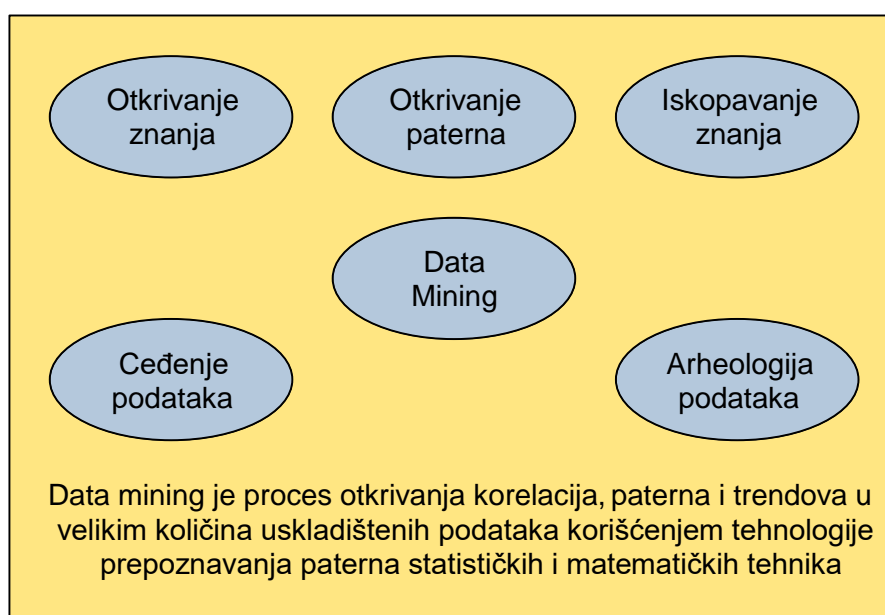
podataka. Uglavnom se koriste za analizu istorijskih podataka kako bi se otkrile zakonitosti u prošlosti i na osnovu njih mogla izraditi planovi za naredni vremenski period. Nakon osnovnog opisa skladišta podataka, detaljan opis istog sledi u Poglavlju 5.3.

Već je istaknuto da je drugi osnovni elemenat poslovne inteligencije, pronalaženje znanja u prethodno pomenutim podacima, te se u nastavku posvećuje pažnja ovoj oblasti, prvo na teoretskoj osnovi a potom u nastavku i na aplikativnom nivou.

Pronalaženje znanja u podacima je proces kreiranja najrazličitijih upita i ekstrakcija korisnih informacija, uzora i trendova prethodno nepoznatih, sadržanih u velikim bazama podataka.

Sastoji se od skupa tehnika za analizu podataka, čiji je cilj da u podacima pronađe određene zavisnosti, veze i pravila vezana za podatke i isti protumači u novi, viši nivo kvalitetne informacije. Za razliku od skladišta podataka koji ima jedinstven prilaz podacima, otkrivanje zakonitosti u podacima, daje rezultate koji predstavljaju veze i zavisnosti između podataka, koji se ne bi mogli otkriti na drugi način, npr. pomoću SQL upita ili prostim posmatranjem podataka.

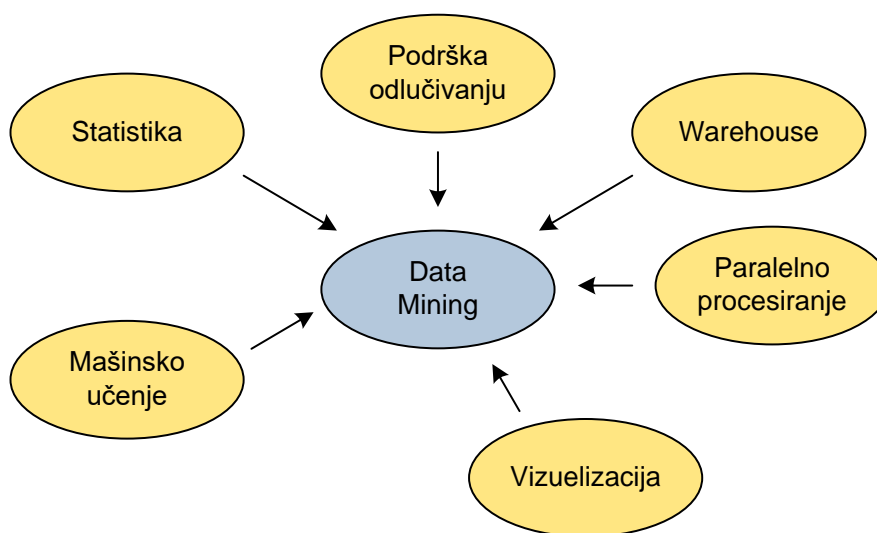
Otkrivanje znanja u podacima se često susreće pod različitim nazivima. Na Slici 5.4., prikazani su najčešći nazivi koji se koriste u literaturi, kao i jedna od najpotpunijih definicija vezana za otkrivanje znanja u podacima, videti u [128].



Slika 5.4. Sinonimi za otkrivanje znanja u podacima

Otkrivanje znanja u podacima predstavlja integraciju više tehnologija kao što je prikazano na prethodnoj slici. Njime je obuhvaćeno i upravljanje podacima kao što je upravljanje bazama, skladišta podataka, statistika, podrška odlučivanju, mašinsko učenje, vizuelizacija, itd. U istraživanjima otkrivanja znanja u podacima se koriste znanja iz mnogih oblasti i disciplina. Tako npr. skladišta podataka, kao jedna od ključnih tehnologija analize podataka, integriše različite izvore podataka i organizuje ih radi efikasnije analize.

Proces otkrivanja znanja u podacima se sastoji od nekoliko važnih koraka. Ti koraci obuhvataju organizovanje podataka za istraživanje, određivanje željenog rezultata, izbor alata za istraživanje, izvodjenje istraživanja nad podacima, selekciju rezultata kako bi se odvojili oni korisni, preduzimanje konkretnih akcija, i evaluacija akcija kako bi se izdvojilo ono što je korisno, videti [35].

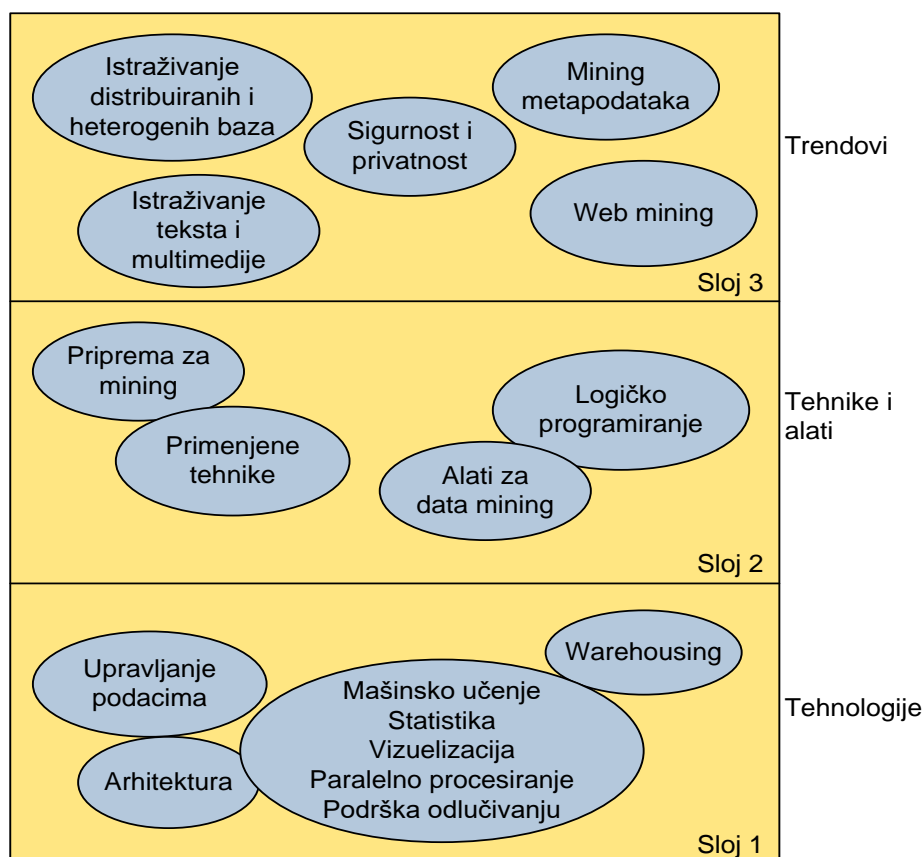


Slika 5.5. Uticaj raznih disciplina na DM

Različite mining tehnike se koriste za sređivanje i struktruiranje tih podataka. Zatim se koriste različiti DM alati koji operišu sa ovako strukturiranim podacima. Ipak, znatan broj podataka je nestrukturiran. Takvi su na primer, podaci u multimedijalnim bazama. Za njih je potrebno razviti odgovarajuće data mining alate. Isto tako, podaci do kojih se dolazi preko interneta su mnogobrojni, pa je neophodno razviti i alate za ekstrahovanje bitnih podataka.

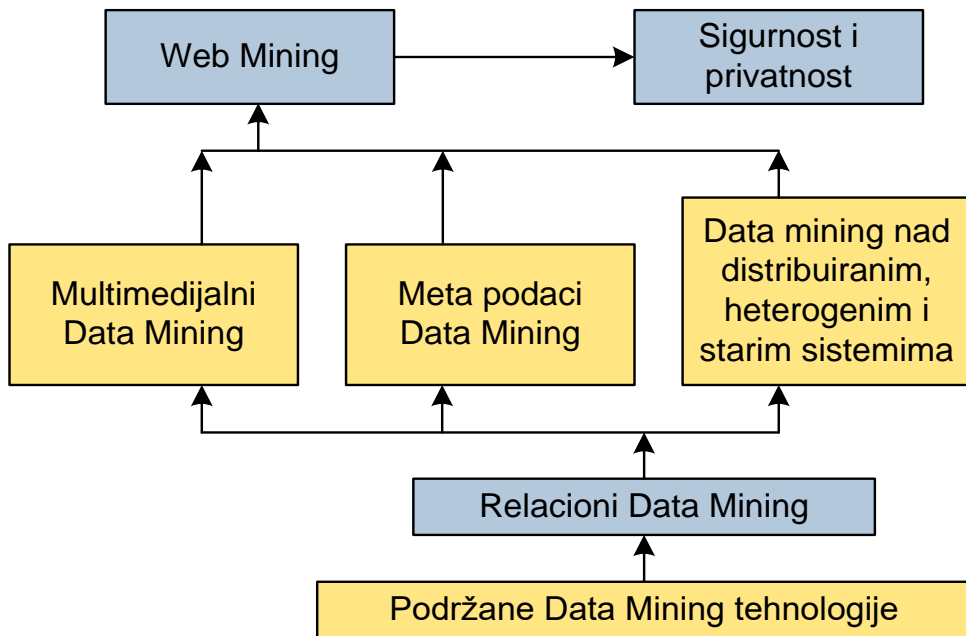
Na sledećoj slici prikazana je evolucija sistema pronalaženja znanja u podacima. Na *prvom* nivou nalaze se tehnologije koje se koriste: projektovanje baza i upravljanje podacima, skladište podataka, mašinsko učenje, statistika, vizualizacija, paralelno procesiranje i sistemi za podršku odlučivanju. Na *drugom* nivou predstavljene su tehnike i alati: priprema za otkrivanje znanja u podacima, alati za otkrivanje znanja u podacima, istraživanje podataka i

logičko programiranje. Na *trećem* nivou prikazani su trendovi: istraživanje distribuiranih i heterogenih baza, tekstulanih i multimedijalnih baza, sigurnosti, privatnosti i WWWa.



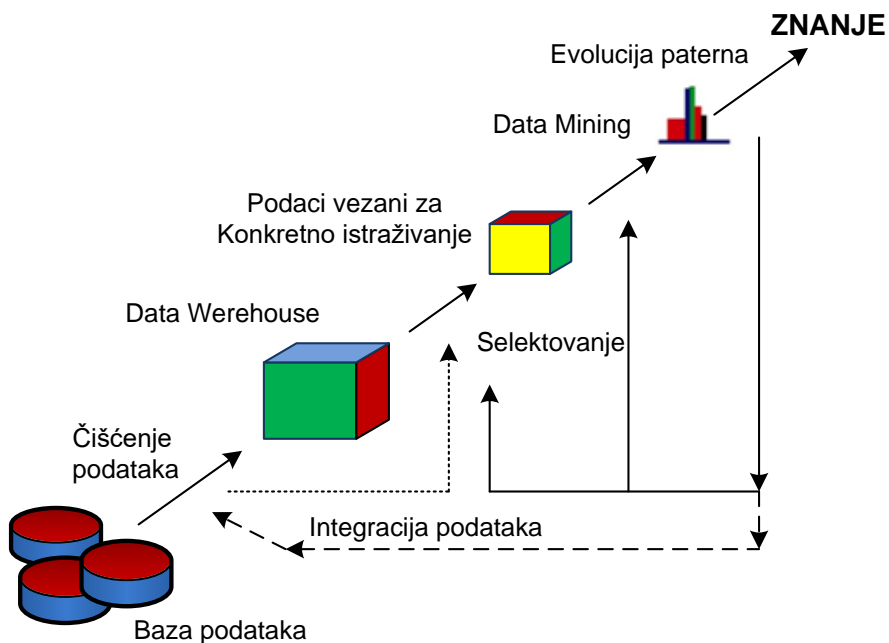
Slika 5.6. Evolucija algoritama za otkrivanje znaja

Na sledećoj slici prikazani su pravci razvoja. U početku, otkrivanje znanja u podacima je bilo bazirano isključivo na relacione izvore podataka. Kasnije se njegova upotreba proširila nad multimedijalnim, metapodacima, distribuiranim, heterogenim i starim bazama podataka. Najnoviji trendovi su zasnovani na WEB miningu, sigurnosti i privatnosti na internetu.



Slika 5.7. Pravci razvoja DM

Očigledno je da je rezultat algoritama za otkrivanje znanja, *pre svega otkriveno i protumačeno znanje za posmatranu oblast, što predstavlja osnovu kvalitetnog odlučivanja*. Na sledećoj slici, (jedna slika hiljadu reči) ukazuje se na evoluciju znanja, nastalu kao simbiozu oblasti skladišta podataka i oblasti otkrivanja znanja u podacima.



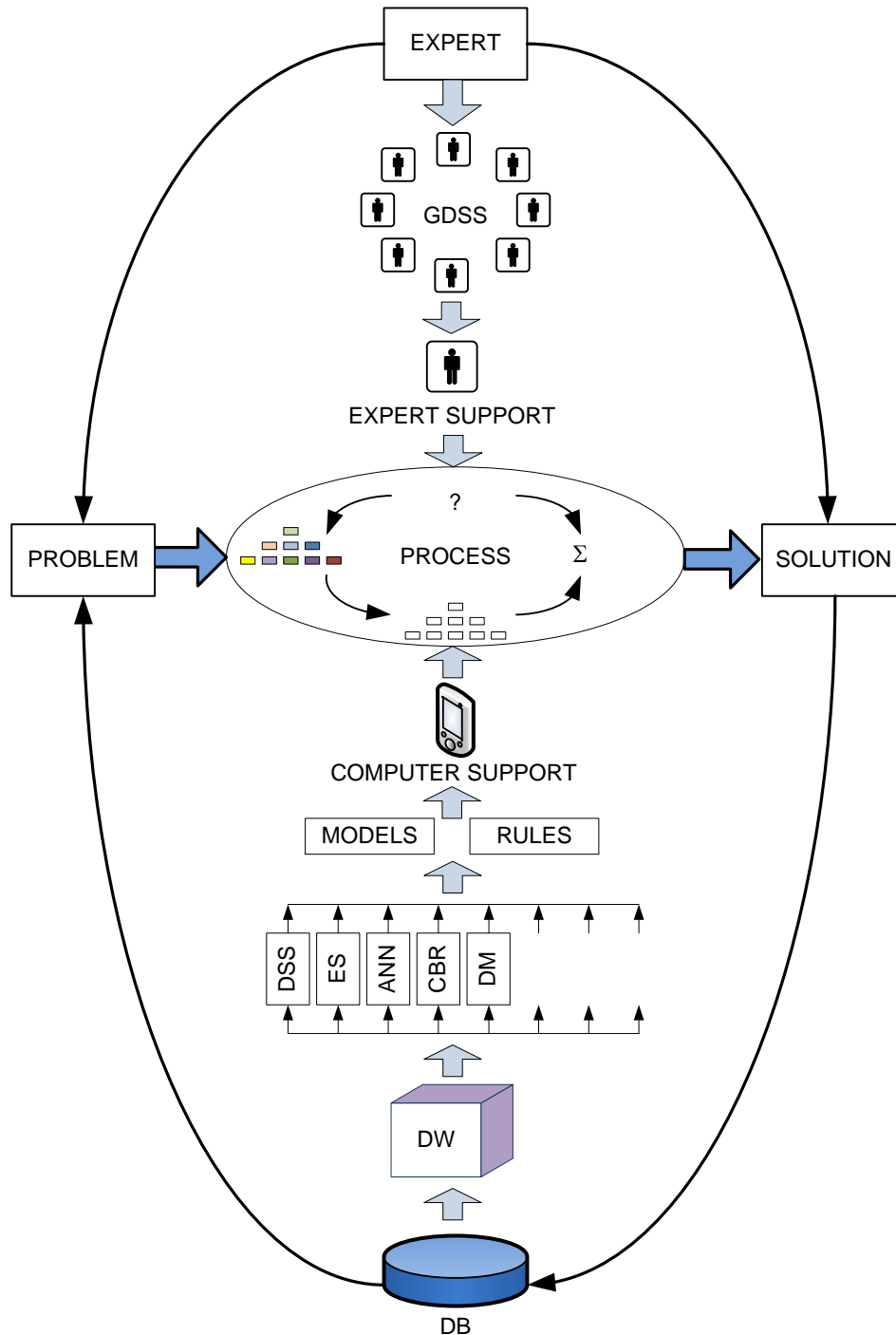
Slika 5.8. Proces otkrivanja znanja

U nastavku sledi proširenje pomenutog osnovnog modela poslovne inteligencije u model sistema poslovne inteligencije.

5.2. Sistemi poslovne inteligencije

Nakon prikaza osnovnih koncepata poslovne inteligencije u nastavku se ukazuje na Sisteme poslovne inteligencije (SPI). Praktično prethodno opisani osnovni model sa elementima skladišta podataka i elementima otkrivanja zakonitosti u podacima usložnjava se kako sa nekim preostalim alatima i tehničkom elementima sa jedne strane, tako i sa ljudskom komponentom sa druge strane. Posebno ističem, da ljudska komponenta može da bude na nivou pojedinca kao što je, analitičar, inženjer znanja, menadžer ili donosioc odluke. Sve pomenute četiri profesije mogu da budu u potpunosti razdvojene i da ih nose sa sobom upravo pomenute bezlično četiri osobe, a može da bude i sinonim za jednu osobu.

Osim pojedinačnog (individualnaog) pristupa, u sistemima poslovne inteligencije, posebno mesto zauzima tim ili grupa eksperata izvesne oblasti. Upravo, u okviru disertacije, poseban akcenat će biti posvećen pomenutom timu, procesu *prikupljanja, ažuriranja i distribucije znanja*, sve sa ciljem predloga najprihvatljivijeg rešenja i izbora adekvatne poslovne odluke. *Jer, nedvosmisleno (jednoznačno) se potvrđuje da je znanje ključni resur za donošenje ispravnih upravljačkih odluka.* Upravo na sledećoj slici, prikazuje se jedan integralni pristup sistemu poslovne inteligencije, sa svim elementima i vezama između elemenata posmatranog sistema.



Slika 5.9. Sistem poslovne inteligencije, [127]

U nastavku se detaljnije daje objašnjenje prikazanog sistema poslovne inteligencije. Za svaki poslovni problem, gradi se njegov adekvatan model za rešavanje. Što je model više verna slika realnog problema, to su izlazi iz modela verodostojniji i mogu biti dobra osnova za rešavanje pomenutog realnog problema. Profesor Petrić navodi, *Ako je model verna slika realnog problema, to je 50% realnog rešenja.*

Sa prethodne slike uočava se centralno mesto, a to je da se za posmatrani problem, grade modelske procedure za njegovo rešavanje u okviru samog procesa rešavanja, sa željom analitičara, inženjera znanja, da se dođe do rešenja, samim tim i do predloga odluke. Sa donje strane slike, ispod procesa rešavanja problema, uočava se pre svega set tehničkih komponenti modela sistema poslovne inteligencije.

Na najnižem konceptualnom nivou, nalazi se pre svega baza podataka, u kojoj se beleže svi relevantni podaci o prirodi problema koji se rešava. Narastanjem poslovnih podataka iz jedne u drugu poslovnu godinu, stvaraju se preduslovi za razvoj i projektovanje adekvatnog skladišta podataka, dobro predstavljenog dela sistema poslovne inteligencije u prethodnom poglavlju. Kao što se vidi sa slike, na pomenuto skladište podataka, naslanja se čitava lepeza alata za podršku procesu donošenja odluka. Od sistema za podršku odlučivanju, ekspertnih sistema, veštačkih neuronskih mreža, genetskih algoritama, zaključivanja na osnovu slučajeva, itd. *Ono što je zajednički imenilac pomenutih alata, a to je da oni podpomažu procese donošenja odluka a ne nikako zamenjuju donosioca odluke.* Osim toga, za sve pomenute alate, zajednički imenilac jeste visok stepen razvijenosti, odnosno pre svega dobro potpora informaciono komunikacionih komponenti.

Sledeći nivo, opisuje u kom je obliku moguće modelovati podatke kao izvorne promenljive, zarad dobijanja izvedenih promenljivih odnosno, *znanja*. Kao što je poznato, to su dva pravca i to: kvantitativni ili modelski i kvalitativni ili u obliku pravila odlučivanja.

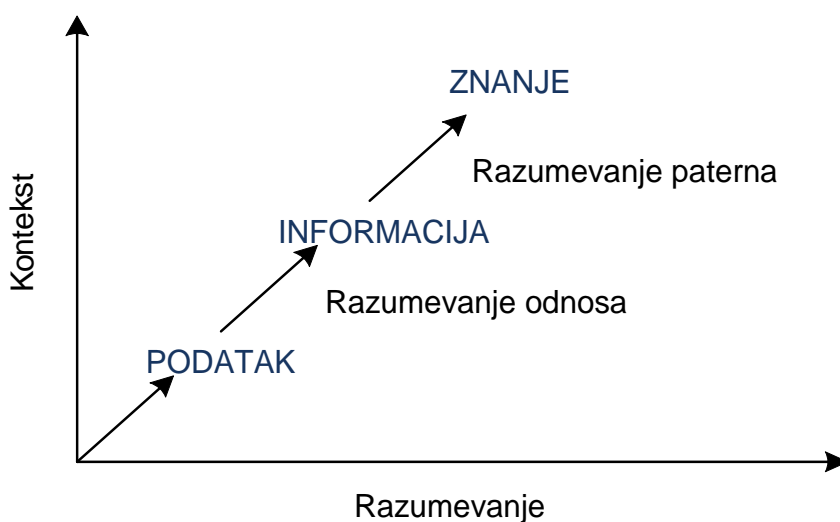
Kao primer, navodi se slučaj, ako je moguće na postojeće podatke, primeniti adekvatan matematički model, moguće je projektovati alat za podršku odlučivanju iz klase sistema za podršku odlučivanju, detaljno u [127]. Osnovne komponente svakoga sistema za podršku odlučivanju su baza podataka, baza modela i korisnički interfejs. Praktično, u ovom slučaju moguće je projektovati adekvatnu bazu modela. Kao primer, moza se navesti, baza modela upravljanja zalihama, sa projektovanim i razvijenim sistemom za upravljanje zalihama u posmatranom poslovnom sistemu.

Sa druge strane, ako podaci nisu kvantitativni, već kvalitativni, stvara se preduslov za razvoj drugog alata SPI, a to su ekspertni sistemi. Njegove osnovne komponente su baza podataka, baza znanja i korisnički interfejs. Upravo pomenuta baza znanja, implementirana najčešće sistemom produkcionih pravila, predstavlja dobru osnovu za modelovanje sa kvalitativnim podacima. Kao primer, može se navesti, baza znanja o kreditnoj sposobnosti ili odlivu i prilivu klijenata u finansijskom sektoru. Sam ekspertni sistem, može biti, ekspertni sistem za podršku odlučivanju o odlivu klijenata, videti u [140].

Kao što se dalje sa slike vidi, u vidu kompjuterske podrške predočavaju se informacije i znaja o problemskoj oblasti, bilo ekspertu ili timu, koji je angažovan u procesu rešavanja poslovnog problema. Sa gornje strane procesa rešavanja problema, egzistira pojedinac (ekspert) ili tim (grupa) za rešavanje problema.

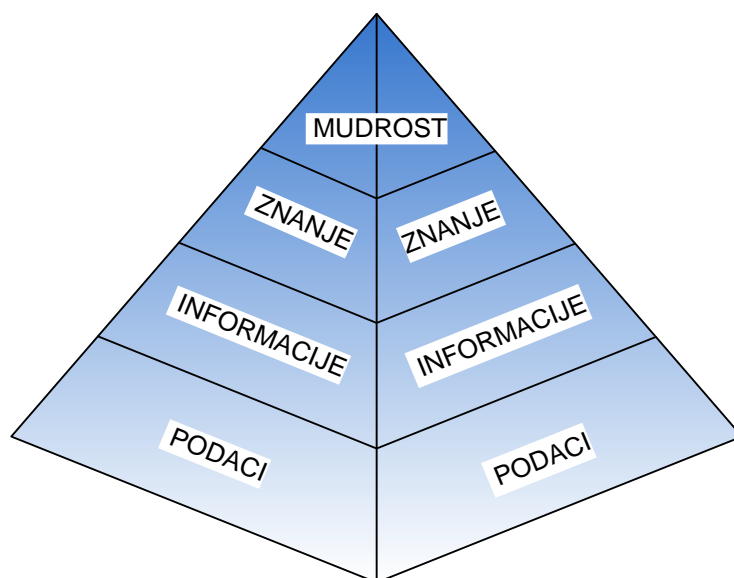
Ukoliko se radi o jednostavnijim poslovnim odlukama, one mogu da budu prepuštene individualnom donosiocu odluke, odnosno pojedincu. Ipak, takvih odluka u realnoj praksi danas je sve manje. Naime, prevlađuje stanje donošenja složenih, sofisticiranih odluke, čije su podloge za donošenje, pre svega nagomilani kvantum znanja. To je i osnovni razlog da se uključi tim ili grupa za njihovo donošenje, jer je već istaknuto da je *zbir znanja članova tima uvek po definiciji veći od znanja bilo kog pojedinca*.

Na sledećoj slici, iako možda sa ponavljanjem, prikazuje se koncept tumačenja podataka, prevođenja u informacije i napokon dobijanje i dolazak do znanja.



Slika 5.10. Veza između konteksta i razumevanja znanja

Ili u formi piramide, znanje može biti predstavljeno i sledećom slikom:



Slika 5.11. Piramida znanja

Podatak je osnovni činilac piramide znanja. Podatak predstavlja osnovnu gradivnu jedinicu znanja. Podatak sam za sebe ne znači mnogo, ali kada se stavi u kontekst može da dobije značenje. Kada se utvrde neke relacije tog podatka sa stvarnim problemom, tada se dobija informacija. Informacija je podatak koji je interpretiran, to je podatak koji ima određeni smisao za određenu osobu. Znanje je informacija sa dodatkom akcije koja omogućava da se nešto uradi. Mudrost predstavlja najviši nivo znanja, posedovanje vizije, zdravog razuma, predviđanja i vizualizacije. Mudrost se stiče iskusvom i shvatanjem zakonitosti u određenoj oblasti.

Mudrost se približava idealnoj situaciji da za svaki problem u određenom kontekstu odredi najbolje rešenje. U nastavku disertacije, biće posvećeno poglavlje resursu znanja, preko svih bitnih aspekata u formi menadžmenta znanja.

U ovom delu u nastavku ukazuje se na osnovne korake projektovanja skladišta podataka, kao osnove SPI, mesta odakle se uzima podatak, tumači u formi informacije i dobija znanje, sve zarad donošenja kvalitetne poslovne odluke. Odluke koja treba da rezultuje ostvarenjem postavljenog definisanog (željenog) cilja.

5.3. Koraci projektovanja skladišta podataka

Može se konstatovati da se u poslednje vreme intenzivno razvija proces skladištenja podataka, kao osnovni koncept sistema poslovne inteligencije. Koncept koji donosi ideju aktivnog pronalaženja i pružanja informacija potrebnih u procesu savremenog odlučivanja. Na temeljima informacionih tehnologija želi se postići “inteligentno” poslovanje preduzeća u kompleksnim tržišnim uslovima, kroz obezbeđivanje dodatnih, neotkrivenih, neopaženih znanja i mogućnosti.

Sisteme skladišta podataka, koriste donosioci odluka da analiziraju stanje i razvoj poslovnog sistema. Ovi sistemi se zasnivaju na velikoj količini podataka integrisanih iz heterogenih izvora u multidimenzionalne šeme koje su optimizovane za pristup podacima na način koji odgovara analitičarima. Uopšteno govoreći, multidimenzionalna šema se sastoji od činjenica, mera i dimenzija. Činjenice su fokus interesovanja za proces donošenja odluka i mogu se pratiti kroz mere i dimenzije. Mera je numeričko svojstvo činjenice i opisuje jedan od njenih kvantitativnih aspekata koji su od interesa za analizu (na primer, broj saobraćajnih nezgoda, broj evidentiranih lakših povreda i slično), a dimenzije predstavljaju kontekst za analizu ovih mera (na primer, mesto, vreme, teretno vozilo, pešak, itd.). Navodi se primer o saobraćajnim nezgodama, jer će isti biti analiziran u aplikativnom delu disertacije.

Sistemi skladišta podataka imaju sledeće važne specifičnosti, videti u [10]:

- ✓ Sistem skladišta podataka se oslanja na operativne baze podataka koje predstavljaju izvore podataka,
- ✓ Korisnički zahtevi se teško prikupljaju i obično menjaju u toku projekta,
- ✓ Projekti skladišta podataka su najčešće veliki projekti: prosečno vreme za njihovu izgradnju je veće od 12 meseci,
- ✓ Menadžeri, donosioci odluka su zahtevni korisnici koji traže pouzdane rezultate u vremenu koje odgovara poslovnim potrebama, itd.

Skladište podataka je priznato kao jedno od najsloženijih modula informacionog sistema i njegov dizajn i održavanje karakteriše nekoliko faktora kompleksnosti koji su, u ranim fazama ove discipline prouzrokovali visok procenat neuspeha sprovedenih projekata. Lista rizika sa kojima se mogu suočiti projekti skladišta podataka je duga. Neki od najčešće isticanih rizika su sledeći, videti [82]:

- ✓ Nepostojanje posvećenosti menadžmenta,
- ✓ Nerazumevanje šta je organizaciji potrebno od skladišta podataka i kako skladište podataka može da pomogne u postizanju strateških ciljeva organizacije,
- ✓ Neuključenost korisnika u proces prikupljanja zahteva i dizajna skladišta podataka,

- ✓ Nedovoljno posvećena pažnja kvalitetu podataka kojima se puni skladište podataka, Ako korisnici nemaju poverenje u odgovore koje dobijaju pomoću skladišta podataka, oni ga neće ni koristiti,
- ✓ Neadekvatno finansiranje,
- ✓ Poslovni sukobi između i unutar organizacionih jedinica,
- ✓ Nepostojanje ljudi koji zaista poznaju i razumeju izvorne sisteme,
- ✓ Loše upravljanje projektima,
- ✓ Nerealna očekivanja korisnika,
- ✓ Nepostojanje sklada između želja korisnika i podataka kojima raspolažu u svojim izvornim sistemima,
- ✓ Nedostupnost zaposlenih i krajnjih korisnika kada zatrebaju razvojnom timu,
- ✓ Promenljivost zahteva i promenljivost prioriteta,
- ✓ Nepostojanje procedura (postupak) za rešavanje spornih situacija, itd.

Zbog svoje specifičnosti, razvoj sistema skladišta podataka je naročito složen i zahteva specifične metodologije i odgovarajući životni ciklus. Svest o kritičnoj prirodi problema i iskustvo koje su praktičari stekli doveli su do razvoja različitih metodologija projektovanja i usvajanja odgovarajućih životnih ciklusa koji bi povećali verovatnoću završetka projekta i ispunjenja zahteva korisnika. Analiza zahteva i konceptualni dizajn su u velikoj meri odgovorni za uspeh projekata skladištenja podataka, jer bi u toku ove dve faze trebalo u potpunosti definisati multidimenzionalne šeme. Dok su delimično bili zanemareni u prvoj eri skladištenja podataka, ovi koraci su u poslednjih deset godina dobili veću pažnju.

Faza *analize zahteva* je jedan od prvih koraka u razvoju sistema skladišta podataka i stoga dovodi do značajnih problema ako se sprovodi na pogrešan način ili nekompletno. Analizi zahteva treba posvetiti posebnu pažnju i ona treba da bude sveobuhvatno podržana efikasnim metodama. Međutim, u razvoju skladišta podataka ovoj fazi nije posvećeno mnogo pažnje. Pored toga, raznovrsnost postojećih pristupa specifikaciji zahteva dovela je do situacije da veliki broj projekata skladištenja podataka, preskače ovu fazu, koncentrišući se na tehnička pitanja, kao što su modelovanje baze podataka ili performanse upita. Kao posledica toga, velike su šanse da rezultati projekta ne zadovolje potrebe korisnika i ne obezbede očekivanu podršku procesu donošenja odluka.

Specifikacija zahteva utvrđuje, između ostalog, koji podaci bi trebalo da budu dostupni i na koji način bi trebalo da budu organizovani, kao i koji su upiti od interesa. Ova faza treba da vodi dizajnera da otkrije bitne elemente multidimenzionalne šeme, odnosno činjenice sa njihovim merama, dimenzije i hijerarhije, koje su potrebne da se olakšaju buduće manipulacije podacima i izračunavanja. Faza analize zahteva postavlja temelj za sve buduće

aktivnosti u razvoju skladišta podataka. Ona ima veliki uticaj na uspeh projekta skladištenja podataka, jer direktno utiče na tehničke aspekte, kao i strukturu i aplikacije skladišta podataka.

Faza analize zahteva treba da se sprovodi zajedno sa fazom *konceptualnog dizajna*. Ove dve faze se u stvarnosti često preklapaju. U mnogim slučajevima, čim se neki inicijalni zahtevi dokumentuju, inicijalna konceptualna šema počinje da se oblikuje. *Kako zahtevi postaju kompletniji, tako se kompletira i konceptualna šema.*

Literatura predlaže nekoliko originalnih pristupa konceptualnom modelovanju u oblasti skladištenja podataka. Međutim, predložena rešenja nisu uvek u sprezi sa odgovarajućom tehnikom za analizu zahteva da bi se formirao metodološki pristup koji obezbeđuje da rezultujuća baza bude dobro dokumentovana i da u potpunosti zadovoljava zahteve korisnika. Specifičnosti skladišta podataka čine da ova dva koraka budu još više povezana nego u tradicionalnim sistemima baza podataka. Zapravo, nedostatak definisanih korisničkih zahteva i utvrđenih izvora operativnih podataka, koji određuju skup dostupnih informacija dovodi do toga da je teško razviti odgovarajuće multidimenzionalne šeme, koje, s jedne strane, ispunjavaju zahteve korisnika i s druge strane, mogu da se napune podacima iz operativnih izvora.

Generalno, mogu se razlikovati četiri kategorije pristupa analizi zahteva i konceptualnom dizajnu, detaljnije u [38]:

- ✓ data-driven, razvoj konceptualne šeme se oslanja na podatke koji su dostupni u izvornim sistemima,
- ✓ user-driven, razvoj konceptualne šeme se oslanja na informacione zahteve poslovnih korisnika,
- ✓ business-driven, razvoj konceptualne šeme se oslanja na poslovnu strategiju i ciljeve poslovanja, kao i poslovne procese, i
- ✓ kombinovani, kombinacija prethodna tri pristupa. Mnogi od predloženih pristupa su ili suviše specifični ili suviše kompleksni da bi se na univerzalni način mogli primeniti u različitim realnim okruženjima.

Istraživanja još uvek treba da obezbede sveobuhvatan pristup analizi korisničkih zahteva, pri čemu se efikasnost ovog koraka može iskoristi na najbolji način samo ako se on sprovodi zajedno sa fazom konceptualnog dizajna, tako da formiraju jedinstven okvir. Usvajanje strukturiranog pristupa prilikom analize korisničkih zahteva i konceptualnog dizajna još uvek je veoma ograničeno u stvarnim projektima i oseća se snažna potreba za rešenjima koja

bi smanjila napore projektovanja i time, u isto vreme, smanjila rizik od neuspeha, proizvodeći dugovečno skladište podataka koje ispunjava korisničke zahteve i očekivanja.

Raznovrsnost aktuelnih pristupa analizi zahteva i konceptualnom dizajnu može da bude zbunjujuća za dizajnere, čak i za one koji su iskusni. Iz tog razloga, u ovom delu disertacije daje se pregled pristupa koje je predložila istraživačka zajednica i predstavljene su prednosti i mane različitih opcija.

5.3.1. Aktuelni pristupi dizajnu skladišta podataka

Istaknuto je da je PI deo informacionog sistema, prevashodno razvijen da omogućí upravljanje performansama poslovnog sistema i da stavi ukupne informacione potencijale preduzeća u funkciju donošenja najkvalitetnijih poslovnih odluka, kako bi se ostvarili utvrđeni strateški ciljevi kompanije. Ona koristi podatke koje kompanija skupljaju, najčešće u skladištima podataka i pretvara ih u informacije korisne za poslovanje. Podaci mogu biti strukturirani i nestrukturirani i mogu biti iz internih ili eksternih izvora. Najčešće je reč o otkrivanju prikriivenih znanja iz poslovnih podataka koje neka kompanija prikuplja rutinski, obavljajući svoje svakodnevne poslovne transakcije.

Već je istaknuto, koncept PI, zasniva se na sledećim idejama, videti [98]:

- ✓ Cilj poslovne inteligencije nije stvaranje što veće količine informacija, već generisanje kvalitetnijih, boljih informacija koje su potrebne pri donošenju poslovnih odluka.
- ✓ Potrebne informacije moraju biti pravovremene i iskazane na način koji korisnicima najviše odgovara.
- ✓ Ako se primeni na pravi način, koncept poslovne inteligencije smanjiće količinu informacija kojoj se zaposleni u kompanijama izlažu, uz istovremeno povećanje kvaliteta tih informacija, itd.

Skladištenje podataka je proces prikupljanja, transformisanja i integrisanja podataka, tipično iz više heterogenih izvora, koji kao rezultat daje skladište podataka, jednu od najobimnijih i najkompleksnijih vrsta baza podataka. Uskladišteni podaci su pročišćeni, konzistentni i integrisani. Skladište podataka je baza podataka koja sadrži istorijske podatke i to detaljne i sumarne. Ti podaci dolaze iz raznih izvora, uglavnom iz transakcionih baza. Skladište podataka je fizički odvojeno i logički izvedeno iz izvora podataka. Cilj je da se podaci o poslovanju neke kompanije integrišu u jednu bazu podataka iz koje krajnji korisnici mogu praviti izveštaje, postavljati upite i analizirati podatke.

Treba naglasiti da je skladištenje podataka proces koji se ne završava inicijalnim učitavanjem podataka, već se skladište podataka osvežava novim podacima u nekim, više ili manje pravilnim, vremenskim intervalima (npr. svakog dana, nedelje, meeseca). Iz toga sledi da je skladištenje podataka kontinuiran i dugotrajan proces.

Skladište podataka je informaciono okruženje sa sledećim karakteristikama, detaljno u [88]:

- ✓ Obezbeđuje integrisan i potpun pogled na kompaniju,
- ✓ Trenutne i istorijske informacije iz preduzeća čini lako dostupnim za donošenje odluka,
- ✓ Čini donošenje odluka mogućim bez ometanja rada operativnih sistema,
- ✓ Pruža konzistentne informacije,
- ✓ Predstavlja fleksibilan i interaktivan izvor strateških informacija, itd.

U transakcionim sistemima koristi se relacioni model podataka koji je normalizovan i optimizovan za postizanje velikih brzina obrade. Takav model podataka pokazao se izvanrednim kada je reč o transakcionim obradama u kojima se upiti odnose na najviše nekoliko desetina zapisa odjednom. Međutim, za potrebe skladišta podataka, u kojima se izvršavaju kompleksni upiti nad milionima zapisa istovremeno, taj model je neprihvatljiv. Zato se u skladištu relacioni model podataka zamenjuje dimenzionim modelom. Dimenzioni model podataka obezbeđuje da baza podataka bude jednostavna i razumljiva. Dimenziona baza podataka može se zamisliti kao multidimenzionalna kocka, tako da korisnici mogu pristupiti delu baze podataka preko bilo koje od njenih dimenzija. Osnovni koncepti multidimenzionalnog modela su, detaljno u [45]:

- ✓ Činjenice
- ✓ Mere
- ✓ Dimenzije
- ✓ Hijerarhije

Činjenica je fokus interesovanja procesa donošenja odluka. Ona obično predstavlja skup događaja koji se pojavljuju u kompaniji. Primeri činjenica u domenu saobraćajnih nezgoda su: učesnici nezgode, vreme događaja, lokacija, itd. Od suštinske važnosti za činjenicu je da ima određene dinamičke aspekte, odnosno, da se na neki način razvija kroz vreme. Koncepti koji se u izvoru podataka često ažuriraju dobri su kandidati za činjenice, dok oni koji su skoro statički, nisu. Izbor činjenica treba da se zasniva ili na prosečnoj učestalosti promena ili na specifičnim interesima analize, videti [43].

Mera je numeričko svojstvo činjenice i opisuje jedan od njenih kvantitativnih aspekata koji su od interesa za analizu. Na primer, svaka saobraćajna nezgoda ako ima materijalnu štetu, vrši se visina same štete. Mere bi trebalo da budu numeričke jer se koriste za izračunavanja

i daju odgovore na pitanja „koliko“. Činjenica takođe može da nema nikakve mere, ako je jedina interesantna stvar koja treba da se registruje sama pojava događaja, u kom slučaju se obično upitima samo broje događaji koji su se desili.

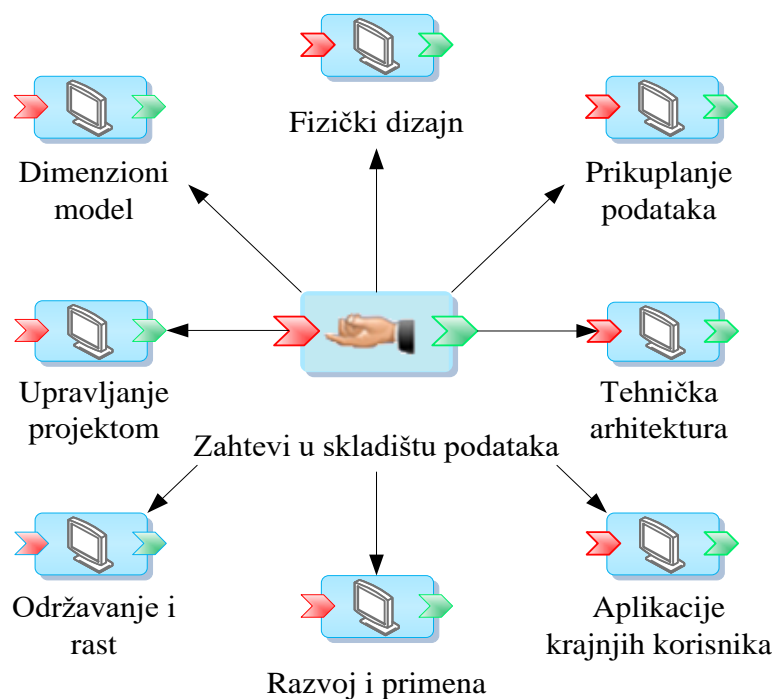
Dimenzija je svojstvo činjenice sa konačnim domenom i opisuje jednu od njenih koordinata analize. Skup svih dimenzija jedne činjenice određuje njenu granularnost (nivo detalja). Na primer, tipične dimenzije za činjenicu “nezgoda” su lica, vozila, mesto i slično. Zahvaljujući dimenzijama, mere iz tabele činjenica dobijaju značenje, jer dimenzije daju odgovore na pitanja “ko”, „šta“, „kada“ i „gde“. Obično jedna od dimenzija činjenica predstavlja vreme koje je neophodno da se na osnovu podataka u skladištu podataka izdvoje vremenske serije.

Odnos između mera i dimenzija je izražen, na nivou instance, preko koncepta događaja. Primarni događaj je pojava činjenice koja se identifikovuje pomoću n-torke u okviru koje postoji jedna vrednost za svaku dimenziju. Svaki primarni događaj je opisan preko jedne vrednosti za svaku meru. Primarni događaji su elementarne informacije koje se mogu predstaviti (koristeći metaforu kočke, oni odgovaraju ćeliji unutar kočke). Na primeru saobraćajne nezgode to je evidentiranje učenika u nezgodi, na jednoj lokaciji u jednoj vremenskoj dimenziji.

Agregacija je osnovna operacija u skladištu podataka jer omogućava da se značajne informacije izvedu na osnovu velike količine podataka. Sa konceptualnog stanovišta, agregacija se izvodi nad primarnim događajima, zahvaljujući definiciji atributa dimenzija i hijerarhija. Atribut dimenzije je svojstvo dimenzije sa konačnim domenom. Na primer, nezgoda se opisuje pomoću mesta, vremena, učesnika i slično. Odnosi između atributa dimenzije izraženi su pomoću hijerarhija. Tako bi na primeru mesta jedna hijerarhija mogla biti definisana na sledeći način: mesto -> opština -> okrug -> država.

5.3.2. Zahtevi u razvoju skladišta podataka

Na sledećoj slici predstavljen je položaj zahteva u celokupnom razvoju skladišta podataka. Slika prikazuje da zahtevi direktno utiču na aspekt dizajna sistema skladišta podataka, na tehnički aspekt, kao i na pitanja koja se odnose na upravljanje projektom. Zahtevi određuju koji podaci moraju biti dostupni u skladištu podataka, kako bi oni trebalo da budu organizovani, koliko se skladište često ažurira, ko ga, zbog čega i kako koristi, detaljno u [88]:



Slika 5.12. Zahtevi u skladištu podataka

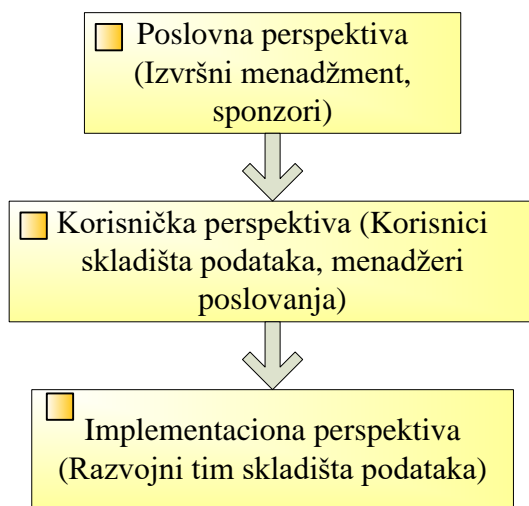
Zahtevi u razvoju skladišta podataka predstavljaju polaznu tačku za sve buduće projektne aktivnosti i značajno utiču na uspeh celokupnog projekta skladišta podataka. Zahtevi utiču skoro na svaku odluku koja se donese tokom dizajna i implementacije skladišta podataka. Veliki broj projekata skladišta podataka nikada ne budu realizovani, a još veći deo njih ne ispunjava poslovne ciljeve. Najčešće, neuspeh skladišta podataka dolazi kao posledica loše komunikacije između *dizajnera skladišta podataka i poslovnih korisnika, kao i slabih upravljačkih sposobnosti menadžera*. Kako bi se postigla uspešna implementacija skladišta podataka, potrebno je uložiti veliki trud vezan za inženjering zahteva, videti [124].

Zahtevi određuju funkcionalno ponašanje skladišta podataka kao i informacije koje ono treba da obezbedi, na primer, koje podatke je potrebno ubaciti u skladište podataka, kako se oni transformišu i organizuju, kao i kako se agregiraju i preračunavaju. Zahtevi omogućavaju zainteresovanim stranama da odrede svrhu, definišu smer i uspostave očekivanja koja imaju od budućeg skladišta podataka.

Zainteresovane strane često izražavaju svoje potrebe kroz generalizovana očekivanja da će sistem skladišta podataka unaprediti njihovo poslovanje. Ovaj “poslovni pogled” opisuje ciljeve i očekivanja zainteresovanih strana, koji predstavljaju osnovu zahteva u sistemu skladišta podataka. Sa druge strane, razvojni tim sistema skladišta podataka očekuje kompletnu, tačnu i nedvosmislenu specifikaciju sistema koji je potrebno razviti, što zahteva

dodatnu doradu poslovnih zahteva zainteresovanih strana. Zato je neophodno poslovne zahteve transformisati u detaljnu i kompletnu specifikaciju.

Iz ovog razloga, zahtevi u razvoju skladišta podataka imaju različite nivoe apstrakcije. Svaki nivo apstrakcije ima svoje zainteresovane strane i prikazuje sistem skladišta podataka iz druge perspektive. Sledeća slika prikazuje tri različite perspektive: poslovnu, korisničku i implementacionu, [124]:



Slika 5.13. Nivoi apstrakcije zahteva u skladištu podataka

Sledi kratak opis evidentiranih zahteva.

Poslovni zahtevi

Zahtevi iz poslovne perspektive predstavljaju ciljeve organizacije za sistem skladišta podataka, na najvišem nivou. Oni su obično obuhvaćeni u dokumentu kojim se opisuje vizija i opseg projekta. Ovi poslovni zahtevi identifikuju primarne benefite koje će skladište podataka pružiti organizaciji i njenim korisnicima. Oni predstavljaju najviši nivo apstrakcije u lancu zahteva. Izražavaju poslovne šanse i mogućnosti, poslovne ciljeve i opisuju tipične zahteve korisnika i organizacije, kao i vrednost koju oni dobijaju sistemom na najvišem nivou. Sledeća tabela prikazuje detaljan obrazac za obuhvatanje poslovnih zahteva. prema [123]:

Poslovni zahtevi		
Poslovna perspektiva	<i>Pokretači projekta</i>	<i>Obim i ograničenja</i>
	- Poslovne potrebe	- Dijagram konteksta sistema
	- Poslovni ciljevi	- Obim početne verzije
	- Zahtevi korisnika ili klijenata	- Obim narednih verzija i proširenja
	- Vrednost koja se obezbeđuje korisnicima	- Ograničenja i isključenja
	- Pozadina, trenutno stanje	
		<i>Poslovni kontekst</i>
	<i>Vizija / Poslovni slučajevi</i>	- Profil korisnika / klijenata
	- Dokument vizije	- Prioriteti projekta
	- Poslovni slučajevi	
- Značajni doprinosi / karakteristike	<i>Faktori uspeha projekta</i>	
- Pretpostavke i zavisnosti	- Faktori uspeha	
- Kriterijumi sprovođenja / napuštanja projekta	- Poslovni rizici	

Tabela 5.1. Poslovni zahtevi, [123]

Korisnički zahtevi

Zahtevi posmatrani iz perspektive korisnika opisuju zadatke koje korisnik mora da bude u stanju da postigne uz pomoć sistema skladišta podataka. Ovi zahtevi moraju da budu prikupljeni od ljudi koji će zaista koristiti sistem skladišta podataka i raditi sa njim. Zahtevi korisnika moraju da budu u skladu sa ciljevima postavljenim u poslovnim zahtevima. Oni mogu biti obuhvaćeni u slučajevima korišćenja ili opisima scenarija, koji se fokusiraju na to šta korisnik treba da uradi sa sistemom skladišta podataka, pa su zbog toga mnogo moćniji od tradicionalnog pristupa prikupljanju zahteva kada korisnika pitaju šta oni žele da sistem radi. Sledeća tabela predstavlja obrazac za zahteve korisnika, prema [122]:

Korisnički zahtevi	
Korisnička perspektiva	<i>Pokretači korisnika</i> - Ciljevi korisnika / aktera - Poslovna pitanja
	<i>Slučajevi korišćenja / Testni slučajevi</i> - Slučajevi korišćenja - Testni slučajevi
	<i>Poslovni procesi</i> - Definicija poslovnih procesa - Poslovna pravila - Mogućnosti za poboljšanje procesa
	<i>Korisnički kontekst</i> - Profili korisnika - Tipovi korisnika

Tabela 5.2. Korisnički zahtevi, [122]

Implementacioni zahtevi

Zahtevi posmatrani iz implementacione perspektive predstavljaju zahteve u razvoju skladišta podataka na veoma detaljnom nivou. Visok nivo detalja obuhvata kompletnu, detaljno razrađenu specifikaciju zahteva, koji predstavljaju značajan input razvojnom timu skladišta podataka. Oni moraju biti usklađeni sa poslovnim i korisničkim zahtevima.

Funkcionalni zahtevi definišu funkcionalnost koju razvojni tim mora ugraditi u sistem skladišta podataka, kako bi on korisnicima omogućio ispunjenje njihovih zadataka, a time posredno i ispunjenje poslovnih zahteva. Funkcionalni zahtevi obuhvataju očekivano ponašanje sistema skladišta podataka. Ovo ponašanje može da se izrazi pomoću servisa, zadataka ili funkcija koje bi sistem trebalo da izvršava. Oni opisuju šta sistem za analizu mora da uradi - mere koje sistem treba da preduzme da bi svojim korisnicima obezbedio korisnu funkcionalnost, videti [22].

Informacioni zahtevi definišu potrebu poslovnog sistema za informacijama. Oni opisuju informacije i podatke koje skladište podataka treba da omogući korisnicima, kao i informacije i podatke kojima skladište podataka treba da pristupi. Oni specificiraju podatke u skladištu podataka, određujući kvalitet koji oni treba da imaju, odakle dolaze, kako bi trebalo da budu obrađeni, kako bi ih trebalo kombinovati za proces analize i koji metodi analize će se koristiti.

Pored funkcionalnih i informacionih zahteva, mogu se specificirati i drugi zahtevi kojima se opisuju ostali relevantni aspekti sistema skladišta podataka, kao što su zahtevi vezani za

interfejs ili okruženje (kulturni, politički, pravni). Sledeća tabela prikazuje obrazac za specificiranje funkcionalnih, informacionih i zahteva interfejsa.

Funkcionalni zahtevi		Informacioni zahtevi		Zahtevi interfejsa	
Implementaciona perspektiva	<i>Priprema podataka</i> - Ekstrakcija podataka - Transformacija podataka - Čišćenje podataka - Punjenje podataka - Prikupljanje podataka - Arhitektura	Implementaciona perspektiva	<i>Izvori podataka</i> - Interni sistemi - Eksterni sistemi - Provajderi informacija / podataka - Mapiranja podataka	Implementaciona perspektiva	<i>Interni interfejsi</i> - Softverski interfejsi - Hardverski interfejsi - Interfejsi baza podataka <i>Eksterni interfejsi</i> - “Front-end” interfejsi - Interfejsi eksternih sistema <i>Korisnički interfejsi</i> - Ugođaj (“Look and Feel”) - Upotrebljivost
	<i>Upravljanje podacima</i> - Administracija / održavanje - Pravljenje rezervne kopije podataka i oporavak - Upravljanje metapodacima		<i>Upiti / izveštavanje</i> - Izveštavanje - Data mining - EIS / MIS - Vizuelizacija podataka - Ad hok upiti		
	<i>“Font-end” zahtevi</i> - Pregled - Izvoz / uvoz podataka - Raspored / formatiranje - Manipulacija podacima - Pretraga - “Roll Out”		<i>OLAP Zahtevi</i> - Indikatori performansi, mere - Dimenzije analize - Usaglašene dimenzije analize - Vremenske dimenzije - Strategije za promenljive dimenzije - Kočke podataka		

Tabela 5.3. Implementacioni zahtevi, [22]

Dodatno, atributi zahteva proširuju opis funkcionalnih, informacionih ili drugih zahteva, opisujući karakteristike u različitim dimenzijama koje su značajne, bilo za korisnike bilo za razvojni tim. Oni predstavljaju obeležja ili kvalitete koje bi sistem skladišta podataka trebalo da ima. Oni mogu da obuhvate standarde, propise i uslove koje sistem skladišta podataka mora da ispuni, opise eksternih interfejsa, zahteve vezane za performanse, ograničenja dizajna i implementacije i napokon attribute u vezi sa kvalitetom. Atributi zahteva su obično vezani za detaljne zahteve sistema. Na primer, kada su funkcionalni zahtevi poznati, moguće je utvrditi kako će se sistem ponašati, koja obeležja treba da ima, koliko veliki i brz treba da bude. Kada su informacioni zahtevi poznati, mogu se odrediti njihovi atribti kao što su kvalitet ili granularnost podataka. Sledeća tabela prikazuje obrazac za specificiranje atributa funkcionalnih i informacionih zahteva.

Funkcionalni atributi		Informacioni atributi	
Implementaciona perspektiva	<i>Operativni atributi</i>	Implementaciona perspektiva	<i>Kvalitet informacija</i>
	- Dostupnost i pouzdanost		- Suštinski kvalitet informacija
	- Efikasnost		* Ispravnost
	- Fleksibilnost		* Preciznost
- Interoperabilnost	* Doslednost	- Kvalitet informacija u pogledu konteksta	
- Prenosivost	* Kompletnost	- Kvalitet informacija u pogledu reprezentacije	
- Robustnost		- Kvalitet informacija u pogledu dostupnosti	
<i>Performanse</i>		<i>Bezbednost informacija</i>	
- Performanse upita		- Opšta bezbednosna pravila	
- Performanse punjenja		- Bezbednosna pravila izveštavanja	
<i>Bezbednost</i>		- Bezbednosna pravila OLAP sistema	
- Identifikacija / autentifikacija			
- Autorizacija / Kontrola pristupa			
- Praćenje akcija			
- Šifrovanje / Dešifrovanje			
- Bezbednost mreže			

Tabela 5.4. Atributi implementacionih zahteva, [22]

5.3.3. Faze dizajna skladišta podataka

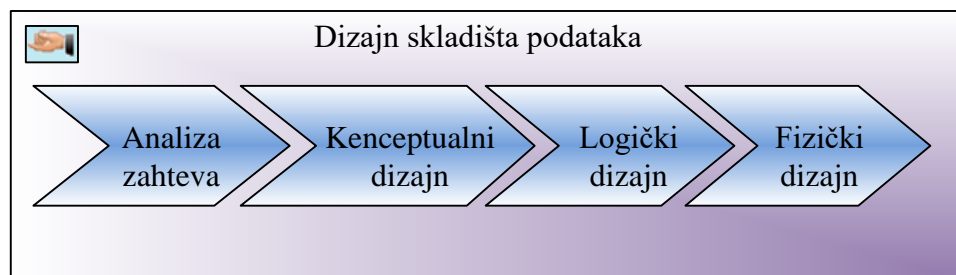
Postoji nekoliko referenci u kojima je predložen opšti metod za dizajn skladišta podataka. Međutim, ove publikacije nisu usaglašene vezano za faze koje bi trebalo slediti u projektovanju skladišta podataka. Neki autori smatraju da tradicionalne faze razvoja operativnih baza podataka (specifikacija zahteva, konceptualni dizajn, logički dizajn i fizički dizajn) mogu da se koriste i u razvoju kompletnog skladišta podataka ili lokalnih skladišta podataka. Drugi autori ignorišu neke od ovih faza, posebno fazu konceptualnog dizajna. Mnoge publikacije se odnose samo na jednu od faza, bez razmatranja koje bi transformacije bile potrebne kako bi se postiglo rešenje koje bi se moglo implementirati.

Neki predlozi smatraju da je razvoj sistema skladištenja podataka, prilično drugačiji od razvoja operativnih sistema baza podataka. S jedne strane, oni uključuju dodatne faze, kao što je ETL proces. Više o samom ETL procesu i njegovoj važnosti za skladište podataka biće u nastavku disertacije. S druge strane, oni obezbeđuju različite metode za fazu specifikacije zahteva. Nekoliko publikacija pominje važnost stvaranja metapodataka kao deo procesa dizajna skladišta podataka, što uključuje ne samo opis modela skladišta podataka, već i informacije o izvornim sistemima i ETL procesima.

Većina predloženih pristupa podrazumeva sledeće podfaze u okviru faze dizajna kompletnog ili lokalnog skladišta podataka, videti [45]:

- ✓ Analiza zahteva: Identifikuje se koje informacije su značajne za proces odlučivanja uzimajući u obzir potrebe poslovanja i korisnika ili dostupnost podataka u operativnim izvorima.
- ✓ Konceptualni dizajn: Kao rezultat ove faze trebalo bi da se izvede implementaciono nezavisna konceptualna šema skladišta podataka.
- ✓ Logički dizajn: Polazeći od konceptualne šeme stvara se odgovarajuću logička šemu u izabranom logičkom modelu. Dok je danas većina sistema skladišta podataka zasnovana na relacionom logičkom modelu (ROLAP), sve veći broj proizvođača softvera takođe predlaže čista ili mešovita multidimenzionalna rešenja (MOLAP/HOLAP).
- ✓ Dizajn ETL procesa: Utvrđuju se mapiranja i transformacije podataka koje su neophodne da bi se podaci dostupni u operativnim izvorima učitali u logičku šemu skladišta podataka.
- ✓ Fizički dizajn: Bavi se svim pitanjima koja su specifična za paket alata koji su izabrani za implementaciju, kao što su, na primer, indeksiranje, alociranje i slično.

Sledeća slika prikazuje faze u dizajnu skladišta podataka.



Slika 5.14. Faze u dizajnu skladišta podataka, [45]

Važno je napomenuti da, iako prethodna slika različite faze prikazuje redom, u realnosti postoji više interakcija između njih, posebno ako se usvoji iterativni proces razvoja u kome se sistem razvija u inkrementalnim verzijama sa povećanim funkcionalnostima.

Pregled pristupa analizi zahteva i konceptualnom dizajnu

U zavisnosti od toga da li se kao pokretačka snaga koriste izvorni sistemi, korisnici, poslovni ciljevi (i/ili procesi) ili njihova kombinacija, predloženo je nekoliko različitih pristupa za specifikaciju zahteva i konceptualni dizajn. U nastavku su ovi pristupi opisani, pri čemu su grupisani u: data-driven, user-driven, business-driven i kombinovane pristupe.

Data-driven pristupi

Data-driven pristup, koji se takođe naziva i source-driven ili supply-driven pristup, je bottom-up tehnika koja počinje od analize operativnih izvora podataka, kako bi se utvrdili svi raspoloživi podaci. Analizom postojećih izvornih sistema dobija se šema skladišta podataka. Neke od predloženih tehnika zahtevaju konceptualne reprezentacije operativnih izvornih sistema, u većini slučajeva baziranih na modelu objekti-veze. Ostale tehnike koriste relacione tabele i pomoću njih predstavljaju izvorne sisteme. Ove izvorne šeme treba da pokazuju dobar stepen normalizacije da bi se olakšala ekstrakcija činjenica, mera, dimenzija i hijerarhija.

U principu, učešće korisnika nije izričito potrebno, ali u nekim tehnikama korisnici bi trebalo da analiziraju dobijenu šemu da bi potvrdili ispravnost izvedenih struktura ili da identifikuju određene činjenice i mere. Kao polaznu tačku za dizajn multidimenzionalnih šema, određujući koji su od raspoloživih podataka od značaja za proces donošenja odluka. S obzirom da se podaci predstavljaju na niskom nivou detalja, uglavnom se uključuju korisnici koji se nalaze na operativnom ili administrativnom organizacionom nivou. Nakon kreiranja početnih šema, korisnici mogu da definišu svoje informacione zahteve izborom elemenata koje su od interesa.

S druge strane, ovaj pristup zahteva visoko obučene i iskusne dizajnere. Pored uobičajenih sposobnosti modelovanja, oni dodatno treba da imaju dovoljno znanja o poslovanju kako bi bili u stanju da razumeju poslovni kontekst i potrebe poslovanja, oslanjajući se uglavnom na operativne podatke. Oni takođe treba da imaju sposobnost da shvate strukturu postojećih operativnih baza podataka.

Data-driven pristupi su primenljivi kada su su ispunjeni sledeći uslovi, videti u [10]:

- ✓ Detaljno znanje o izvorima podataka je dostupno ili se može lako obezbediti,
- ✓ Izvorne šeme pokazuju dobar stepen normalizacije, i
- ✓ Kompleksnost izvornih šeme nije previše velika.

Data-driven pristup ima nekoliko prednosti:

- ✓ Osigurava da skladište podataka odražava postojeće relacije u podacima.
- ✓ Skladište podataka ispočetka sadrži sve neophodne podatke.
- ✓ Razvijanje skladišta podataka na osnovu postojećih operativnih baza podataka pojednostavljuje ETL proces, jer svaki podatak u skladištu podataka odgovara jednom ili više izvornih atributa.
- ✓ Korporativni model podataka implementiran u izvornom sistemima može pružiti stabilniju osnovu za dizajn, nego zahtevi korisnika koji mogu biti podložni

promenama. Tako se dobija multidimenzionalni model koji je stabilan, jer je baziran na šemi operativnih izvora podataka koji se ne menjaju tako često kao lični zahtevi poslovnih korisnika.

- ✓ Da bi projekat počeo potrebno je minimalno vreme korisnika.
- ✓ Razvojni proces može da bude brz i jednostavan ako postoje dobro strukturirani i normalizovani operativni sistemi.
- ✓ Data-driven pristup je jednostavniji i jeftiniji (u smislu vremena i novca) nego drugi pristupi, jer njegovo trajanje zavisi samo od veština dizajnera i složenosti izvora podataka.
- ✓ Ako su operativne baze podataka predstavljene modelom objekti-veze ili normalizovanim relacionim tabelama, mogu se primeniti automatske ili poluautomatske tehnike.

Rodonačelnik ove oblasti **Inmon**, smatra da se okruženje skladišta podataka dobija polazeći od raspoloživih izvornih podataka. Prema njemu, zahteve treba razumeti nakon što se skladište podataka popuni podacima i koristi od strane analitičara za podršku odlučivanju. Model podataka se izvodi tako što se korporativni model podataka prevodi u šemu skladišta podataka i dodaju se faktori performansi.

User-driven pristupi

User-driven pristup, koji se takođe naziva i demand-driven pristup, je bottom-up tehnika koja polazi od utvrđivanja informacionih zahteva različitih poslovnih korisnika. U ovom pristupu smatra se da korisnici imaju ključnu ulogu u toku analize zahteva i da moraju biti aktivno uključeni u rasvetljavanju relevantnih činjenica i dimenzija. Njihova gledišta se zatim integrišu i čine konzistentnim u cilju dobijanja jedinstvenog skupa multidimenzionalnih šema. Naglasak je na procesu analize zahteva i na pristupima kojima se omogućava uključivanje korisnika. Za specifikaciju informacionih zahteva biraju se korisnici iz različitih nivoa organizacije i koriste se razne tehnike, kao što su intervjui i sesije.

Identifikacija ključnih korisnika za specifikaciju zahteva je krucijalan zadatak. Potrebno je uzeti u obzir nekoliko aspekata, detaljnije u [98]:

- ✓ Ciljni korisnici treba da budu svesni sveukupnih poslovnih ciljeva da bi se izbegla situacija u kojoj zahtevi predstavljaju lične percepcije korisnika u skladu sa njihovim ulogama u organizaciji ili njihovim specifičnim poslovnim jedinicama.
- ✓ Treba izbeći ili ublažiti uticaj korisnika koji bi dominirali procesom specifikacije zahteva kako bi se osiguralo da će se razmotriti informacione potrebe različitih korisnika.

- ✓ Korisnici treba da budu dostupni i treba da se slože da učestvuju tokom celog procesa prikupljanja zahteva i konceptualnog dizajna.
- ✓ Korisnici treba da imaju ideju o tome šta sistemi skladišta podataka i OLAP sistemi mogu da ponude. Ako ovo nije slučaj, korisnicima bi to trebalo predstaviti putem objašnjenja, demonstracija ili prototipova.

User-driven pristup ima nekoliko prednosti:

- ✓ Poslovni korisnici veoma cene ovaj pristup, jer su uključeni u dizajn i mogu da razumeju opravdanost napravljenih izbora.
- ✓ Efektivno učešće korisnika obezbeđuje bolje razumevanje činjenica, dimenzija, kao i odnosa koji postoji između njih.
- ✓ Kako su korisnici uključeni od samog početka, povećava se verovatnoća prihvatanja sistema skladišta podataka.
- ✓ Pristup je koristan za kreiranje prototipova i primenljiv je u domenima koji zavise od korisnika.

User-driven pristup ima sledeće nedostatke:

- ✓ Mapiranje korisničkih zahteva na dostupne izvore podataka obavlja samo naknadno i može da ne uspe, na taj način dovodeći do razočarenja korisnika.
- ✓ Izabrane mere, dimenzije i nivo granularnosti zavisi od korisnika koji su uključeni u razvoj.
- ✓ Razvojni proces može da bude vremenski zahtevan, jer poslovni korisnici na taktičkom nivou retko imaju jasno i zajedničko razumevanje poslovnih ciljeva, procesa i organizacije. Time i troškovi mogu biti veći.
- ✓ Specifikacija korisničkih zahteva koji nisu usklađeni sa poslovnim ciljevima može da proizvede veoma složenu šemu koja ne podržava procese donošenja odluka korisnika, na svim organizacionim nivoima.
- ✓ Ovaj pristup obično zahteva velike napore od strane menadžera projekta, koji mora biti prilično umeren i imati veoma dobre liderske veštine, u cilju integrisanja različitih tačaka gledišta.
- ✓ Rizik od zastarelosti dobijene šeme je veliki, ako se zahtevi zasnivaju na ličnim gledištima korisnika i ne izražavaju kulturu kompanije niti radne procedure.

Business-driven pristupi

Business-driven pristup, koji se takođe naziva i process-driven, goal-driven ili requirements-driven pristup, je top-down tehnika koja polazi od toga da korisnici često nisu u mogućnosti da jasno formulišu svoje potrebe. Stoga izvođenje struktura skladišta podataka počinje od analize poslovnih zahteva ili analize poslovnih procesa. U prvom slučaju se razmatraju poslovni zahtevi na najvišem nivou organizacije, dok se ne identifikuju neophodni

multidimenzionalni elementi. Shodno tome, dobijena specifikacija će obuhvatiti zahteve korisnika na svim organizacionim nivoima i biće usklađena sa opštim poslovnim ciljevima. S druge strane, analiza poslovnih procesa zahteva da se odrede različiti poslovni servisi ili aktivnosti koje proizvode određeni izlaz. Različiti elementi koji učestvuju u ovim aktivnostima mogu se razmatrati kao dimenzije u šemi skladišta podataka. Metrike koje donosioci odluka koriste za evaluaciju poslovnih aktivnosti, mogu se razmatrati kao mere. Primenljivost ovog pristupa strogo zavisi od volje vrha rukovodstva da učestvuje u procesu projektovanja i obično zahteva sposobnost projektnog tima da prevede zahtev prikupljen na visokom nivou u merljive ključne indikatore efikasnosti.

Neke od prednosti business-driven pristupa su sledeće, videti u [38]:

- ✓ Pristupi koji su orijentisani na ciljeve povećavaju verovatnoću ispravnosti identifikacije relevantnih indikatora, čime se smanjuje rizik od zastarelosti multidimenzionalne šeme.
- ✓ Pošto se zahtevi za skladištenjem podataka izvode iz poslovne perspektive, oni obezbeđuju sveobuhvatnu i preciznu specifikaciju potreba zainteresovanih strana sa njihove poslovne tačke gledišta.
- ✓ Korišćenje formalnih tehnika za definisanje modela poslovnih procesa, obezbeđuje formalni opis informacionih zahteva korisnika.
- ✓ Ovaj pristup može da poveća prihvatanje sistema ako postoji kontinuirana interakcija sa donosiocima odluka.
- ✓ Mogu se specificirati eksterni podaci koje bi trebalo uključiti u skladište podataka.
- ✓ Mogu se specificirati dugoročni strateški ciljevi.

Međutim, neki nedostaci ovog pristupa mogu da odigraju važnu ulogu u određivanju njegove upotrebljivosti za konkretni projekat skladištenja podataka:

- ✓ Specifikacija poslovnih ciljeva može biti težak proces, a rezultat zavisi od tehnika koje se primenjuju i veština razvojnog tima.
- ✓ Specifikacija zahteva na osnovu poslovnih procesa može da postane komplikovanija ukoliko ovi procesi prelaze granice organizacije.
- ✓ Projekat uglavnom traje duže. Prema tome, troškovi projekta mogu biti veći.
- ✓ Definisani zahtevi možda neće biti ispunjeni informacijama koje postoje u izvornim sistemima.

U mnogim slučajevima iz prakse, razlika između usvajanja business-driven umesto user-driven pristupa može postati veoma nejasna. Ipak, trebalo bi da bude jasno da je business-driven proces top-down i baziran na progresivnom prečišćavanju nekoliko ciljeva definisanih od strane menadžmenta na najvišem nivou, dok se u user-driven pristupu zahtevi

dobijaju spajanjem nekoliko jednostavnijih zahteva, prikupljenih od poslovnih korisnika u bottom-up stilu. Rezultat business-driven pristupa razlikuje se od rezultata user-driven pristupa, uvek kada korisnici ne razumeju jasno poslovnu strategiju i organizacione ciljeve.

Kombinovani pristupi

Da bi se izbegli nedostaci jednog “čistog” pristupa razvijeni su neki mešoviti pristupi koji predstavljaju kombinaciju business-driven ili user-driven pristupa i data-driven pristupa, uzimajući u obzir ono što poslovne i korisničke zahteve i ono što izvorni sistemi mogu da obezbede. U idealnoj situaciji, ove dve komponente treba, da se uklapaju, to jest, sve informacije koje korisnici ili poslovanje zahtevaju za svrhe analize treba da se obezbede na osnovu podataka koji se nalaze u izvornim sistemima. Ovaj pristup se naziva i top-down/ bottom-up analiza.

S jedne strane, pristupi se mogu koristiti sekvencijalno, tako da se rezultati jednog koraka koriste kao ulaz u drugi korak, a sa druge strane, koraci mogu biti nezavisni, pri čemu se njihovi rezultati naknadno kombinuju.

Preporuke u vezi sa korisnicima i razvojnim timom koje su prethodno navedene takođe treba razmatrati i ovde. Kombinovani pristupi imaju nekoliko važnih prednosti, detaljnije u [10]:

- ✓ Generiše se izvodljivo rešenje (odnosno rešenje je podržano postojećim izvorima podataka) koje bolje odražava ciljeve korisnika.
- ✓ Mogu da ukažu na podatke koji nedostaju u operativnom bazama podataka, a koji su potrebni za podršku procesu donošenja odluka.
- ✓ Ako izvorni sistemi nude više informacija nego što poslovni korisnici u početku traže, analiza se može proširiti i na nove aspekte koji do sada nisu uzimani u obzir.

Međutim, kombinovani pristup ima sledeće nedostatke:

- ✓ Razvojni proces je komplikovaniji, pošto su potrebne dve šeme, jedna se dobija na osnovu definicije poslovnih zahteva, a druga proizilazi iz postojećih izvornih sistema.
- ✓ Proces integracije kojim se utvrđuje da li izvori podataka pokrivaju potrebe poslovanja i korisnika može da zahteva složene tehnike.

Poređenje aktuelnih pristupa

U prethodnom delu ukazano je na prednosti i mane aktuelnih pristupa analizi korisničkih zahteva i konceptualnom dizajnu. U sledećoj tabeli daje se komparativni prikaz tri osnovna pristupa i može biti korisna prilikom izbora onog koji najviše odgovara datom projektu.

Glavni tehnički element koji utiče na izbor odgovarajućeg pristupa odnosi se na dostupnost i kvalitet šeme operativnih izvora podataka, ali je u izbor uključeno i nekoliko netehničkih

faktora. Na primer, ograničenja u pogledu troškova i vremena zahtevaju smanjenje vremena posvećenog intervjuima i razgovorima sa korisnicima. Slično, kada poslovni korisnici nemaju dovoljno znanja o poslovnim procesima i strategiji, trebalo bi izbegavati user-driven pristup. Sledi tabelarni prikaz osnovnih principa, prema [10]:

	Data-driven	User-driven	Business-driven
<i>Osnovni pristup</i>	Bottom-up	Bottom-up	Top-down
<i>Uključenost korisnika</i>	Mala: Administratori baza podataka	Velika: Poslovni korisnici	Umerena: Najviši menadžment
<i>Podrška projektu</i>	Ne postoji	Odeljenje	Najviši menadžment
<i>Fokus</i>	Kratkotrajan	Kratkotrajan	Dugotrajan
<i>Oblast primene, domen zahteva</i>	Istraživanje podataka i data mining	Povećanje prihvatanja sistema	Osnova za podršku odlučivanju
<i>Ograničenja</i>	Postojanje usklađenih nivoa podataka	Poslovni korisnici moraju dobro da poznaju procese i organizaciju kompanije	Spremnost najvišeg menadžmenta da učestvuju u procesu dizajna
<i>Snaga</i>	Obezbeđena je dostupnost podataka	Povećanje prihvatanja sistema	Povećanje verovatnoće ispravnosti identifikovanja relevantnih ključnih indikatora performansi
<i>Rizici</i>	Multidimenzionalne šeme nisu u skladu sa zahtevima poslovnih korisnika	Brzo zastarevanje multidimenzionalnih šema zbog promene poslovnih korisnika	Teškoće da se dobije podrška najvišeg menadžmenta i da se poslovna strategija prevede u merljive ključne indikatore performansi
<i>Ciljni organizacioni nivo</i>	Operativni i delimično taktički	Zavisi od nivoa ispitanih korisnika; obično taktički	Strateški i taktički
<i>Veštine članova projektnog tima</i>	Dizajneri skladišta podataka	Moderatori, dizajneri skladišta podataka	Moderatori, ekonomisti, dizajneri skladišta podataka

<i>Rizik od zastarelosti</i>	Nizak	Visok	Nizak
<i>Broj izvornih sistema</i>	Mali	Umeren	Veliki
<i>Broj mera</i>	Mnogo	Mnogo	Nekoliko
<i>Tip mera</i>	Nefinansijske i kvantitativne	Nefinansijske i kvantitativne	Balansirano: Finansijske i nefinansijske kao i kvalitativne i kvantitativne
<i>Nivo granularnosti</i>	Nizak	Nizak	Visok
<i>Broj dimenzija</i>	Nekoliko	Mnogo	Nekoliko
<i>Tip dimenzija</i>	Predstavlja osnovnu strukturu aplikacije	Predstavlja osnovnu strukturu aplikacije i eksternih izvora	Predstavlja strateške gradivne elemente organizacije
<i>Dugovečnost, stabilnost modela podataka</i>	Velika	Mala	Velika
<i>Troškovi</i>	Niski	Visoki	Visoki
<i>Trajanje projekta</i>	Kratko	Veoma dugo	Dugo

Tabela 5.5. Poređenje osnovnih pristupa

Data-driven metodologija razvoja se preporučuje za svrhe “data mining”-a i istraživanja podataka. Dobijeni konceptualni model jeste dugovečan i stabilan, jer je baziran na šemi operativnih izvora podataka koji, se ne menjaju tako često kao lični zahtevi poslovnih korisnika, ali može da ne bude u skladu sa zahtevima poslovnih korisnika. Postoji rizik generisanja informacija o performansama poslovanja koje ne odgovaraju ciljnim korisničkim grupama. Kako se oslanja na postojeće podatke, ovaj pristup ne može da se koristi kada je potrebno analizirati ili odrediti dugoročne strateške ciljeve. Značajna prednost ovog pristupa u odnosu na druge su niski troškovi i kratko trajanje projekta, jer njegovo trajanje zavisi samo od veština dizajnera i složenosti izvora podataka.

U suštini, monopolizacija user-driven pristupa je rizična i treba je izbegavati, jer ona generiše informacije o performansama koje odražavaju organizacioni nivo ljudi, koji su uključeni. Zbog toga su izabrane mere, dimenzije, nivo granularnosti i ciljani nivo organizacione hijerarhije veoma nestabilni. Metodologija ima bottom-up tendenciju, jer većina zaposlenih ne vidi organizaciju iz šireg ugla, ima sužene vidike i egocentričnu perspektivu. Projekat može da bude vema dug i skup, jer učesnici u projektu zahtevaju duge rasprave o mnogim nepotrebnim merama i dimenzijama. Analiziranje kriterijuma user-driven pristupa nema smisla, jer se rezultati menjaju u zavisnosti od ljudi koji su uključeni. Ovaj pristup može povećati prihvatanje sistema, ali se mora kombinovati sa data-driven ili business-driven pristupom da bi se povećala dugovečnost sistema. Što korisnici više odbijaju

sistem, to je potrebnije uključivanje korisnika, pored fokusa na organizacione strategije ili korporativni model podataka.

Business-driven pristup podržava savremene upravljačke metode i osnova je za podršku odlučivanju na svim organizacionim nivoima. Nivo granularnosti je mnogo viši u odnosu na data-driven pristup. Trajanje razvoja projekta ima tendenciju da bude veoma obimno i skupo, kako mnogo visoko kvalifikovanih profesionalaca i menadžera učestvuju u brojnim radionicama i iz strategija izvodi pokazatelje performansi. Krajnji korisnici su retko uključeni. Oni su samo potrebni kada su u pitanju operativni detalji. Kako je model usklađen sa korporativnom strategijom, on je stabilan. Mere i dimenzije su uravnotežene, uzimaju se u obzir finansijski, nefinansijski, kvalitativni i kvantitativni aspekti. Integriše se mnogo izvora podataka, jer je holistički pristup baziran na svim aspektima organizacije.

Data-driven i business-driven pristupi ne podrazumevaju uzajamnu isključivost. Kako slede različite namene mogu da postoje paralelno. Data-driven pristup se čak može posmatrati kao niži nivo detalja business-driven pristupa. Ovi pristupi su komplementarni i benefiti su čak veći kada se koriste paralelno.

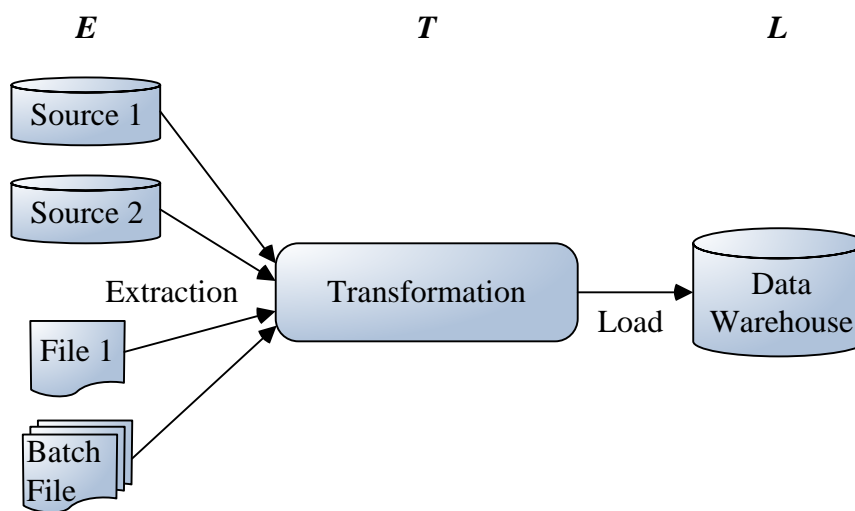
Analiza zahteva i konceptualni dizajn su u velikoj meri odgovorni za uspeh projekata skladištenja podataka, jer bi u toku ove dve faze trebalo u potpunosti definisati multidimenzionalne šeme. Većina faktora rizika može se uspešno tretirati usvajanjem odgovarajućih metodologija. U literaturi su dostupni neki osnovni pristupi analizi korisničkih zahteva i konceptualnom dizajnu. Jedni se oslanjaju isključivo na podatke dostupne u izvornim operativnim sistemima, drugi na zahteve korisnika, a treći na ciljeve ili procese poslovanja. Međutim, usvajanje “čistog” pristupa nije u stanju da se zaštiti od sopstvenih slabosti. Zato su neki autori predložili i testirali na realnim projektima mešovite tehnike čime su pokazali da tri osnovna pristupa nisu međusobno isključiva, već su, umesto toga, komplementarna i kada se oni koriste paralelno mogu da se prevaziđu mnogi problemi. Spajanje data-driven, user-driven i business-driven koraka može da dovede do prikupljanja kompletne specifikacije i obezbeđivanja većeg stepena dugovečnosti, kao i prihvatanja od strane korisnika.

5.4. ETL (Extraction, Transformation, Loading) procesi

Već je ranije istaknuto, da se podaci skladište u skladište podataka iz različitih izvora, najčešće iz transakcionih sistema, poslovnog sistema. Takođe u četvrtom poglavlju disertacije dat je opsežan prikaz važnosti podatka u PI. U ovom delu disertacije, ideja je da se ukaže na važnost ehstrahovanja podataka iz različitih spoljnih izvora u skladište podataka.

Najopsežniji posao u aktivnostima skladištenja podataka predstavlja procese integriranja podataka i organizovanja njihovog sadržaja. Zadatak ovih procesa je da izvrše celovito transformisanje i punjenje tj. unošenje podataka iz jednog ili više transakcionih sistema u skladište podataka. Od ukupnog vremena potrebnog za projektovanje skladišta podataka, smatra se da čak 70% se utroši na pomenuti proces.

ETL procesima prethode pripremne aktivnosti vezane za skladištenje i čišćenje podataka. Najčešće su izvorni podaci nekompletni (nedostaju vrednosti atributa ili postoje samo agregirani podaci), neprecizni (netačne vrednosti atributa) ili nekonzistentni (neslaganje u imenima ili šiframa). Stoga je potrebno standardizovati ih, odnosno prevesti u jedinstven format. U tom formatu će se podaci koristiti u svim daljim fazama obrade. Standardizacijom se izbegava redundansa podataka. Čišćenje, kao pripremna aktivnost ETL procesa, ima zadatak da ukloniti sve one podatke koji se pojavljuju kao posledica ranijih grešaka u radu informacionih sistema. Dalje, pri implementaciji je potrebno voditi računa o skalabilnosti ETL sistema. Ova vrlo bitna karakteristika se odnosi na razumevanje obima podataka koji će biti procesuirani. Povećanje količine podataka može prouzrokovati neophodne promene u dizajnu sistema.



Slika 5.15. Model ETL procesa

Za izgradnju ETL procesa, kompanije obično kupuju ETL alate. ETL alati moraju biti takvi da omoguće komunikaciju sa različitim relacionim bazama podataka i čitanje fajlova različitih formata. Karakteristika standardnih ETL alata je da imaju bolje performanse vezane za procese čišćenja podataka, ali su im tada slabiji kapaciteti vezani za procese transformacije ili obrnuto. Stoga je pre izbora potrebno znati karakteristike podataka koji će se unositi u skladište podataka. Što je kompleksnija transformacija podataka, to je veća potreba za nabavkom dodatnih ETL alata. Sledi opis svake pomenute aktivnosti ETL-a.

Ekstrakcija

Ekstrakcija podataka je proces prikupljanja podataka iz različitih izvora i platformi, radi smeštanja istih u skladište podataka. Predstavlja jako bitan selektivni proces unošenja informacija bitnih za proces odlučivanja u skladište podataka. Proces ekstrakcije podataka je potrebno izvesti na način, da redovni operativni poslovi i operativne transakcije što manje trpe. Iz tog razloga su programi i alati za ekstrakciju, oblikovani tako da ETL procese mogu obavljati što produktivnije, uz nastojanje da potrebne podatke iz operativnih procesa preuzimaju što je moguće brže. Pri tom se kao problem može pojaviti potencijalno visok stepen redundanse podataka u transakcionim sistemima. Zato treba odabrati pristup ekstrakciji kojim se vrši zahvatanje samo onih podataka koji će se koristiti u aplikacijama sistema poslovne inteligencije.

Metode izvlačenja podataka iz operacionih baza su sledeće, videti u [7]:

- ✓ Izvlačenje podataka iz log fajlova.
- ✓ Neki sistemi kreiraju poseban fajl promena, koja se pri izvlačenju koristi na isti način kao log fajl.
- ✓ Ako izvorni podaci imaju oznake vremena, u procesu izvlačenja podataka mogu se odabrati samo oni podaci koji su se promenili od zadnjeg izvlačenja podataka.
- ✓ Metoda koja se godinama koristila je poređenje datoteka.

Transformacija

Kad se podaci izvuku iz operacionih baza podataka i spoljašnjih izvora, treba ih pripremiti za učitavanje u skladište podataka, tj. treba ih transformirati u prikladan format. Podatke treba prilagoditi modelu podataka odredišne baze podataka, proveriti njihovu tačnost i kvalitet, pa ih potom integrisati. U okviru ETL procesa, najviše vremena se troši na postupak transformacije podataka, koji prema stručnim procenama traje i do 80% od ukupnog ETL procesa. U postupku transformacije se mogu pojaviti različiti problemi koji usporavaju proces, a najčešći su, videti [7]:

- ✓ *Nekonzistentne vrednosti podataka* - pojavljuju se prilikom kopiranja podataka, pri čemu kopija u tom slučaju ne bude verna originalu.
- ✓ *Nepodudarnost primarnih ključeva* korišćenih u izvornim datotekama i bazama podataka s primarnim ključevima koje pretpostavljaju aplikacije poslovne inteligencije.
- ✓ *Netačne vrednosti podataka* - potrebno je definisati logiku čišćenja za ispravljanje netačnih vrednosti podataka. Čišćenje je potrebno izvoditi stalno, odnosno svakim novim ciklusom punjenja podataka.

- ✓ *Različiti formati podataka* - format podataka, npr. broj žiro-računa kupca nije evidentiran u operativnoj bazi u istom formatu kako ga očekuje alat poslovne inteligencije, pa ga je potrebno prilagoditi.
- ✓ *Problem sinonima i homonima* - ponekad se isti podaci pojavljuju pod različitim nazivima. Isto tako događa se da različiti podaci nose isti naziv tj. govori se o homonimima. Oba slučaja su nepoželjna i treba ih uklanjati.
- ✓ *"Skrivena" procesna logika* - može predstavljati problem koji usporava proces transformacije kad se za zahvatanje podataka koriste zastareli softverski sistemi. U njima se mogu naći podaci koje nije lako protumačiti i često je potrebno izvršiti transformaciju u neki drugi poznati tip odnosa.

Metode transformacije su sledeće:

- ✓ Selektovati samo odgovarajuće tabele za unošenje (izbaciti null kolone),
- ✓ Prevođenje kodiranih podataka (npr. ako se u izvornom sistemu ženski pol obeležava sa *F*, muški sa *M*, a skladište podataka koristi obeležja 1 i 2),
- ✓ Šifrovanje slobodnih formi podataka (npr. mapiranje *Male* ili *Mr.* Kao 1),
- ✓ Stvaranje nove vrednosti (npr. iznos kupovine = količina x prodajna cena),
- ✓ Spajanje podataka iz raznih izvora,
- ✓ Sumiranje više redova podataka (npr. ukupna prodaja po regionu),
- ✓ Kreiranje vrednosti surogat ključeva, itd.

Punjenje

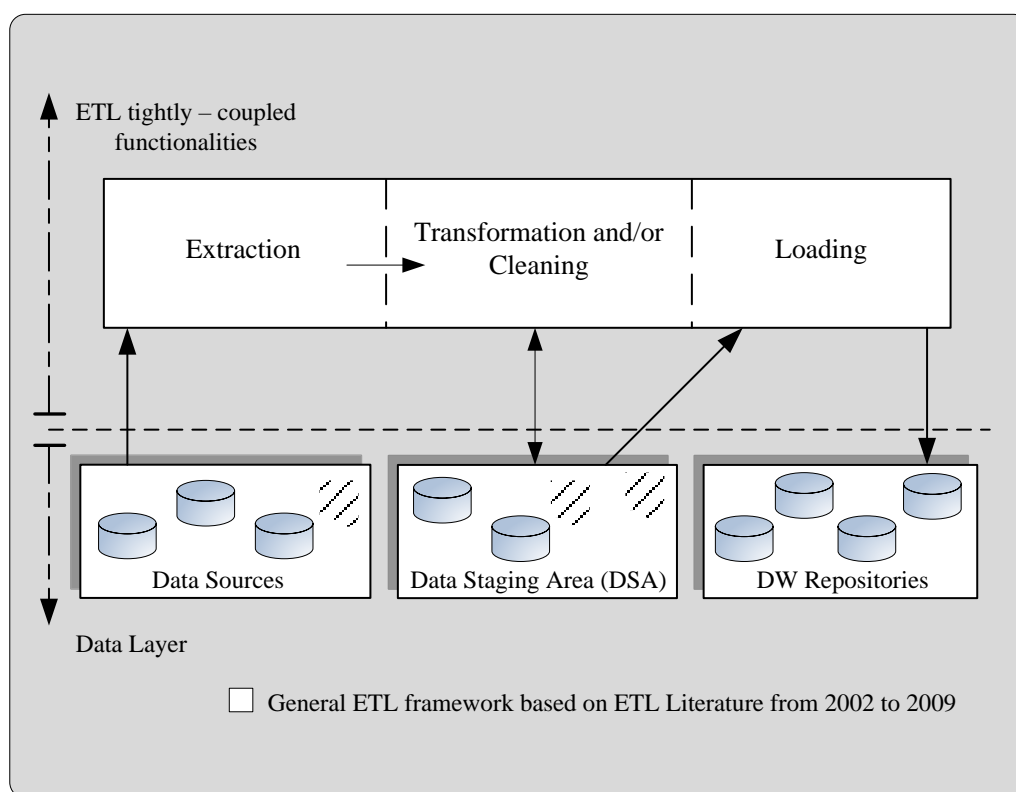
Za procese punjenja skladišta podataka se koristi više vrsta ETL alata, kao što su:

- ✓ programi za inicijalno punjenje,
- ✓ za punjenje istorijskih podataka,
- ✓ programi za inkrementalno punjenje.

Osnovne karakteristike programa za inicijalno punjenje skladišta podataka su rutine za čišćenje i usklađivanje podataka, kako bi se iz podataka uklonile greške. Kod istorijskih podataka ponekad nije moguće primeniti postupke čišćenja koji se primenjuju za "on-line" podatke, jer je od vremena nastanka tih podataka do sada možda došlo do različitih promena u slogovima i formatima podataka. Za tu kategoriju podataka primenjuju se programi za punjenje istorijskih podataka, koji čine nastavak inicijalnog punjenja. Za razliku od on-line, istorijski podaci su statičnog karaktera i sada čine samo sadržaj arhivskih datoteka. Treću vrstu predstavljaju programi za inkrementalno punjenje podataka, a aktiviraju se nakon što su prethodna dva programa odradila postupak čišćenja i usklađivanja podataka. Njihova karakteristika je da se pokreću periodično i oni predstavljaju stalno aktivan mehanizam punjenja skladišta podataka odgovarajućim sadržajima.

Tradicionalna arhitektura ETL procesa

ETL komponente u tradicionalnoj arhitekturi su čvrsto spregnute jedna sa drugom u okviru postojećih okruženja. Nedostatak mogućnosti za proširenje postojećeg ETL procesa rezultuje stvaranju prepreka prilikom dodavanja nove komponente, kada zahtevi poslovanja to zahtevaju. Veza između ključnih komponenata ETL procesa u tradicionalnoj arhitekturi je prikazana na sledećoj slici. Podaci se ekstrahuju iz različitih izvora, zatim se podaci transformišu i čiste, a potom smeštaju u privremeno skladište podataka (eng. Data Staging Area). Konačno, podaci se iz privremenog skladišta u fazi punjenja prebacuju u skladište podataka. Kao što se može videti na slici svaki od podprocesa se izvršava u okviru integrisanog okruženja. Ukoliko bi bilo potrebe da se ovo okruženje proširi, nastao bi problem zbog čvrste spregnutosti komponenti.

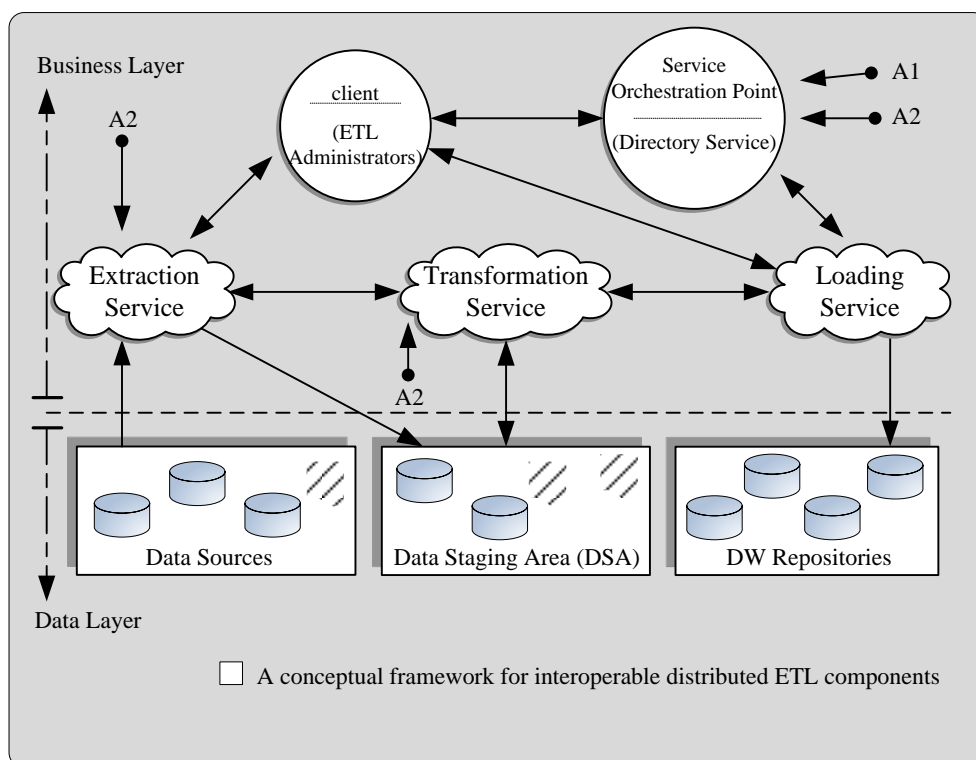


Slika 5.16. Tradicionalna ETL arhitektura

Distribuirana arhitektura ETL procesa

Distribuirana arhitektura ETL procesa se sastoji iz više samostalnih komponenti, međusobno nezavisnih koje komuniciraju jedna sa drugom, sa namerom postizanja zajedničkog cilja. Upravljanje ovim komponentama se obavlja uvođenjem novog koncepta orkestracije SOP (eng. Service Orchestration Point takođe poznata kao Directory Service ili

Service Registry). Servisi su podeljeni prema poslu koji obavljaju na ekstrakciju, transformaciju i punjenje. Oni nisu čvrsto spregnuti kao u tradicionalnoj arhitekturi ETL procesa. Šematski prikaz ove arhitekture je dat na sledećoj slici. Vidi se da su ovde procesi ekstrakcije, transformacije i punjenja podeljeni na međusobno nezavisne servise koji su povezani labavim vezama. Njima se upravlja principom orkestracije. Orkestracija predstavlja mehanizam kojim se upravlja komunikacijom i povezivanjem servisa u funkcionalnu celinu. Zahvaljujući ovakvom modelu distribuirane arhitekture, omogućava se prevazilaženje postojećih izazova vezanih za probleme proširenja koji su karakteristični za tradicionalnu arhitekturu ETL procesa, prema [7]:



Slika 5.17. Distribuirana arhitektura ETL procesa

Konkretizacija primene ETL procesa biće u nasatvku disertacije u delu koji se odnosi na projektvanje skladišta podataka o saobraćajnim nezgodama.

5.5. Otkrivanje zakonitosti u podacima

Otkrivanje zakonitosti u podacima ili data mining je proces automatskog otkrivanja korisnih informacija iz velikih količina podataka. Wikipedia daje dve definicije: Data mining je definisan kao *netrivijalno izvlačenje implicitnih, prethodno nepoznatih, i potencijalno korisnih informacija*, dok druga kaže za data mining da je *nauka izvlačenja korisnih informacija iz velikih skupova podataka ili baza podataka*.

Postojeći tehnološki trendovi neminovno dovode do poplave podataka. Sve više i više podataka se kreira npr., o saobraćajnim nezgodama, zatim bankarskim, telekomunikacijskim i drugim poslovnim transakcijama. Sve više podataka se generiše naučnim eksperimentima u astronomiji, biologiji, istraživanjem svemira... Sve više podataka se kreira na webu, posebno tekstovi, slike i drugi multimedijalni formati.

Na primer, procenjuje se da je samo u 2002. godini kreirano 5 miliona terabajta novih podataka. Prema analizama, duplo više informacija je generisano u 2002. nego u 1999., što dakle, predstavlja rast od 30% na godišnjem nivou. Predviđanja analitičara procenjuju još brži rast količine podataka. U svakom slučaju, jasno je da količine podataka rastu munjevito, a posledica tog rasta je da je veoma mala verovatnoća da će ti podaci ikad biti pregledani od strane nekog analitičara. Alati i algoritmi za otkrivanje znanja u podacima su dakle, potrebni kako bi se izvukao neki smisao, šablon i korist iz svih tih podataka. Takođe, oni obezbeđuju mogućnost da se predvide ishodi nekih budućih posmatranja, kao što je, na primer, predviđanje da li će i pod kojim okolnostima (uslovima) doći do saobraćajne nezgode.

Data mining se često meša sa 'pisanjem gomile izveštaja i upita', a zapravo *data mining* aktivnosti ne uključuju tradicionalno pisanje izveštaja, niti pravljenje upita. *Data mining* se vrši kroz specijalizovane alate, koji izvršavaju predefinisane data mining operacije bazirane na analitičkim modelima. Ova tehnologija spaja tradicionalne metode analize podataka sa sofisticiranim algoritmima za procesiranje velikih količina podataka i na taj način pruža sjajne mogućnosti za istraživanje i analizu novih tipova podataka, kao i za analizu starih tipova podataka, ali na potpuno nov način.

Data mining je analiza podataka, sa namerom da se otkriju dragocene skrivene informacije u masi kvantitativnih podataka koji su bili sakupljeni tokom normalnog ciklusa vođenja posla. Drugačiji je od konvencionalnih statističkih analiza, i to po osobinama koje se vide u sledećoj tabeli. Obe imaju svoje prednosti i slabosti, videti u [90]:

<i>Statističke analize</i>	<i>Data Mining</i>
Statističari obično kreću sa hipotezom (pitanjem ili pretpostavkom)	Data mining ne zahteva hipotezu
Statističari moraju da razviju svoje jednačine koje će odgovarati hipotezama.	Data mining algoritmi u okviru odgovarajućeg alata mogu automatski da razviju svoje jednačine
Statističke analize koriste samo numeričke podatke.	Data mining alati mogu koristiti različite tipove podataka, ne samo numeričke.
Statističari mogu da pronađu i filtriraju prljave podatke tokom njihovih analiza.	Data mining zavisi o čistim, uredno dokumentovanim podacima.
Statističari tumače svoje rezultate i prenose ih svojim izvršnim direktorima i menadžerima.	Data mining rezultate nije lako tumačiti, i statističar mora učestvovati takođe u analizi data mining rezultata i saopštavanju nalaza izvršnim direktorima i menadžerima.

Tabela 5.6. Poređenje Statističke analize i Data Mininga

Značaj data mining-a

Mnoge kompanije sakupljaju velike količine podataka, koji predstavljaju potencijalan izvor vrednih poslovnih informacija. Mogu se konstruisati analitički modeli, kako bi se pronašli šabloni u tim podacima i kako bi se dobijene informacije iskoristile za sticanje prednosti nad konkurentima. Na ovaj način analitičari, menadžeri i izvršni direktori dobijaju informacije nepohodne za preduzimanje daljih koraka, radi smanjenja saobraćajnih nezgoda, povećanje profita, smanje troškova, kao i kreiranje nove proizvodne strategije i povećanje učešća na tržištu.

Data mining, vođen otkrivanjem znanja iz podataka, nalazi odgovore na pitanja koja menadžeri ne znaju da postave. Zbog ove bitne sposobnosti, data mining čini bitan deo poslovne inteligencije. Može se reći i da data mining, takođe poznat pod nazivom otkivanje znanja iz podataka, koji predstavlja proces konvertovanja sirovih podataka u korisne informacije, je prekretnica u snabdevanju poslovne inteligencije strateški bitnim odlukama.

Ova tvrdnja može zvučati preterano, jer kad se pogleda, veliki broj sistema za podršku odlučivanju obezbeđuje bitne informacije poslovnoj inteligenciji, kao npr., videti [121]:

- ✓ Izvršni informacioni sistemi omogućavaju senior menadžerima da prate, ispituju i menjaju mnoge aspekte njihovih poslovnih operacija.
- ✓ Upiti i alati za izveštavanje daju poslovnim analitičarima mogućnost da ispituju performanse kompanije i ponašanje kupaca,
- ✓ Statistički alati omogućavaju statističarima da izvedu sofisticirane studije o toku i uspešnosti poslovanja
- ✓ Novi OLAP alati isporučuju mogućnost da se izvede 'šta ako' analiza, itd.

Mnogi od ovih alata rade sa aplikacijama poslovne logike i mogu ispitati veliki broj podataka. Šta je toliko razlika između ovih alata i data mining-a? Pa, velika razlika je to što sve tradicionalne tehnike analize, čak i one sofisticirane, oslanjaju na analitičara koji zna šta da traži u podacima. Analitičar kreira i pokreće upite bazirane na nekim hipotezama i pretpostavkama o mogućim vezama, trendovima i odnosima za koje se misli da će biti prezentovani u podacima. Slično, menadžeri se oslanjaju na poslovne poglede ugrađene u izvršne informacione sisteme, koji mogu ispitati samo one faktore za koje je alat programiran da izvidi i ispita. Kako problemi postaju kompleksniji i uključuju više promenljivih za analizu, ove tradicionalne tehnike postaju neefikasne. Nasuprot tome, data mining okrenut otkrivanju podataka, podržava suptilne i kompleksne istrage.

Za razliku od OLAP aplikacija i ad-hoc izveštavanja, čija je glavna funkcija analiza i izveštavanje na postojećim podacima, data mining aplikacije idu korak dalje - njihova je uloga pronalaženje *uzoraka i povezanosti* među podacima koji nisu očigledni. Nadovezujući se na takve uzorke, moguće je predvideti trendove i uzorke u budućnosti. Današnje kompanije suočene su sa eksponencijalnim rastom količina podataka, koje zahtevaju sve naprednije metode analize i obrade. Moćan softver za data mining zadužen je za pronalaženje uzoraka i trendova koji su značajni za pomoć pri definisanju poslovne strategije, pronalaženje povezanosti između raznih varijabli i pronalaženje zanimljivih segmenata i preseka informacija. Kombinacijom tehnika iz područja statističke analize, neuronskih mreža, tehnika modeliranja, inteligentnom analizom podataka, pronalaze se strukture i odnosi među njima, te se izvode pravila i modeli koji omogućuju predviđanje i odlučivanje u novonastalim situacijama. Osim pronalaženja tih veza, rešenja za data mining moraju pružiti i kvalitetnu i smislenu prezentaciju i vizualizaciju takvih informacija, omogućiti njihovo korišćenje u poslovnim procesima, kao što su agencije za bezbednost saobraćaja, marketinške agencije, analize profila kompanija ili analize proizvoda i programa.

Data mining je analitički proces namenjen ispitivanju podataka, kako bi se naišlo na konzistentne šablone i veze između promenljivih. Potom se te otkrivene veze i paterni ocenjuju primenjujući ih direktno na neki novi podskup podataka. Glavni cilj data mininga je predviđanje. Predviđanje u data miningu je najčešće korišćena tehnika u data miningu i ona koja ima konkretne poslovne aplikacije.

Koraci u data mining-u

Proces data mininga ima tri stadijuma:

- ✓ Početno istraživanje,
- ✓ Građenje modela, ili identifikacija paterna zajedno sa validacijom/verifikacijom, i
- ✓ Primena (npr. primena novog modela na nove podatke kako bi se generisala predviđanja).

Stadijum 1: Početno istraživanje. Ova faza obično počinje sa pripremom podataka, što može uključiti čišćenje podataka, transformaciju podataka, odabir podskupova podataka i u slučaju skupova podataka sa velikim brojem promenljivih (polja). Izvođenje nekih operacija preliminarne selekcije kako bi se veliki broj podataka sveo na neki manji opseg s kojim se može raditi (sve to u zavisnosti od statističkih metoda koje se planiraju primeniti). Potom, u zavisnosti od prirode problema, prva faza procesa data mininga može uključiti i jednostavan izbor prediktora za regresioni model, kao i kompleksne analize koje koriste veliki broj grafičkih i statističkih metoda kako bi se identifikovale najvažnije promenljive i utvrdila složenost i priroda modela koja će se uzeti u obzir u sledećem stadijumu.

Stadijum 2: Građenje modela i validacija. Tokom ove faze uzima se u obzir više različitih modela i bira se najbolji, na osnovu njihovih performansi (npr, u zavisnosti od toga koliko stabilne rezultate daje kroz primere). Ovo može zvučati kao jednostavna operacija, ali zapravo, ponekad je vrlo zahtevan proces.

Kako bi se postigao ovaj cilj, razvijen je veliki broj tehnika od kojih je većina zasnovana na takozvanom 'kompetitivnom vrednovanju modela', što predstavlja primenu različitih modela na isti skup podataka i potom poređenje njihovih performansi i biranje najbolje. Ove tehnike, koje se često smatraju osnovom predikativnog data mininga, uključuju: *Bagging* (Voting, Averaging), *Boosting*, *Stacking* (Stacked Generalizations), i *Meta-Learning*.

Stadijum 3: Primena. Poslednja faza uključuje korišćenje selektovanog najboljeg modela u prethodnom koraku i primenu istog na nove podatke kako bi generisao predviđanja ili procene očekivanih ishoda.

Koncept data mininga je postao popularan, kao poslovni alat gde se od njega očekuje da otkrije strukture znanja koje mogu voditi ka odlukama u uslovima ograničene sigurnosti. Nedavno, javila se potreba za razvojem novih analitičkih tehnika posebno dizajniranih kako bi odgovorile na relevantna pitanja poslovnog data mininga, ali sam data mining je i dalje baziran na statističkim principima i tradicionalnom istraživanju podataka i modeliranju, i deli i s jednim i s drugim komponente prilaza problemu i određene tehnike.

Ipak, glavna razlika između data mininga i tradicionalnog istraživanja podataka je ta, što je data mining više orijentisan ka aplikacijama nego ka osnovnoj prirodi problema koji leži ispod svega. Drugim rečima, data mining nešto manje obraća pažnju na određene veze između uključenih promenljivih. Na primer, otkrivanje prirode funkcija ili specifičnih tipova interaktivnih, multivarijantnih zavisnosti između varijabli nije primarni cilj data mininga. Umesto toga, fokus je na izradi rešenja koje može generisati korisna predviđanja. Tako da, data mining prihvata, između ostalog, koncept "crne kutije", kao prilaz istraživanju podataka i otkrivanju znanja i koristi ne samo tradicionalne tehnike istraživanja podataka, već i tehnike kao što su *Neuronske mreže* koje mogu generisati validna predviđanja, ali nisu u stanju da identifikuju specifičnu prirodu međurelacija između varijabli, na osnovu kojih su ta predviđanja napravljena.

Poreklo data mining-a

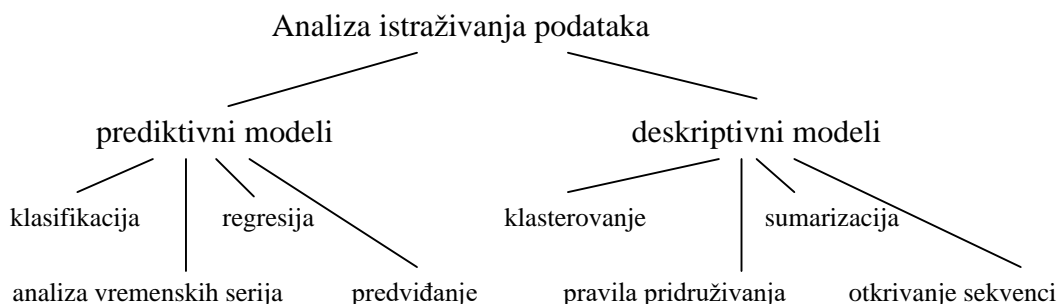
Istraživači iz različitih disciplina, želeći da razviju efikasnije i skalabilnije alate, koji će moći da koriste različite tipove podataka, su našli rešenje u data miningu, koji je zasnovan na metodologiji i algoritmima koji su istraživači ranije koristili. Konkretnije, zasniva se na idejama poput (1) *uzorkovanja i testiranja hipoteza* preuzetim iz statistike i (2) *algoritmima pretrage, tehnikama modelovanja i teoriji učenja* preuzetim iz veštačke inteligencije, prepoznavanja paterni i mašinskog učenja. Data mining je prihvatio ideje i iz drugih oblasti, uključujući optimizaciju, informacionu teoriju, procesiranje signala, vizuelizaciju i prikupljanje informacija. Pored ovih, i mnoge druge oblasti imaju pomoćne uloge. Baze podataka su potrebne kako bi obezbedile podršku za efikasno skladištenje, indeksiranje i procesiranje upita, detaljno u [13].

Zadaci data mining-a

Kao što je istaknuto, zadaci data mininga su generalno podeljeni na dve kategorije:

Predikativni zadaci. Cilj ovih zadataka je da predvide vrednost nekog određenog atributa na osnovu vrednosti drugih atributa. Atribut čija se vrednost predviđa se zove ciljna ili zavisna promenljiva, dok atributi na osnovu kojih se vrši ovo predviđanje se zovu ulazne, opisne ili nezavisne varijable.

Deskriptivni zadaci. Ovde je cilj identifikovati obrasce (korelacije, trendove, klasterne, trajektorije i anomalije) i odnose između podataka. Deskriptivne metode istražuju ispitivane osobine podataka, a ne predviđaju nove osobine.



Slika 5.18. Tehnike data mininga

Četiri glavna data mining zadatka su:

- ✓ model predviđanja,
- ✓ analiza pridruživanja,
- ✓ klaster analiza, i
- ✓ detektovanje anomalija.

Model predviđanja se odnosi na zadatak kreiranja modela u kome će ciljna promenljiva biti funkcija objašnjavajućih, opisnih varijabli. Postoje dva tipa ovog modela: *Klasifikacija*, koja se koristi za diskretne ciljne promenljive i *Regresija*, koja se koristi za kontinualne ciljne promenljive. Npr, predviđanje da li će neki učesnik u saobraćaju imati saobraćajnu nezgodu je zadatak klasifikacije, zato što je ciljna promenljiva binarna. S druge strane, predviđanje buduće cene akcija je zadatak regresije zato što je cena kontinualan atribut. Zadatak oba ova tipa je da naprave model koji minimizira grešku između predviđene i stvarne vrednosti ciljne promenljive. Model predviđanja može poslužiti kako bi se identifikovali kupci koji će odreagovati na marketinške kampanje, kako bi se predvidela pomeranja u Zemljinom ekosistemu ili prosudilo da li pacijent boluje od određene bolesti, u zavisnosti od rezultata medicinskog testiranja.

Analiza pridruživanja (analiza veza, analiza afiniteta) označava otkrivanje obrazaca koji opisuju međusobno čvrsto povezane osobine podataka. Obično se predstavljaju preko implikacije ili osobina podskupova. Kako prostor koji se pretražuje raste eksponencijalno, cilj je izdvojiti najinteresantnije obrasce na najefikasniji način.

Klasterovanje predstavlja proces grupisanje skupova fizičkih ili apstraktnih objekata u klase sličnih objekata. Klaster je kolekcija objekata koji su sličniji jedni drugima, nego što su to objekti u drugim klasterima. Na taj način se ova grupa objekata tretira kao jedan, čime se olakšavaju operacije nad njima. Naziva se i učenje bez nadzora jer klasteri nisu određeni pre ispitivanja podataka.

Detekcija anomalija je zadatak identifikovanja objekata čije karakteristike su značajno drugačije od ostalih objekata u skupu podataka. Takvi objekti se u statistici zovu autlajeri, pojave koja su numerički udaljene od ostalih podataka. Statistički podaci koji se dobijaju iz skupova podataka koji sadrže autlajere često dovode do pogrešnih zaključaka, Zadatak detekcije anomalija je otkrivanje pravih autlajera i na taj način izbegavanje pogrešnog obeležavanja normalnih podataka kao anomalija. Drugim rečima, dobar detector anomalija mora imati visoku stopu detekcije i nisku stopu greške. Koristi se u otkrivanju prevara, upada u mreže, neuobičajenih obrazaca bolesti.

U nastavku disertacije, u narednom poglavlju biće prikazana aplikativna strana skladišta podataka i primena koncepta otkrivanja zakonitosti u podacima. Već je istaknuto da će primer baze podataka biće o saobraćajnim nezgodama.

6. IMPLEMENTACIJA SISTEMA POSLOVNE INTELIGENCIJE

Kao što je u uvodu i u prethodnom poglavlju disertacije istaknuto, aplikativni deo sistema poslovne inteligencije biće predstavljen u ovom (šestom) poglavlju. Baza podataka tretira problematiku o saobraćajnim nezgodama u gradu Beogradu, što je bio deo projekta i zajedničkog rada nastavnika i saradnika Fakulteta organizacionih nauka i Saobraćajnog fakulteta Univerziteta u Beogradu, za naručioca, Skupštinu grada Beograda. Podaci su neimenovani i zaštićeni. Kao što je u prethodnom poglavlju na teoretskom nivou istaknuto, prvi korak ka tome jeste razvoj skladišta podataka i aktiviranje ETL procesa.

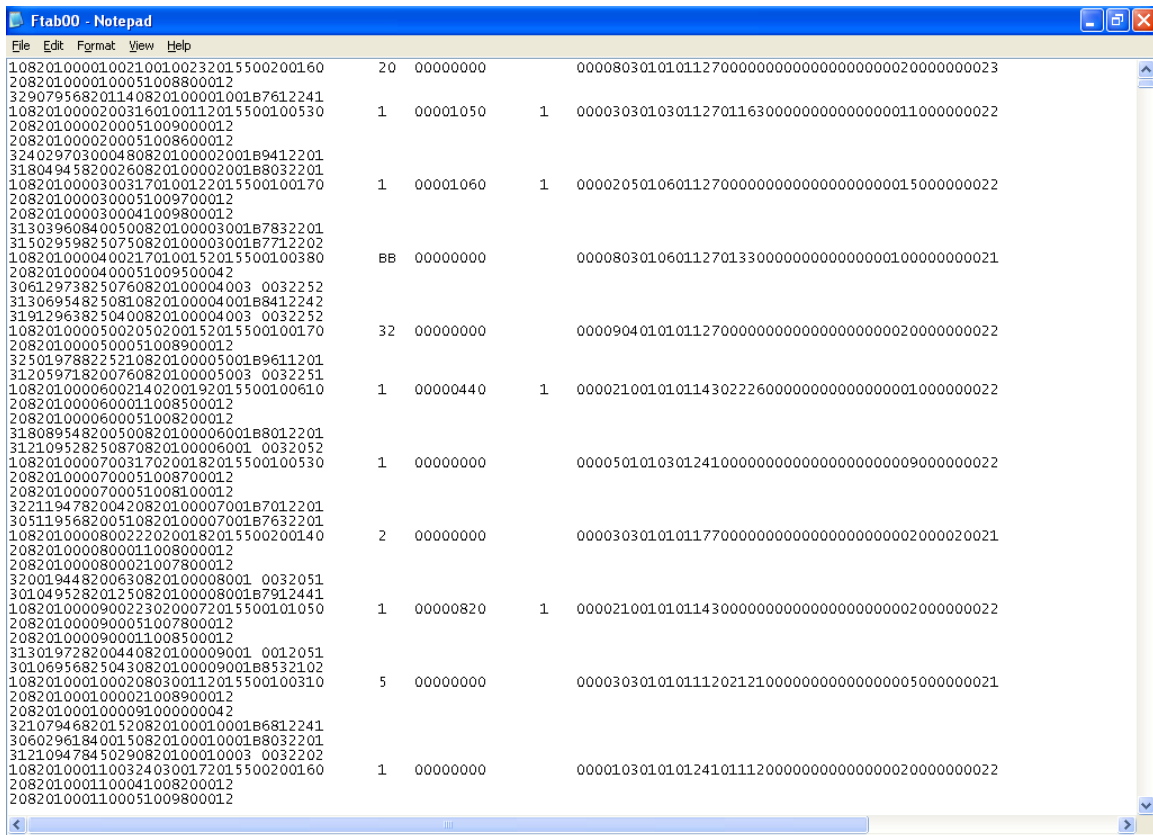
6.1. Razvoj skladišta podataka

Prvi korak u razvoju skladišta podataka jeste, ekstrahovanje, transformacija i učitavanje relevantnih podataka u skladište podataka a shodno opisu datom u prethodnom poglavlju. Sami podaci o saobraćajnim nezgodama se mogu preuzimati iz više izvora, kako je to već napomenuto, pre svega internih i eksternih. Konkretno za ovaj slučaj to su:

- ✓ MUP, što predstavlja primarni izvor podataka;
- ✓ Zdravstvene ustanove;
- ✓ Osiguravajuće kompanije;
- ✓ Direkcija za puteve;
- ✓ Statistički centri, i
- ✓ Razna druga istraživanja, koja prate problematiku bezbednosti saobraćaja.

U nastavku se opisuje unos i transformacija podataka preuzetih pre svega od MUPa Beograd, što predstavlja osnovu za projektovanje i implementaciju skladišta podatak.

Sama baza podataka o saobraćajnim nezgodama, na području grada Beograda, preuzima se u *izvoznom *.txt* formatu. Njen osnovni izgled prikazan je na sledećem ekranskom formom.



Slika 6.1. Sirovi podaci preuzeti iz MUPa Beograd

Kao što se vidi sa prethodne ekranske forme (Slika 6.1.), jedan zapis ima sledeći izgled:

```
10820100001002100100232015500200160          20      00000000
000080301010112700000000000000000000000020000000023
2082010000100051008800012
329079568201140820100001001B7612241
```

Važno je istaći da su ključni akteri o mogućim saobraćajnim nezgodama *triplet*, sama *nezgoda*, zatim *vozilo*, zatim *lice*. Kao što se vidi sa prethodnog zapisa, podaci o jednoj saobraćajnoj *nezgodi* šifrirani su sa brojem **1**, za *vozilo* sa brojem **2**, i za *lice* sa brojem **3**, tako da se svaki saobraćajni akcident može jedinstveno i jednoznačno opisati.

Na taj način su upotpunosti definisani potrebni podaci za evidentirana tri objekta, i to **Nezgoda, Vozilo i Lice**.

Za ovako organizovane sirove podatke kreirana je automatizovana procedura za njihovu transformaciju u analitičku bazu podataka.

Procedura se sastoji iz sledećih koraka:

1. Izgradnja baze SQL Server.
2. Učitavanje svih nomenklatura MUPa u bazu.
3. Razdvajanje podataka iz MUP tekstualnog fajla u tri nova tekstualna fajla (Nezgoda, Vozilo i Lice).
4. Definisane strukture tekstualnih fajlova Nezgoda, Vozilo i Lice.
5. Definisane tipa podataka u tekstualnim fajlovima.
6. Provera validnosti podataka.
7. Upisivanje tekstualnih fajlova u odgovarajuće tabele SQL Servera.
8. Formiranje konekcije ka SQL Server bazi podataka.
9. Formiranje strukture analitičke baze podataka.
10. Učitavanje podataka u analitičku bazu.
11. Procesiranje analitičke baze.
12. Izvoz analitičke baze.

Važno je napomenuti da Koraci 1, 2, 8 se preduzimaju, u načelu samo jednom, i to onda kada se vrši po prvi put definisanje analitičke baze podataka (kocke). Eventualno, Korak 1 može da se ponavlja u zavisnosti od zahteva za dodavanjem novih podataka za analizu.

U nastavku se daje opis i konkretizacija skladišta podataka shodno navedenim koracima.

6.1.1. Izgradnja baze SQL Server

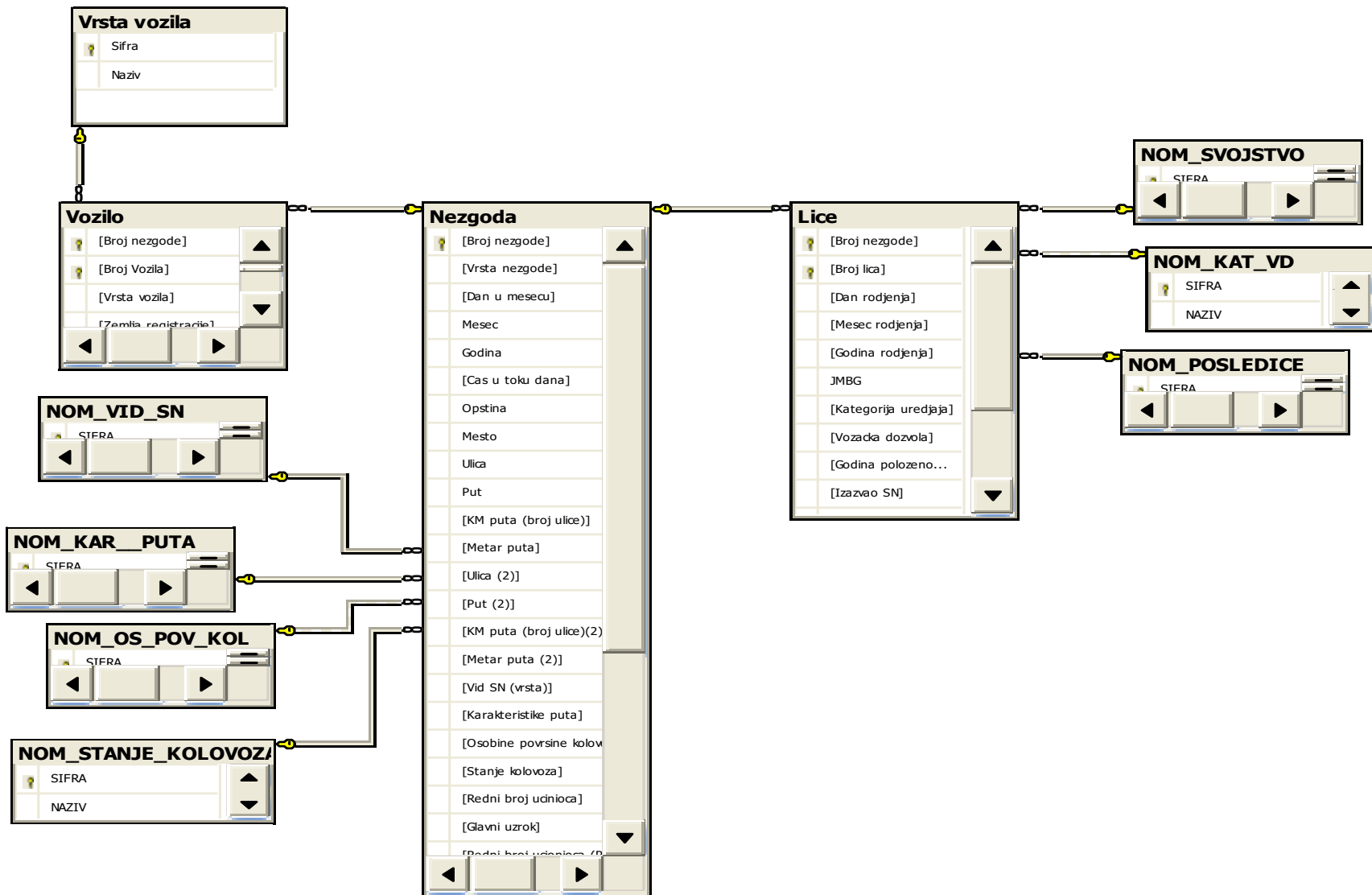
Na sledećoj slici (Slika 6.2.) se prikazuje sama struktura baze podataka koja može da prihvati podatke iz MUPa.

Već je istaknuto da su osnovne tri tabele:

- ✓ Nezgoda,
- ✓ Vozilo, i
- ✓ Lice.

Iste su povezane sa odgovarajućim nomenklaturama:

1. Vid SN, Stanje kolovoza, Osobine površine kolovoza, Karakteristike puta sa tabelom Nezgoda.
2. Vrsta vozila sa tabelom Vozilo.
3. Svojstvo, Kategorija vozačke dozvole i Posledica sa tabelom Lice.



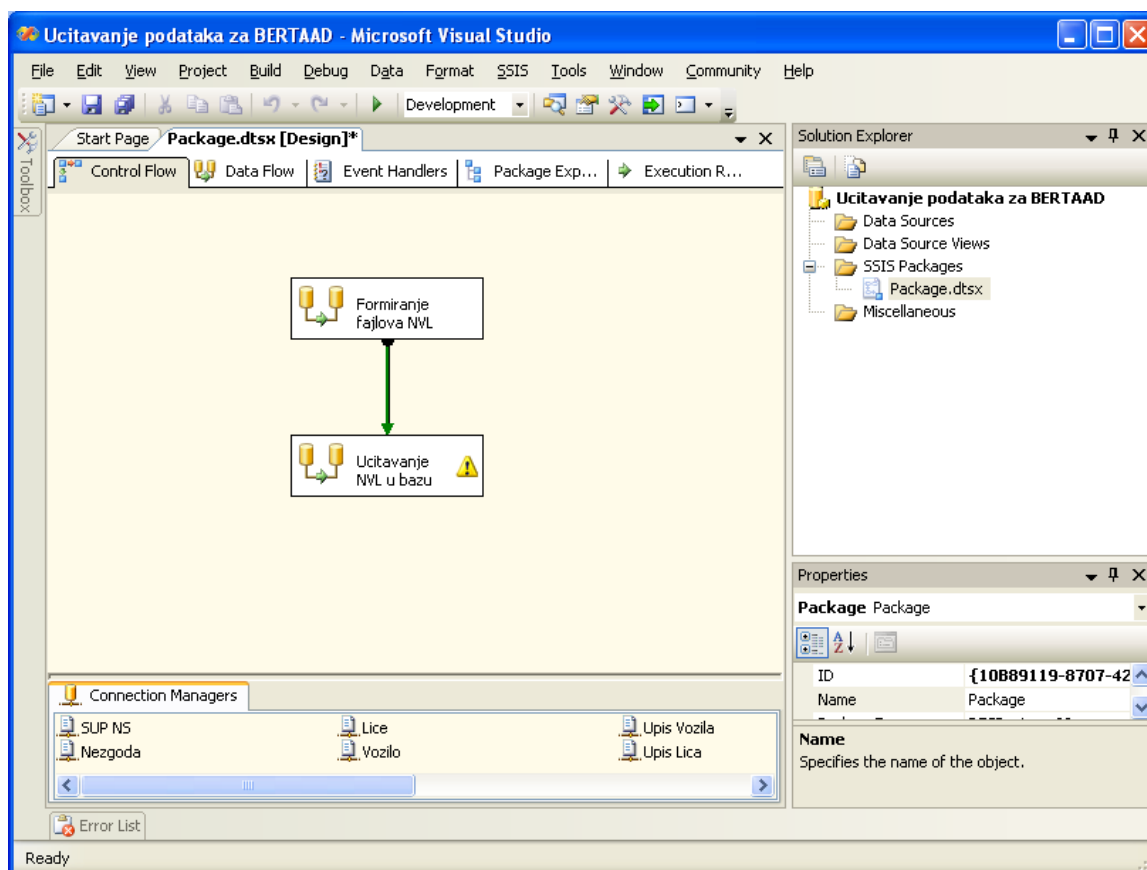
Slika 6.2. Dijagram objekti veze baze SQL Server

6.1.2. Učitavanje nomenklatura (šifarnika) MUPa u bazu

Nakon što je formirana baza podataka, potrebno je u nju učitati odgovarajuće šifarnike MUPa, kako bi podaci dobili odgovarajuće jednoznačno značenje. Nomenklature su učitanje za deset dobijenih šifarnika preuzetih iz MUPa. To su: Opština nezgode, Mesto nezgode, Ulica nezgode, Put nezgode, Glavni uzrok, Ishod (Prijava), Zemlja registracije, Vlasnik (korisnika) vozila, Javni prevoz, i Pol. Svi pomenuti atributi definisani su u nekoj od tri osnovne pomenute tabele.

6.1.3. Razdvajanje podataka iz tekstualnog fajla u tri nova tekstualna fajla (Nezgoda, Vozilo i Lice)

Razdvajanje podataka je automatizovana procedura koja se radi u SQL Server Integration Service okruženju. Na sledećoj ekranskoj formi, prikazuju se objekti razdvajanja i učitavanja.



Slika 6.3. Procedura učitavanja podataka u SQL Server

Učitavanje podataka u SQL Server se sastoji iz dve faze:

Faza 1,

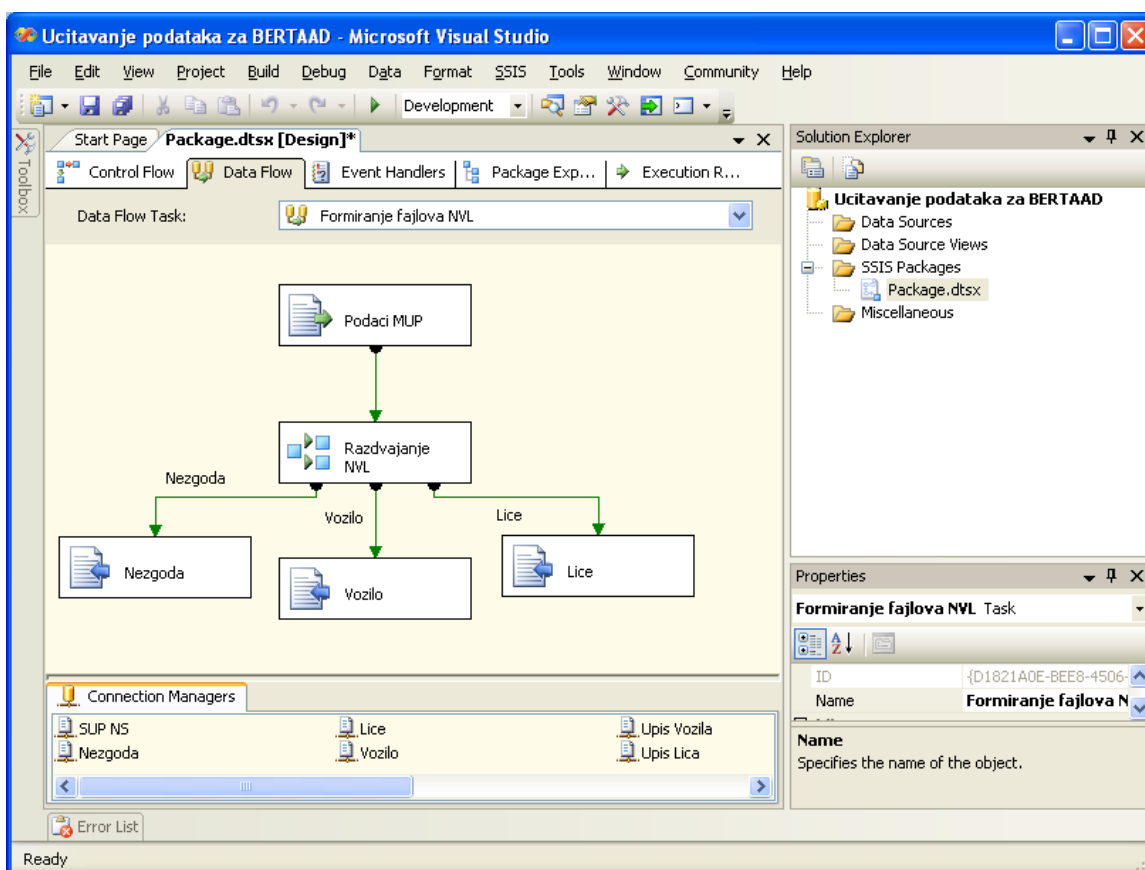
- ✓ Razdvajanje MUP-ovog tekstualnog fajla na tri tekstualna fajla: Nezgoda, Lice i Vozilo, i

Faza2,

- ✓ Učitavanje tri navedena fajla u SQL Server.

Faza 1 se sastoji iz tri koraka (Slika 6.4):

- ✓ Uspostavljanje konekcije sa MUP tekstualnim fajlom;
- ✓ Razdvajanje podataka po identifikacionom ključu, i
- ✓ Učitavanje podataka u odgovarajuće tekstualne fajlove.



Slika 6.4. Koraci transformacije podataka na tri definisana fajla

Učitavanje podataka u SQL Server se sastoji od dva koraka (Slika 6.5):

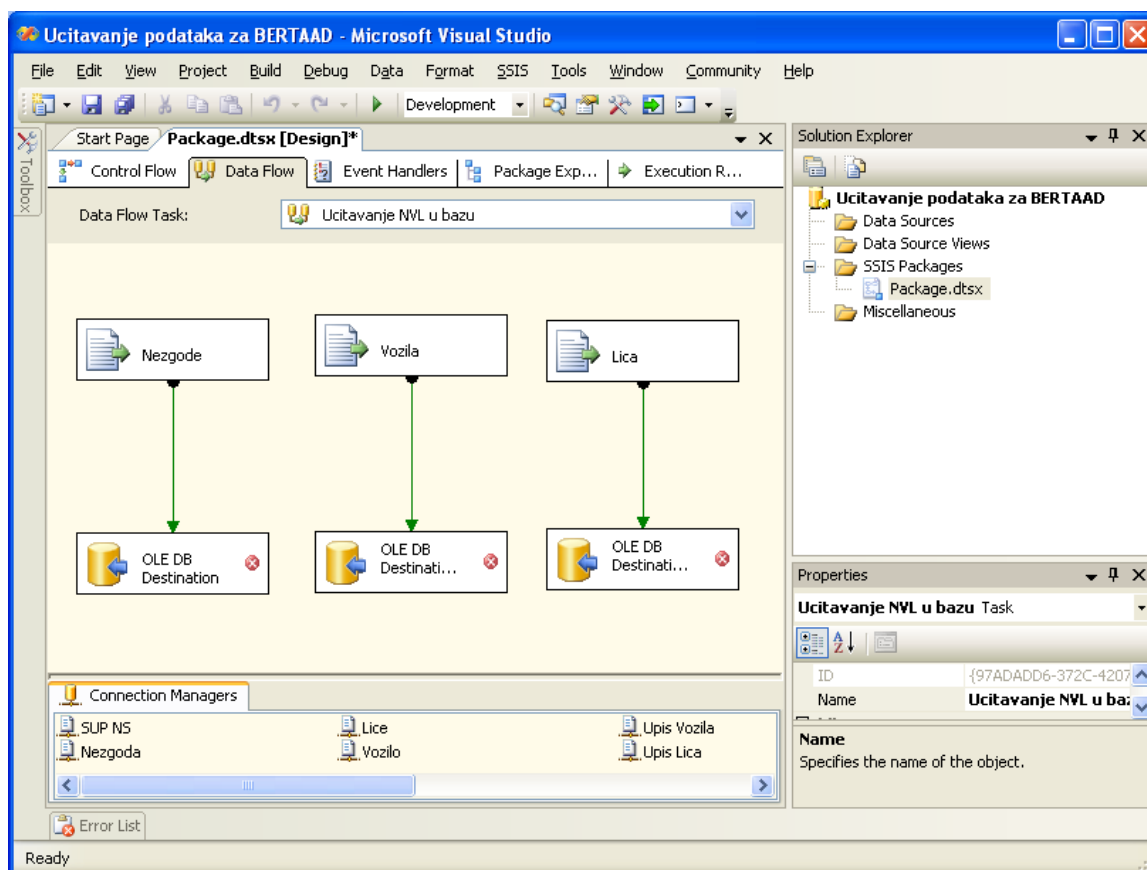
- ✓ Uspostavljanje konekcije sa tekstualnim fajlovima Nezgoda, Vozilo i Lice.
- ✓ Učitavanje navedenih fajlova u tabele SQL Servera.

6.1.4. Definisanje strukture tekstualnih fajlova

Nezgoda, Vozilo i Lice

Svaki slog fajlova Nezgoda, Vozilo i Lice je strukturiran i se sastoji iz odgovarajućih podataka koji zauzimaju određeni broj mesta u bazi podataka. Bilo je bitno i potrebno odrediti taj broj mesta (širina polja), da bi podaci mogli da se tumače i koriste na odgovarajući analitički način.

Širina jednog sloga koji se generiše iz sva tri fajla jeste ukupno **164** karaktera.



Slika 6.5. Učitavanje tekstualnih fajlova Nezgoda, Lice i Vozilo u SQL Server

Fajl **Nezgoda** se sastoji od zapisa sa poljima sledeće strukture:

Naziv podatka	Širina podatka
Tip	1
Broj nezgode	12
Vrsta nezgode	1
Dan nezgode	2
Mesec nezgode	2
Godina nezgode	2
Čas nezgode	2
Opština nezgode	5
Mesto nezgode	3
Ulica nezgode	5
Put nezgode	6
Km puta, broj ulice	4
Metar puta	3
Ulica nezgode 2	5
Put nezgode 2	6
Km puta, broj ulice 2	4
Metar puta 2	3
Vid nezgode	2
Karakteristike nezgode	2
Osobine površine kolovoza	2
Stanje kolovoza	2
Redni broj učinioca	2
Glavni uzrok	3
Redni broj učinioca	2
Posredni uzrok	3
Redni broj učinioca	2
Prateći uzrok	3
Šteta	15
Prekid sat	2
Prekid minut	2
NN nezgoda	1
Ishod (Prijava)	1
Višak (neki nestrukturirani ostali podaci)	54

Tabela 6.1. Atributi zapisa nezgoda

Fajl **Vozilo** se sastoji od zapisa sa poljima sledeće strukture:

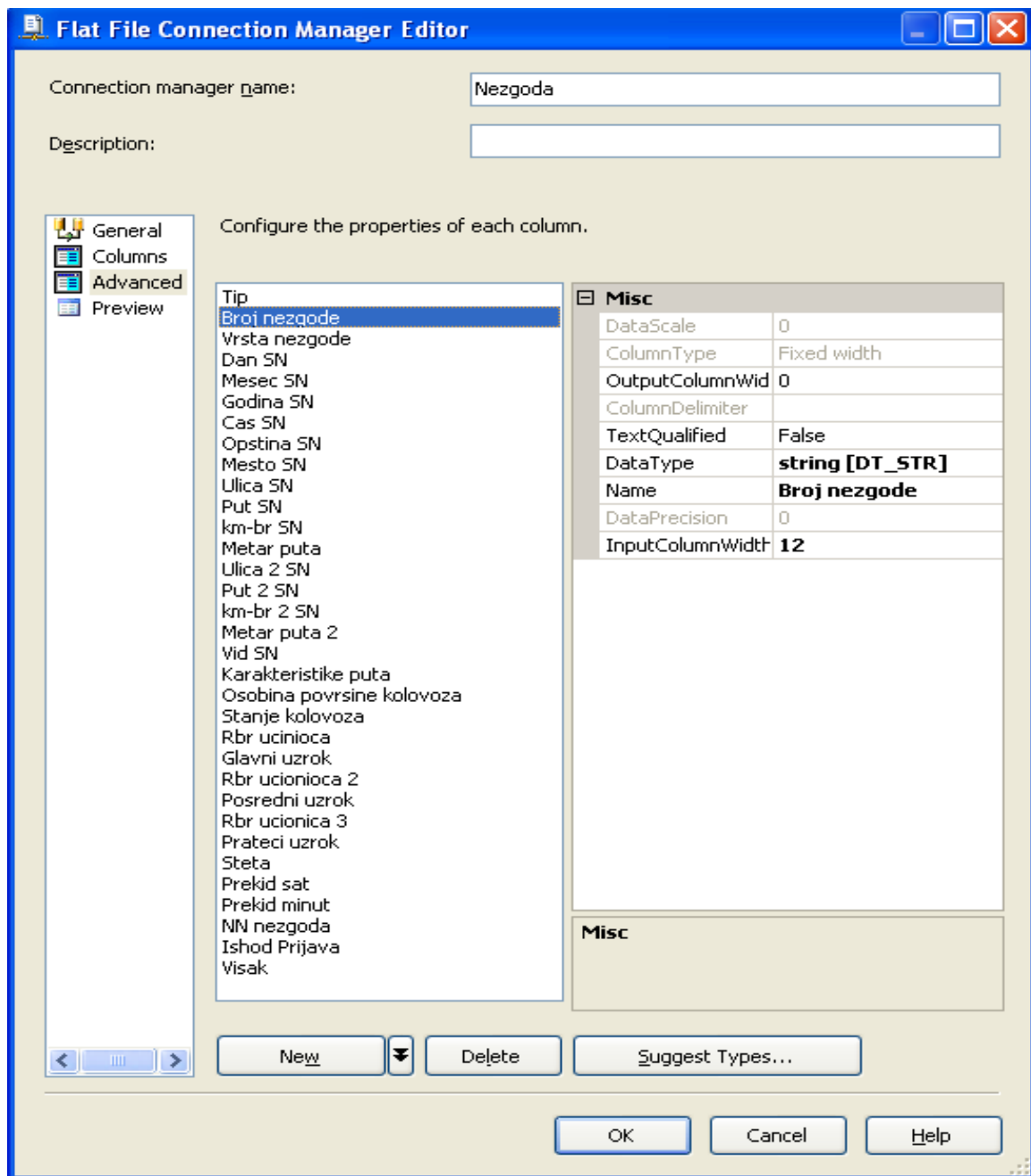
Naziv podatka	Širina podatka
Tip	1
Broj nezgode	12
Vrsta vozila	2
Zemlja registracije	3
Godina proizvodnje	2
Vlasnik (korisnik) vozila	4
Javni prevoz	1
Višak (neki nestrukturirani ostali podaci)	139

Tabela 6.2. Atributi zapisa vozilo

Fajl **Lice** se sastoji od zapisa sa poljima sledeće strukture:

Naziv podatka	Širina podatka
Tip	1
Dan rođenja	2
Mesec rođenja	2
Godina rođenja	3
JMBG	6
Broj nezgode	12
Svojstvo lica	1
Kategorija vozačke dozvole	1
Godina polaganja vozačkog ispita	2
Izazvao nezgodu	1
Pobegao sa mesta SN	1
Posledice	1
Pol	1
Višak (neki nestrukturirani ostali podaci)	129

Tabela 6.3. Atributi zapisa lice



Slika 6.6. Izgled ekrana za definisanje širine polja, strukture i tipa podataka fajla Nezgoda

6.1.5. Definisanje tipa podataka u tekstualnim fajlovima

Sa slike 6.6., se vidi ekran u kome se definiše struktura zapisa tekstualnih fajlova, veličina polja zapisa, kao i odgovarajući tip podataka.

Treba naglasiti da svakom od navedenih podataka se dodeljuje odgovarajući tip podataka, da bi mogle da se vrše odgovarajuće analize. Jer vrste analiza koje se zahtevaju, određuju prevashodno tipovi podataka odgovarajućih polja zapisa.

6.1.6. Provera validnosti podataka (logička kontrola)

Za pojedine podatke potrebno je izvršiti logičku kontrolu (validnosti), kako bi analitička baza podataka bila popunjena ispravnim podacima.

Konkretno, za sledeće podatke je potrebno uraditi proveru validnosti:

Naziv podatka NEZGODA	Provera
Vrsta nezgode	da li su podaci u nomenklaturi
Dan nezgode	da li su podaci u intervalu [1,31]
Mesec nezgode	da li su podaci u intervalu [1,12]
Čas nezgode	da li su podaci u intervalu [0, 24]
Opština nezgode	da li su podaci u nomenklaturi
Mesto nezgode	da li su podaci u nomenklaturi
Ulica nezgode	da li su podaci u nomenklaturi
Put nezgode	da li su podaci u nomenklaturi
Vid nezgode	da li su podaci u nomenklaturi
Karakteristike nezgode	da li su podaci u nomenklaturi
Osobine površine kolovoza	da li su podaci u nomenklaturi
Stanje kolovoza	da li su podaci u nomenklaturi
Glavni uzrok	da li su podaci u nomenklaturi
Posredni uzrok	da li su podaci u nomenklaturi
Prateći uzrok	da li su podaci u nomenklaturi
NN nezgoda	da li su podaci u nomenklaturi
Ishod (Prijava)	da li su podaci u nomenklaturi

Naziv podatka VOZILO	Širina
Vrsta vozila	da li su podaci u nomenklaturi
Zemlja registracije	da li su podaci u nomenklaturi
Vlasnik (korisnik) vozila	da li su podaci u nomenklaturi
Javni prevoz	da li su podaci u nomenklaturi

Naziv podatka LICE	Širina
Svojstvo lica	da li su podaci u nomenklaturi
Kategorija vozačke dozvole	da li su podaci u nomenklaturi
Izazvao nezgodu	da li su podaci u nomenklaturi
Pobegao sa mesta SN	da li su podaci u nomenklaturi
Posledice	da li su podaci u nomenklaturi
Pol	da li su podaci u nomenklaturi

Podatak za koji nije zadovoljen prethodno definisani uslov, se čekira, i stavlja u *listu podataka* koja označava da ne postoji precizna informacija o validnosti istoga. Svi podaci iz ove liste ponovo se proveravaju, da bi se doveli u konzistentno stanje.

6.1.7. Upisivanje tekstualnih fajlova u odgovarajuće tabele SQL Servera

Nakon što je izvršena priprema za učitavanje podataka u bazu SQL Server, moguće je izvršiti istu proceduru. Proces je upotpunosti automatizovan i prikazan je na sledećim Slikama 6.3., 6.4., i 6.5. Tako ekstrahovana baza ima sledeći izgled.

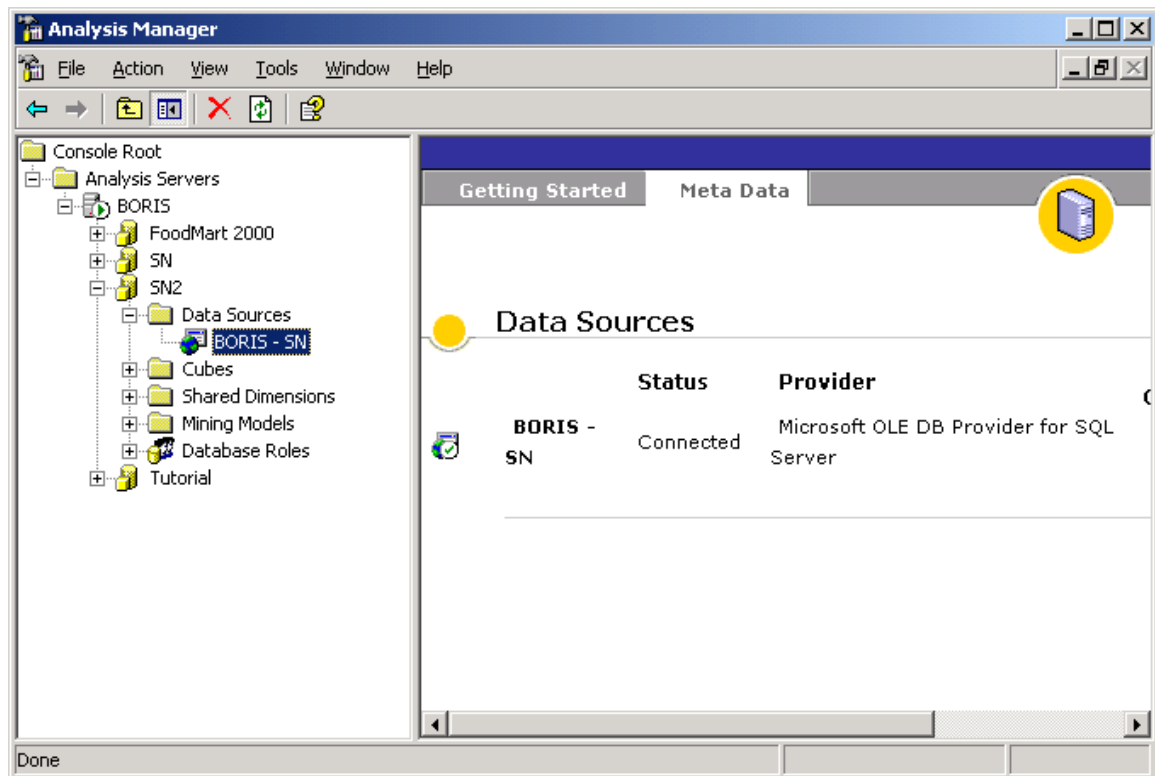
Broj nezgode	Vrsta nezgode	Dan u mesecu	Mesec	Godina	Cas u toku dana	Opstina	Mesto	Ulica	Put
082010000100	2	10	01	00	23	20155	002	00160	
082010000200	3	16	01	00	11	20155	001	00530	
082010000300	3	17	01	00	12	20155	001	00170	
082010000400	2	17	01	00	15	20155	001	00380	
082010000500	2	05	02	00	15	20155	001	00170	
082010000600	2	14	02	00	19	20155	001	00610	
082010000700	3	17	02	00	18	20155	001	00530	
082010000800	2	22	02	00	18	20155	002	00140	
082010000900	2	23	02	00	07	20155	001	01050	
082010001000	2	08	03	00	11	20155	001	00310	
082010001100	3	24	03	00	17	20155	002	00160	
082010001200	2	06	04	00	05	20155	001	00410	
082010001300	3	30	04	00	14	20155	005	00360	
082010001400	3	02	05	00	17	20155	001	00530	
082010001500	3	12	05	00	14	20155	001	00860	
082010001600	3	28	05	00	06	20155	002	00160	
082010001700	2	31	05	00	09	20155	001	00610	
082010001800	2	06	06	00	15	20155	001	00820	
082010002000	2	09	06	00	21	20155	001	00340	
082010002100	2	09	06	00	21	20155	001	00340	
082010002200	2	11	06	00	18	20155	001	00610	
082010002300	3	19	06	00	19	20155	001	00530	
082010002400	2	30	06	00	23	20155	002	00160	
082010002500	2	24	07	00	16	20155	001	00470	
082010002600	3	25	07	00	11	20155	002	00150	

Slika 6.7. Napunjena baza SQL Servera izvornim podacima

6.1.8. Formiranje konekcije ka SQL Server bazi podataka

Da bi podaci mogli da se koriste za pravljenje analitičke baze podataka, potrebno je uspostaviti konekciju između Microsoft Analysis Service i SQL Servera.

Uspostavljena veza je u potpunosti prikazana na Slici 6.8.



Slika 6.8. Veza analitičke baze sa SQL Serverom

6.1.9. Formiranje strukture analitičke baze podataka

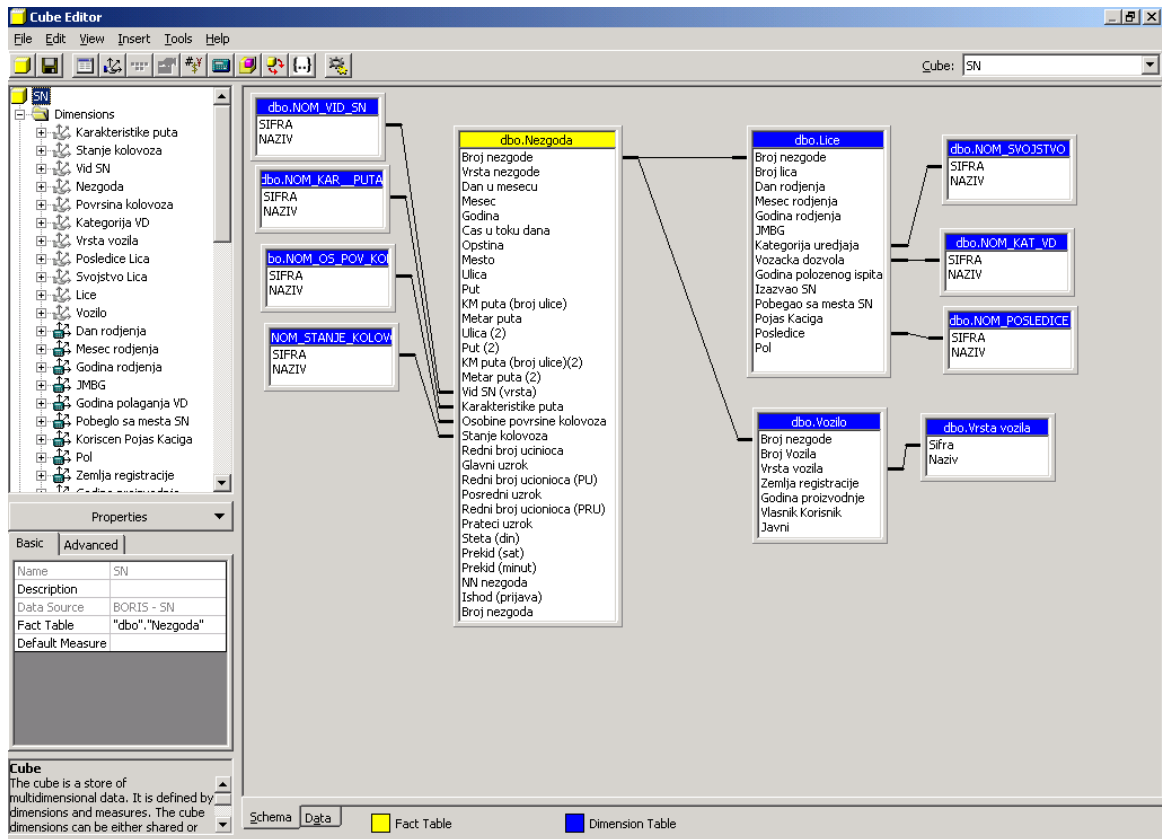
Potreban uslov da se izgradi analitička baza podataka, jeste da se definiše njena struktura, i to konkretno preciziranje dimenzija i mera baze.

Na Slici 6.9., se vidi struktura analitičke baze. Definisana su tri podatka preko kojih se vrši analiza svih posmatranih polja, i to:

- ✓ Ukupan broj saobraćajnih nezgoda,
- ✓ Ukupan broj vozila učesnika u tim nezgodama, i
- ✓ Ukupan broj lica učesnika u nezgodama.

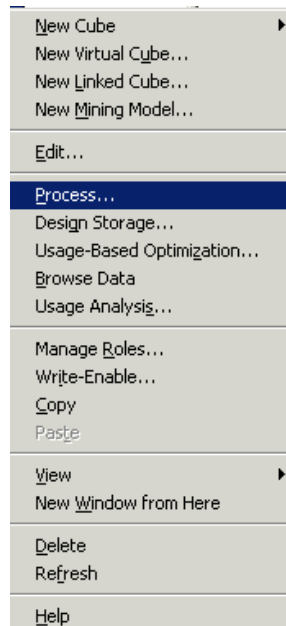
6.1.10. Učitavanje i procesiranje podataka u analitičkoj bazi

Podaci se u analitičku bazu povlače u trenutku kada se uključi opcija za procesiranje baze. Do tog trenutka analitička baza ima samo svoju strukturu.

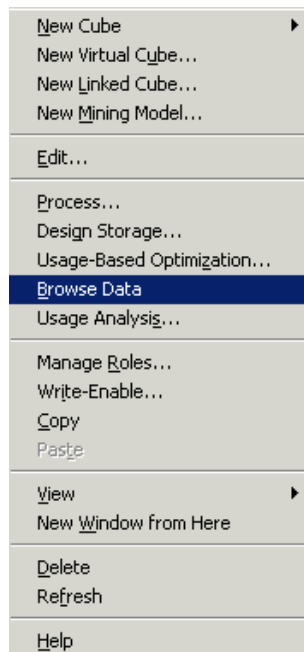


Slika 6.9. Struktura analitičke baze podataka

Procesiranje se vrši naredbom prikazanom na Slici 6.10.



Slika 6.10. Procesiranje baze podataka



Slika 6.11. Naredba za pristup podacima analitičke baze

Kada se završi procesiranje, baza je spremna za upotrebu (Slika 6.11.), tj. postavljanje najraznovrsnijih upita, što je tema narednog poglavlja. Baza može da se pristupi iz Microsoft Analysis Service (naredba na Slici 6.11.) ili iz nekog drugog programa, kao što je Microsoft Excel.

6.1.11. Izvoz analitičke baze podataka

Nakon što je analitička baza procesuirana, spremna je za izvoz u razne alate (programe), od kojih je svakako najpopularniji Microsoft Excel.

Iz navedenog programa je moguće formirati tzv. oflajn (offline) analitičke baze podataka u formi fajla, koji je moguće učitati na svakom računaru na kome postoji instaliran Excel.

Izgled jedne od formi je dat na narednoj Slici 6.12.

Analitičar ili donosilac odluke je sada u mogućnosti da dobije najraznovrsnije izveštaje jednostavnim prevlačenjem polja u Excel tabelu.

Izveštaje može da se generiše (Slika 6.12.) u tabelarnom i/ili grafičkom obliku, da se štampa, analizira, formatira, itd.

Ovim je završen prikaz transformacije i ekstrakovanja podataka za analitičku bazu podataka.

	A	B	C	D	E
1	Drop Page Fields Here				
2					
3	Data	Posledice Lica	Total		
4	Broj Lica	LAKE TELESNE POVREDE	1127		
5		LICE NEPOVREDJENO	19044		
6		SMRT DO 30 DANA	8		
7		SMRT NA LICU	55		
8		SMRT ZA VREME PREVOZA DO BOLNICE	12		
9		TESKE TELESNE POVREDE	464		
10	Broj Nezgoda	LAKE TELESNE POVREDE	1055		
11		LICE NEPOVREDJENO	11379		
12		SMRT DO 30 DANA	8		
13		SMRT NA LICU	54		
14		SMRT ZA VREME PREVOZA DO BOLNICE	12		
15		TESKE TELESNE POVREDE	455		
16	Broj Vozila	LAKE TELESNE POVREDE	1812		
17		LICE NEPOVREDJENO	21129		
18		SMRT DO 30 DANA	12		
19		SMRT NA LICU	91		
20		SMRT ZA VREME PREVOZA DO BOLNICE	22		
21		TESKE TELESNE POVREDE	759		
22	Total Broj Lica		20710		
23	Total Broj Nezgoda		12133		
24	Total Broj Vozila		22071		

Slika 6.12. Pregled ukupnog broja nezgoda, vozila i lica razvrstanih po posledicama lica

Sa prethodne slike se vide izvedeni (sumirani) podaci u koloni C, od reda 22 do reda 24.

6.2. Rad sa analitičkom bazom podataka o saobraćajnim nezgodama

Postupak rada u analitičkoj bazi saobraćajnih nezgoda biće objašnjen u okruženju:

- ✓ Microsoft Analysis Service, i
- ✓ Microsoft Excel.

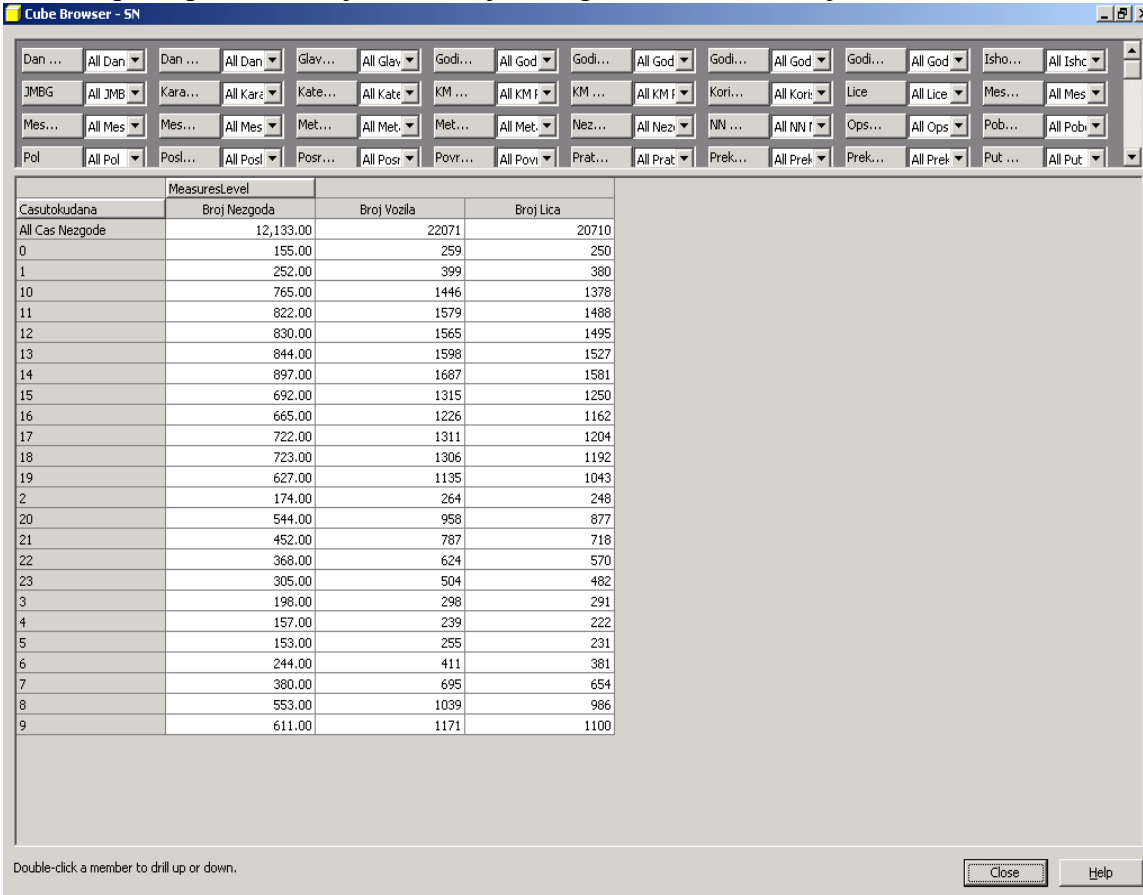
Rad sa bazom podataka moguć je u okruženju u kome je nastala, tj. Microsoft Analysis Service, kao i u jednom broju drugih aplikativnih softvera koji podržavaju OLAP (Online Analytical Processing) tehnologiju. Kao izuzetan alat za analizu, izabran je Microsoft Excel u kome mogu da se pokažu pogodnosti korišćenja analitičke baze podataka.

Naredni deo teksta će biti strukturiran u nekoliko celina:

1. Rad sa bazom u Microsoft Analysis Service, i
2. Rad sa bazom u Microsoft Excel.

6.2.1. Rad sa bazom u Microsoft Analysis Service

Kada se pristupi analitičkoj bazi, dobija se izgled kao na sledećoj slici:



Casutokudana	Broj Nezgoda	Broj Vozila	Broj Lica
All Cas Nezgode	12,133.00	22071	20710
0	155.00	259	250
1	252.00	399	380
10	765.00	1446	1378
11	822.00	1579	1488
12	830.00	1565	1495
13	844.00	1598	1527
14	897.00	1687	1581
15	692.00	1315	1250
16	665.00	1226	1162
17	722.00	1311	1204
18	723.00	1306	1192
19	627.00	1135	1043
2	174.00	264	248
20	544.00	958	877
21	452.00	787	718
22	368.00	624	570
23	305.00	504	482
3	198.00	298	291
4	157.00	239	222
5	153.00	255	231
6	244.00	411	381
7	380.00	695	654
8	553.00	1039	986
9	611.00	1171	1100

Slika 6.13. Pregled distribucije nezgoda, vozila i lica po času nezgode

Željeni izveštaj može da se formira na sledeći način:

Na početku analitičar ili donosilac odluke, selektuje kolone ili redove polja koja su najinteresantnija za izveštaj. To se radi jednostavnim „prevlačenjem“ polja iz gornjeg dela ekrana sa prethodne slike (Slika 6.13.). Moguća demonstracija je na sledeći način. Kada se doda svojstvo lica u red gde se već nalazi čas nezgode, dobija se izveštaj kao na sledećoj slici.

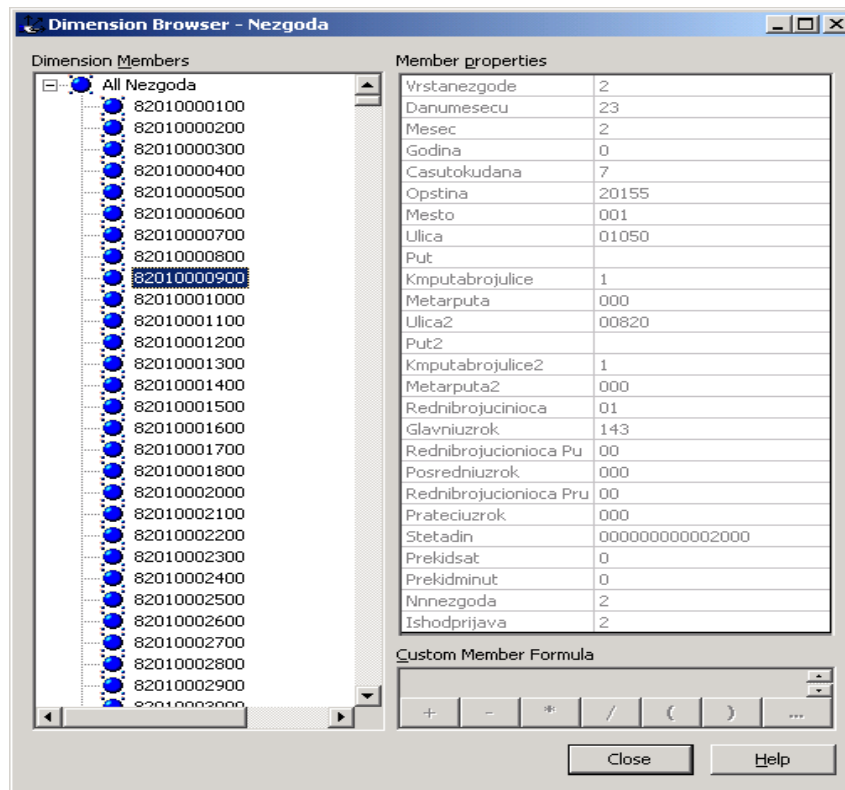
Casutokudana	Svojstvo Lica	MeasuresLevel		
		Broj Nezgoda	Broj Vozila	Broj Lica
0	All Svojstvo Lica	155.00	259	250
	jahac		0	0
	pešak		0	0
	putnik		0	0
	teraoc stoke		0	0
	vozac	155.00	259	250
1	All Svojstvo Lica	252.00	399	380
	jahac		0	0
	pešak	1.00	2	1
	putnik	1.00	2	1
	teraoc stoke		0	0
	vozac	252.00	399	378
10	All Svojstvo Lica	765.00	1446	1378
	jahac		0	0
	pešak	1.00	1	1
	putnik		0	0
	teraoc stoke		0	0
	vozac	765.00	1446	1377
11	All Svojstvo Lica	822.00	1579	1488
	jahac		0	0
	pešak	1.00	2	1
	putnik	1.00	2	1
	teraoc stoke		0	0
	vozac	822.00	1579	1486
12	All Svojstvo Lica	830.00	1565	1495
	jahac		0	0
	pešak	1.00	2	1
	putnik		0	0
	teraoc stoke		0	0
	vozac	830.00	1565	1494

Slika 6.14. Distribucija nezgoda, vozila i lica po času i svojstvu lica

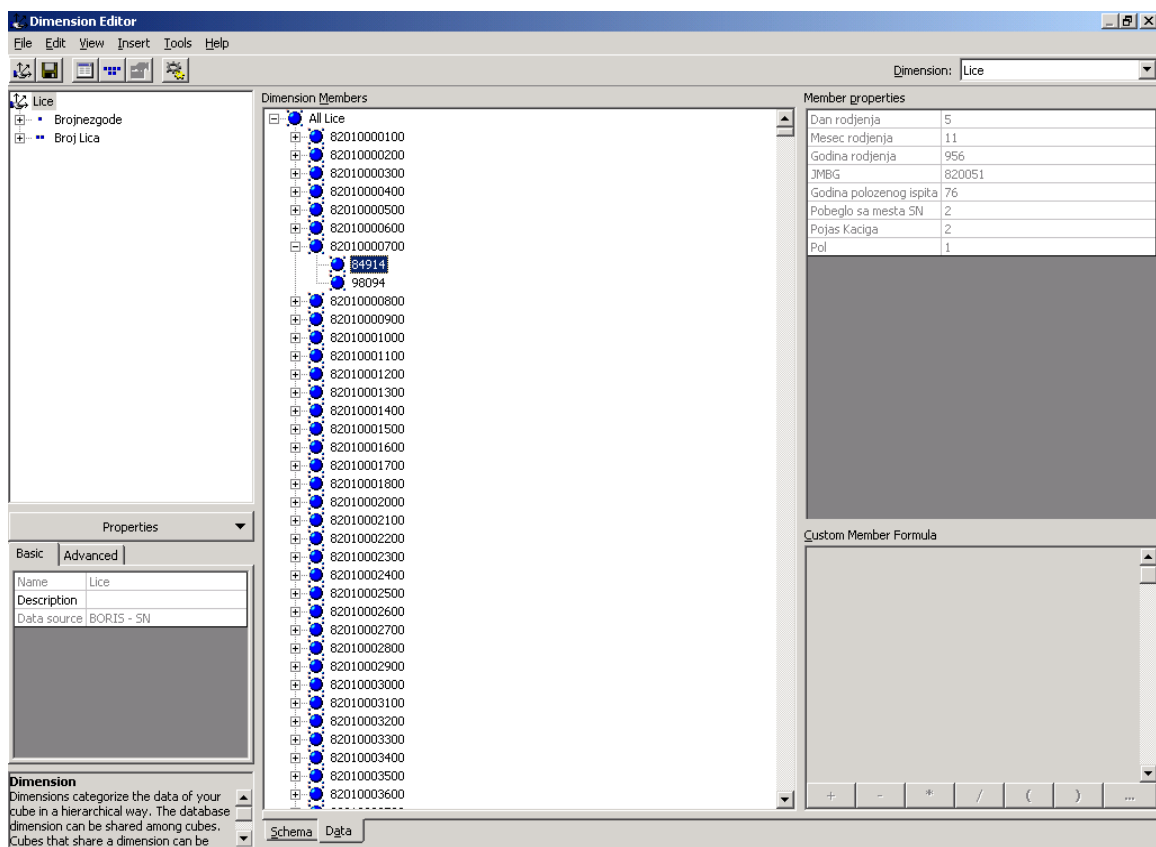
Detaljnije podatke o nezgodi mogu da se dobiju ulaskom u editor dimenzija nezgoda.

Ako se želi posmatrati nezgoda npr. 820100000900, prvo se vrši njena selekcija, pa se u desnom delu ekrana Slike 6.15., pojavljuju bliži podaci o nezgodi.

Takođe, moguće je između ostalog posmatrati i za konkretnu nezgodu koja su lica i vozila učestvovala u nezgodi. Na Slici 6.16., je prikazana nezgoda 820100000700 u kojoj su učestvovala dva lica, pri čemu može da se pristupi njihovim konkretnim podacima.



Slika 6.15. Pristup podacima o konkretnoj nezgodi



Slika 6.16. Pregled zavisnosti nezgoda i lica i pristup podacima lica

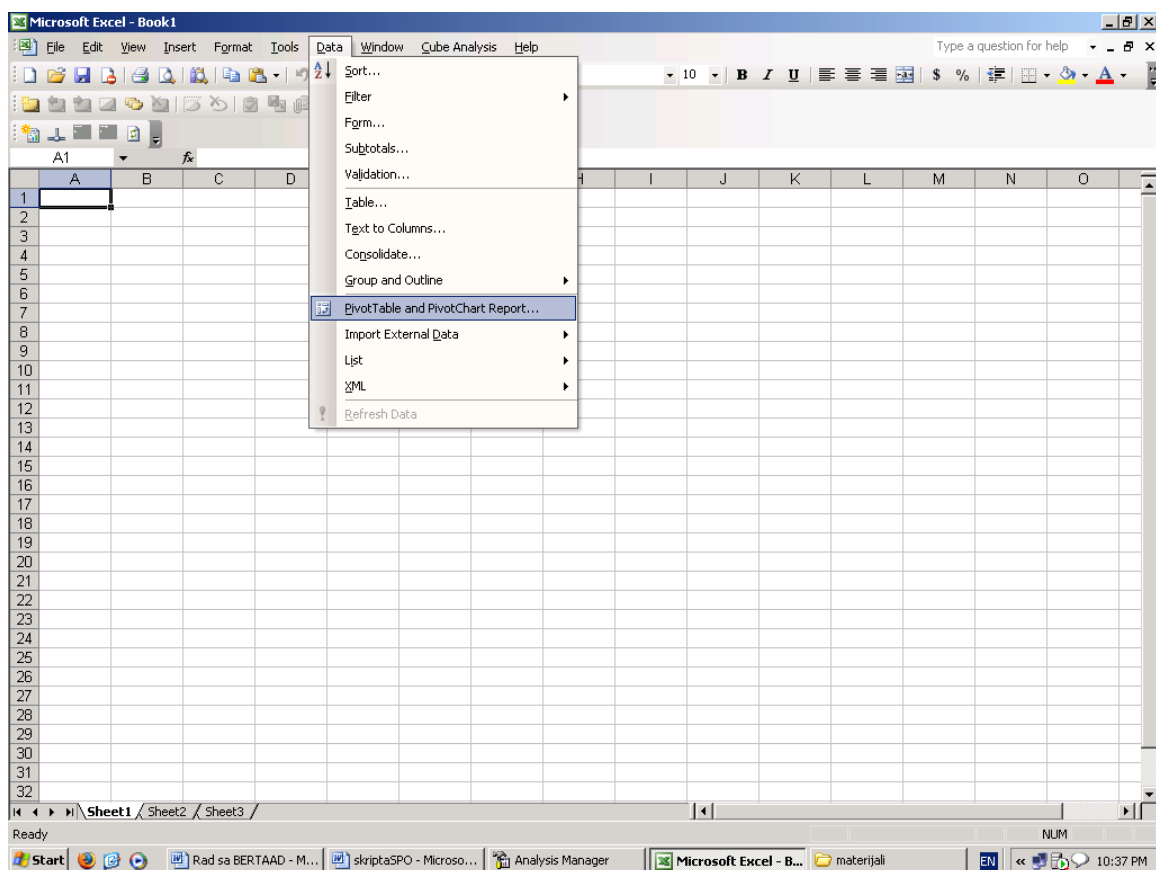
Trenutno u bazi postoje tri podatka (tzv. *mere*) na osnovu kojih se vrši izveštavanje, a to je *broj nezgoda, broj lica i broj vozila* u tim nezgodama. Broj mera nije ograničen i može se proširiti dosta jednostavno i na podatke, kao što je ukupna materijalna šteta, prosečna starost, i sl.

Rad u okruženju Excel je daleko prijatniji, pa se u nastavku ukazuje rad u ovom okruženju.

6.2.2. Rad sa bazom u Microsoft Excel

Pre početka rada u okruženju Excel, potrebno je izvesti određene predradnje. Tabele u kojima se vrši rad sa analitičkom bazom podataka se zovu Pivot tabele.

Alat se pokreće kao na slici 6.17.



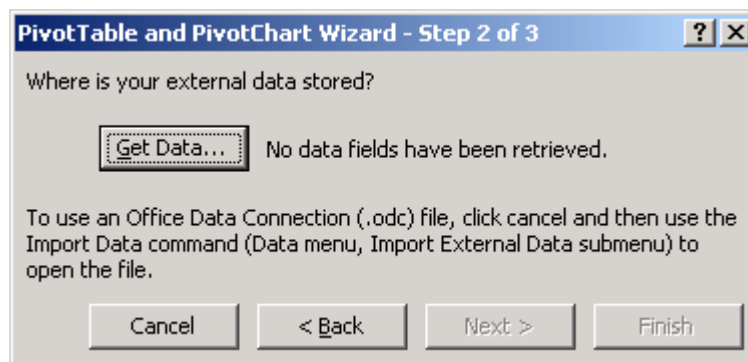
Slika 6.17. Naredba za pokretanje alata za rad sa analitičkom bazom

Zatim se pojavljuje pomoć i podrška, tzv. čarobnjak koji omogućava kreiranje veze sa analitičkom bazom podataka.



Slika 6.18. Izbor izvora podataka i načina rada sa analitičkom bazom

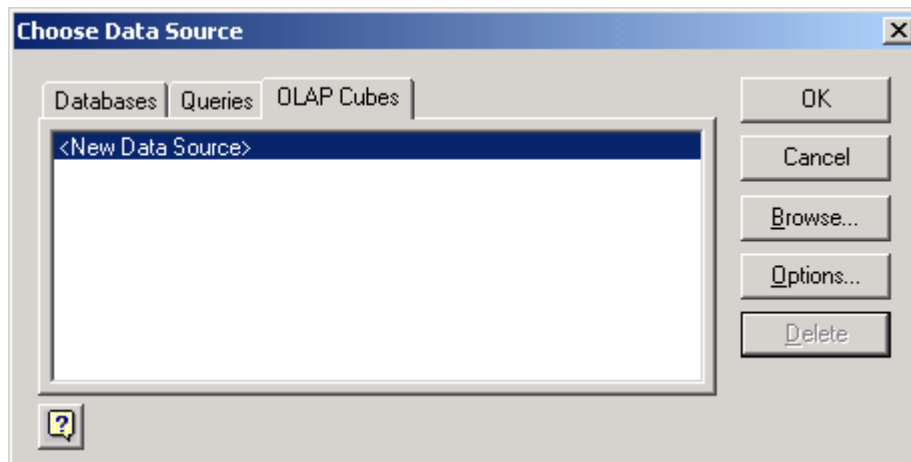
Na Slici 6.18., je izabrana opcija da se podacima pristupa iz eksternih izvora i da se koristi opcija pivot tabele.



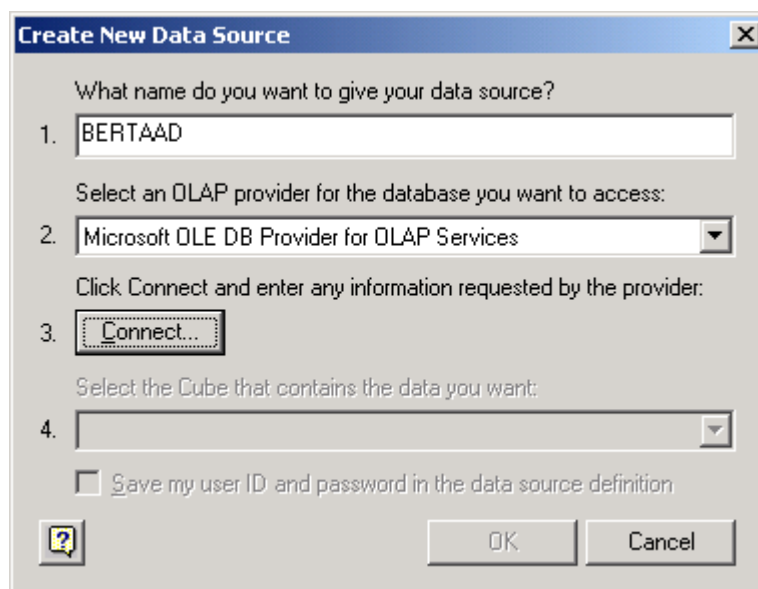
Slika 6.19. Formiranje konekcije sa analitičkom bazaom podataka

Ekran na Slici 6.19., omogućava da se izborom Get Data opcije odredi putanja ka analitičkoj bazi podataka.

Sledeći ekran koji se pojavljuje omogućava da se izabere da li se pravi konekcija sa bazom podataka, upitom, ili analitičkom bazom podataka (tzv. OLAP kockom). Bira se poslednja opcija i potvrđuje se OK.



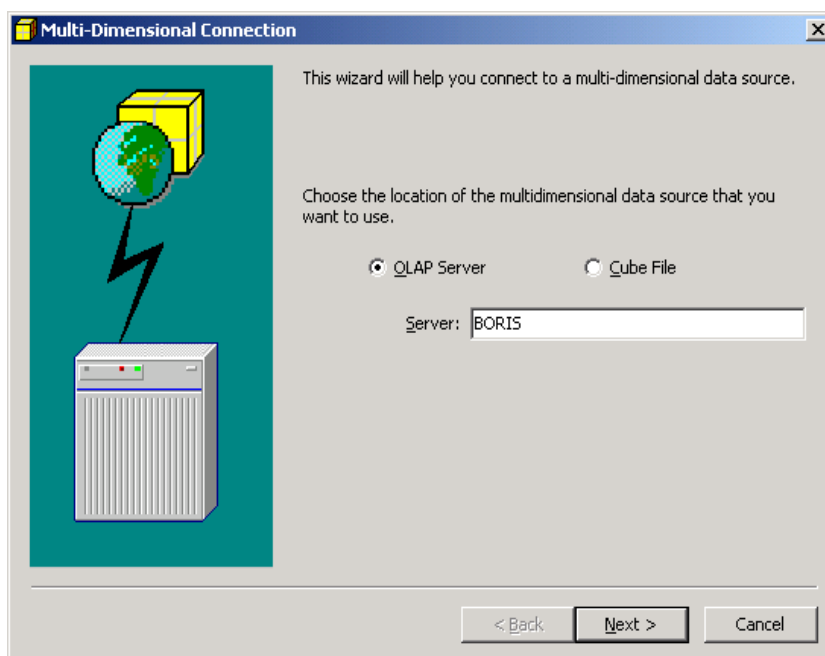
Slika 6.20. Kreiranje veze sa analitičkom bazom



Slika 6.21. Definisanje naziva i servisa za pristup analitičkoj bazi podataka

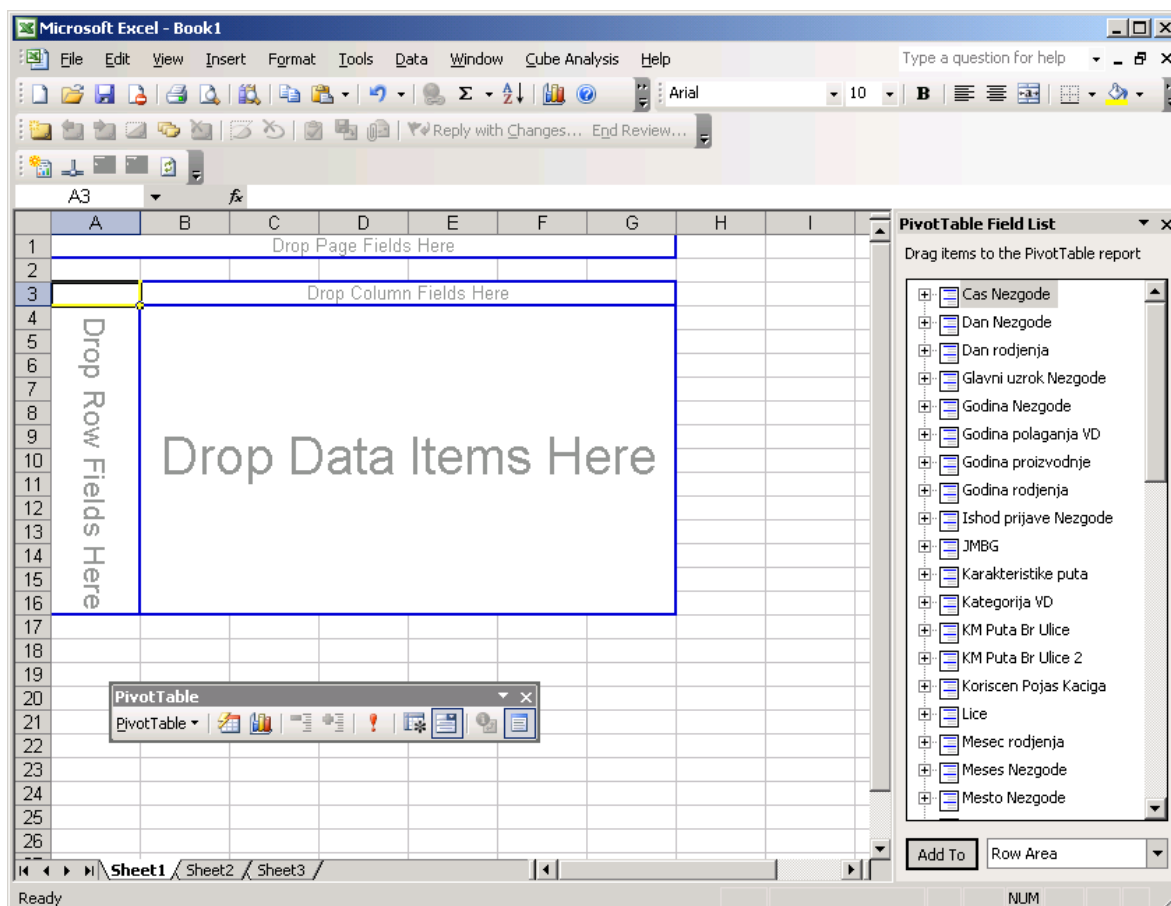
Ekran na Slici 6.21., omogućava da se dodeli ime konekciji i izabere servis za pristup bazi. Treba izabrati opciju kao na slici.

Na sledećem ekranu se bira naziv servera na kome je implementirana analitička baza podataka i zatim se potvrđuje sa OK.



Slika 6.22. Povezivanje sa serverom

Potom se bira naziv baze i ona je spremna za učitavanje u Excel. Excel dobija sledeći izgled, prikazan na Slici 6.23.



Slika 6.23. Izgled ekranske forme Excela, spremnog za rad sa analitičkom bazom

U tzv. Pivot tabeli se nalaze sledeća polja:

- **Data Items:** služe za unos mera kocke. Ako se samo one ubace u polje dobija se ukupan broj nezgoda, ukupan broj vozila u nezgodama ili ukupan broj lica u nezgodama. U ovom polju može biti više mera.
- **Row Fields:** ovde se ubacuju dimenzije koje treba da budu redovi u željenom izveštaju. Treba voditi računa o redosledu ubacivanja dimenzija u ovaj prostor. Sortiranje i prikazivanje u izveštaju se vrši po dimenzijama sa leva na desno. U ovom prostoru, dakle, može biti zastupljeno više dimenzija. Ove dimenzije je moguće rotirati.
- **Column Fields:** ovde se ubacuju dimenzije koje treba da budu kolone u željenom izveštaju. Treba voditi računa o redosledu ubacivanja dimenzija u ovaj prostor. Sortiranje i prikazivanje u izveštaju se vrši po dimenzijama od gore na dole. U ovom prostoru može biti zastupljeno više dimenzija.
- **Page Fields:** Služe za prikazivanje sumarnih podataka za jednu dimenziju. Kada je dimenzija postavljena na ovo polje, tada se može da vrši filtriranje, npr. izaberom samo jedne nezgode ili jedne vrste posledice određenog dan u mesecu. Ovo polje se koristi da bi se dobili rezultati određene dimenzije. U ovom prostoru može biti zastupljeno više dimenzija.

Treba reći da dimenzije služe za sužavanje domena mera (agregacija) i time može da se dođe do željene informacije. OLAP kocka je i smišljena sa namerom lakog postavljanja upita i izveštavanja, što bi u klasičnim bazama podataka iziskivalo daleko više vremena i daleko više znanja za postavljanje upita i pravljenje izveštaja.

6.2.3. Primeri rada sa analitičkom bazom

U nastavku se prikazuju neki od upita koje je moguće dobiti realizovati u analitičkoj bazi podataka.

Prikazaće se sledeće vrste upita:

- ✓ Ukupan broj saobraćajnih nezgoda i lica razvrstanih po posledicima lica.
- ✓ Ukupan broj nezgoda, vozila i lica razvrstanih po mesecima u godini.
- ✓ Ukupan broj nezgoda i vozila, razvrstanih po vrsti vozila i posledicama lica.
- ✓ Ukupan broj nezgoda, vozila i lica, razvrstanih po polu i po kategoriji vozačke dozvole.
- ✓ Ukupan broj nezgoda i lica razvrstanih po danu rođenja lica.

U nastavku se rezultat izveštaja svakog pojedinačnog upit.

A) Ukupan broj saobraćajnih nezgoda i lica razvrstanih po posledicima lica.

Da bi se dobio izveštaj kao na Slici 6.24., potrebno je u Pivot tabelu prikazanu na Slici 6.23., izvršiti sledeće operacije.

1. U polje Drop Data Items ubacuju se iz Pivot Table Field List (desna strana ekrana na Slici 6.23.) mere **Broj Lica** i **Broj Nezgoda**.
2. U polje Drop Row Fields se ubacuju iz Pivot Table Field List dimenzija **Posledice Lica**.

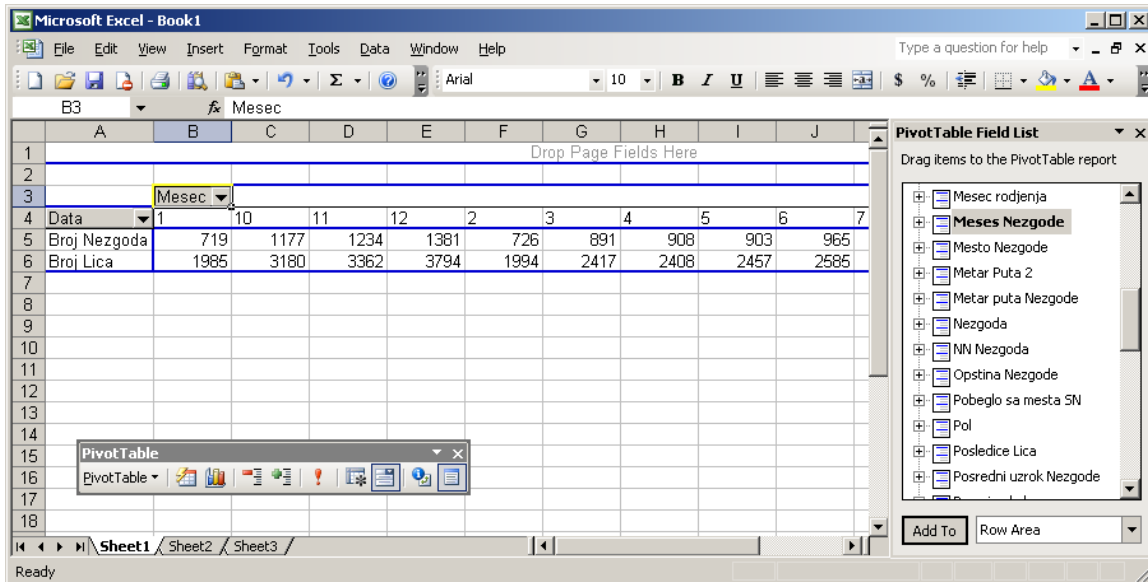
Data	Posledice Lica	Total
Broj Nezgoda	LAKE TELESNE POVREDE	1055
	LICE NEPOVREDJENO	11379
	SMRT DO 30 DANA	8
	SMRT NA LICU	54
	SMRT ZA VREME PREVOZA DO BOLNICE	12
	TESKE TELESNE POVREDE	455
Broj Lica	LAKE TELESNE POVREDE	2183
	LICE NEPOVREDJENO	30424
	SMRT DO 30 DANA	17
	SMRT NA LICU	110
	SMRT ZA VREME PREVOZA DO BOLNICE	25
	TESKE TELESNE POVREDE	920
Total Broj Nezgoda		12133
Total Broj Lica		32844

Slika 6.24. Pregled nezgoda i lica po posledicama lica

B) Ukupan broj nezgoda, vozila i lica razvrstanih po mesecima u godini

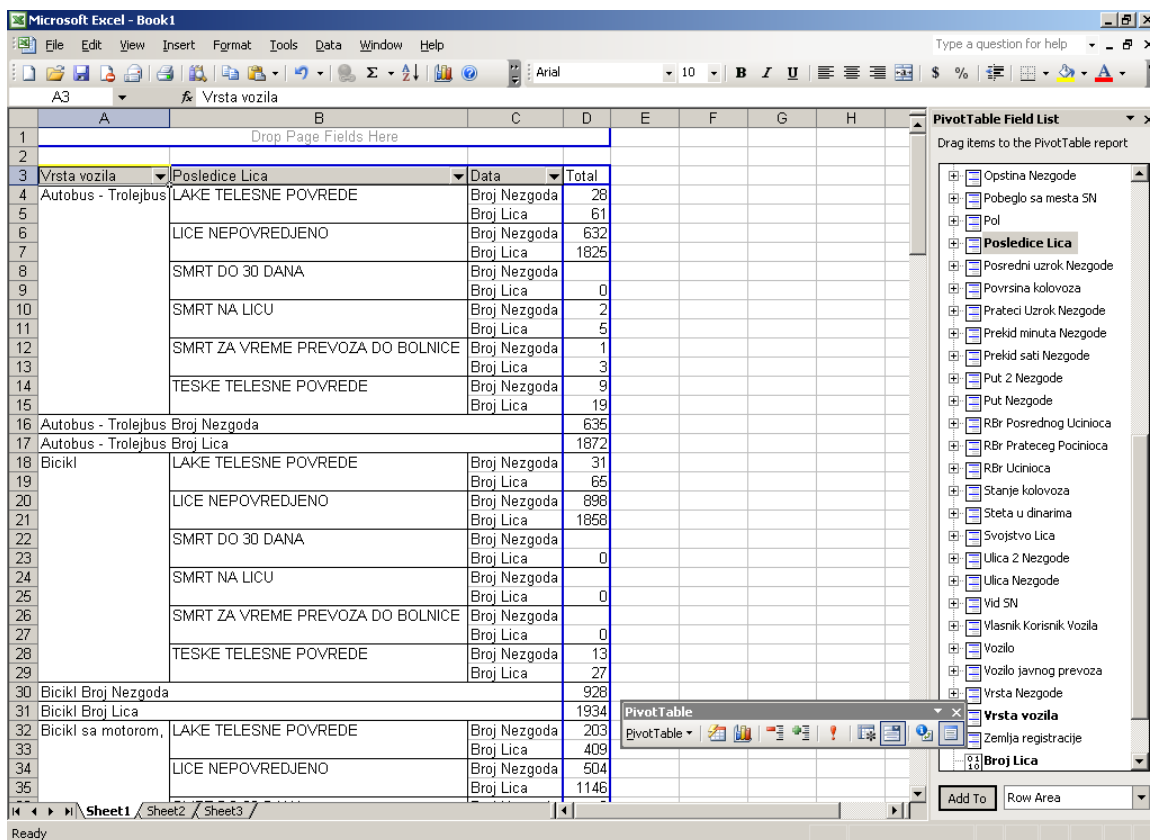
Da bi se dobio izveštaj kao na Slici 6.25., potrebno je u Pivot tabelu prikazanu na Slici 6.23., izvršiti sledeće operacije.

1. U polje Drop Data Items ubacuju se iz Pivot Table Field List (desna strana ekrana na Slici 6.23.) mere **Broj Lica**, **Broj Vozila** i **Broj Nezgoda**.
2. U polje Drop Column Fields se ubacuju iz Pivot Table Field List dimenzija **Mesec Nezgode**.



Slika 6.25. Pregled nezgoda i lica po mesecima u godini.

C) Ukupan broj nezgoda i vozila, razvrstanih po vrsti vozila i posledicama lica



Slika 6.26. Pregled nezgoda i vozila po vrsti vozila i posledicama lica

Da bi se dobio izveštaj kao na Slici 6.26., potrebno je u Pivot tabelu prikazanu na Slici 6.23., izvršiti sledeće operacije.

1. U polje Drop Data Items ubacuju se iz Pivot Table Field List (desna strana ekrana na Slici 6.23.) mere **Broj Lica** i **Broj Nezgoda**.

2. U polje Drop Row Fields se ubacuju iz Pivot Table Field List dimenzije: **Vrsta vozila i Posledice Lica**.

D) Ukupan broj nezgoda, vozila i lica, razvrstanih po polu i po kategoriji vozačke dozvole

Drop Page Fields Here		Pol		Grand Total
Data	Kategorija VD	1	2	
Broj Nezgoda	A	128	3	131
	B	10663	1958	11312
	C	1523	2	1524
	D	668	3	670
	E	162		162
	F	89		89
	potvrda o poznavanju propisa	230	19	249
Broj Lica	A	258	7	264
	B	26282	4031	29003
	C	3189	5	3192
	D	1365	7	1370
	E	333	0	333
	F	190	0	190
	potvrda o poznavanju propisa	464	39	502
Broj Vozila	A	354	9	362
	B	30395	5810	32110
	C	4498	7	4501
	D	2006	10	2012
	E	483	0	483
	F	269	0	269
	potvrda o poznavanju propisa	644	54	697
Total Broj Nezgoda		11667	1980	12133
Total Broj Lica		30279	4080	32844
Total Broj Vozila		33050	5871	34205

Slika 6.27. Pregled nezgoda, lica i vozila po polu i kategoriji vozačke dozvole

Da bi se dobio izveštaj kao na Slici 6.27., potrebno je u Pivot tabelu prikazanu na Slici 6.23., izvršiti sledeće operacije.

1. U polje Drop Data Items ubacuju se iz Pivot Table Field List (desna strana ekrana na Slici 6.23.) mere **Broj Lica**, **Broj Vozila** i **Broj Nezgoda**.
2. U polje Drop Column Fields se ubacuju iz Pivot Table Field List dimenzija: **Pol**.
3. U polje Drop Row Fields se ubacuju iz Pivot Table Field List dimenzija: **Kategorija VD**.

E) Ukupan broj nezgoda i lica razvrstanih po danu rođenja lica

Da bi se dobio izveštaj kao na Slici 6.28., potrebno je u Pivot tabelu prikazanu na Slici 6.23., izvršiti sledeće operacije.

1. U polje Drop Data Items ubacuju se iz Pivot Table Field List (desna strana ekrana na Slici 6.23.) mere **Broj Lica**, i **Broj Nezgoda**.
2. U polje Drop Row Fields se ubacuju iz Pivot Table Field List dimenzija: **Dan Rodjenja**.

The screenshot shows Microsoft Excel with a PivotTable and the PivotTable Field List. The PivotTable is structured as follows:

Dan Rodjenja	Data	Total
1	Broj Nezgoda	873
1	Broj Lica	1764
10	Broj Nezgoda	731
10	Broj Lica	1480
11	Broj Nezgoda	625
11	Broj Lica	1261
12	Broj Nezgoda	674
12	Broj Lica	1360
13	Broj Nezgoda	629
13	Broj Lica	1271
14	Broj Nezgoda	624
14	Broj Lica	1257
15	Broj Nezgoda	791
15	Broj Lica	1609
16	Broj Nezgoda	623
16	Broj Lica	1259
17	Broj Nezgoda	693
17	Broj Lica	1393
18	Broj Nezgoda	676
18	Broj Lica	1361
19	Broj Nezgoda	643
19	Broj Lica	1298
2	Broj Nezgoda	723
2	Broj Lica	1458
20	Broj Nezgoda	701
20	Broj Lica	1416
21	Broj Nezgoda	608
21	Broj Lica	1228
22	Broj Nezgoda	658
22	Broj Lica	1328
23	Broj Nezgoda	637
23	Broj Lica	1288

The PivotTable Field List on the right shows the following items:

- Cas Nezgode
- Dan Nezgode
- Dan rodjenja**
- Glavni uzrok Nezgode
- Godina Nezgode
- Godina polaganja VD
- Godina proizvodnje
- Godina rodjenja
- Ishod prijave Nezgode
- JMBG
- Karakteristike puta
- Kategorija VD
- KM Puta Br Ulice
- KM Puta Br Ulice 2
- Koriscen Pojas Kaciga
- Lice
- Mesec rodjenja
- Meses Nezgode
- Mesto Nezgode
- Metar Puta 2
- Metar puta Nezgode
- Nezgoda
- MN Nezgoda
- Opstina Nezgode
- Pobeglo sa mesta SN
- Pol
- Posledice Lica

Slika 6.28. Pregled nezgoda i lica po danu rođenja lica.

Imajući u vidu činjenicu, da kroz sva poglavlja disertacije provejava važnost posedovanja znanja prilikom odlučivanja u nastavku sledi poglavlje posvećeno upravo oblasti menadžmenta znanja.

7.0. MENADŽMENT ZNANJA U ORGANIZACIJI

U ovom delu disertacije sledi opis *resursa znanja*, kao ključne komponente za donošenje ispravnih upravljačkih poslovnih odluka u datom poslovnom okruženju. Sledi upravo opis znanja.

7.1. Znanje

Preduslov za uspešno poslovanje jedne kompanije u konkurentnim uslovima, jeste da ona na vreme shvati u kakvom se okruženju nalazi kao i koje mogućnosti poseduje kako bi postigla konkurentsku prednost. Takođe treba da zna šta uslovljava njeno ponašanje i šta ona može da učini kako bi upravljala promenama. Prošli vek je sigurno obeležila informaciona tehnologija, pa je intelektualni kapital postao glavno oružje kompanija u borbi za opstanak i stvaranje profita. U budućnosti će preživeti samo one kompanije koje nauče da upravljaju znanjem i koje shvate da je ulaganje u obuku i razvoj kadrova od najveće važnosti

U poslovnom procesu, prilikom donošenja poslovnih odluka u ključnim trenucima organizacije u smislu opstanka i daljeg progressa, menadžeri se najviše oslanjaju na znanje. Problemi tokom životnog ciklusa preduzeća mogu biti različiti. Svaki od njih zahteva poseban pristup i poseban napor upravljačke ili kadrovske strukture u cilju njegovog otklanjanja. Idealno rešenje problema za svakog menadžera predstavlja razumevanja okolnosti, konteksta pod kojim se dati problem javlja, kao i sama akcija koja će dovesti do rešavanja sporne situacije. Ta ista akcija bazira se na znanju.

U današnjoj organizacionoj praksi, znanje je shvaćeno kao jedan od najvažnijih resursa za uspeh. Poznati g-din, Bill Gates ide i korak dalje i ističe da je ovo vreme ekonomije znanja, koje zahteva od kompanija da integrišu svoje izvore informacija i poslovne procese. Uz pomoć IT, mogu u najmanje da budu sposobne da na izazove odgovore brzo, efikasno i najprihvatljivije. Najveći značaj u tome pridaje se „glatkom“ toku informacija između različitih komponenti poslovanja i koji mora omogućiti kompanijama da reaguju na promene u okruženju, predvide konkurentske poteze i odgovore na potrebe potrošača.

Ako se znanje kao ključna dimenzija procesa odlučivanja restrukturira na sastavne elemente, može se uočiti da se znanje sastoji od skupa činjenica, odnosno jasnih i dokazanih tvrdjenja

i heuristika, odnosno zapisa iskustava koje omogućavaju uočavanje pravila. Može se zaključiti da strukturu znanja čine skup činjenica i određena organizacija među tim činjenicama.

Pojam znanja potrebno je razlikovati od pojma *inteligencije*. Često u definicijama autori pojam znanja poistovećuju sa inteligencijom. Inteligencija se odnosi na sposobnost prihvatanja i primene znanja od strane pojedinca. Ona predstavlja sposobnost razumevanja i korišćenja jezika kao i sposobnost pamćenja, brzog rešavanja problema. Znanje sadrži komponentu akcije i ima u sebi karakteristike inteligencije. U procesu donošenja poslovnih odluka, važnu ulogu igra sistem poslovne inteligencije, koji se bazira na znanju, za koje je više reči o samoj poslovnoj inteligenciji i njenim oblastima bilo u prethodnom poglavlju.

Interesantni opisi znanja

Znanje predstavlja osnovu za donošenje ispravnih upravljačkih poslovnih odluka. Ne postoji opšta i jednoznačna saglasnost oko definicija znanja. Slede neki pogledi, autora čije su reference date u poglavlju posvećenom Literaturi.

Tiwana, A., definiše znanje kao informaciju sa akcijom raspoloživom u pravom formatu, u pravo vreme i na pravom mestu za odlučivanje. Znanje bi u tom slučaju predstavljalo pravovremenu, adekvatnu akciju, preduzetu od strane kompetentnog učesnika poslovnog procesa, na bilo kom delu organizacione strukture u cilju prilagođavanja i usvajanja promena turbulentnog okruženja.

Definicija koja znanje posmatra kao razumevanje je definicija koju je dao Awad, D., po kojoj je znanje razumevanje stečeno iskustvom ili učenjem. Razumevanje daje odgovor na pitanje šta uraditi i kako reagovati u određenoj situaciji ako dođe do problema koji zahteva znanje.

Firestone, J., u svojoj definiciji navodi da je znanje informacija koj je prošla proces validacije.

Postoji i definicija po kojoj je znanje sposobnost pretvaranja informacija i podataka u efektivnu akciju. Acharya, N., navodi da je znanje ljudska interakcija sa stvarnošću.

Karakteristika znanja je da je ono uvek vezano za kontekst. Primena znanja za rešavanje određenog problema u određenoj situaciji, ne znači da će isto znanje pomoći u nekoj drugoj situaciji sa istim problemom ili za rešavanje nekog drugog problema nastalog u istoj situaciji. Kada se javi određeni problem koji treba rešiti bitan je kontekst u kome se on javlja,

jer on zajedno sa problemom određuje rešenje. Može se zaključiti da trojka problem, kontekst i rešenje predstavlja znanje.

U svakoj organizaciji, prilikom donošenja poslovnih odluka, koristi se znanje, a znanje se vezuje za ljude. Znanja pojedinaca se razlikuju, s obzirom da svaki pojedinac ima razvijen drugačiji sistem vrednosti, verovanja i iskustava. Što je više pojedinaca uključeno u rešavanje poslovnog problema, samom rešenju se prilazi iz više perspektiva i na više načina što je i jeste cilj organizacije. Međusobnom saradnjom i učestvovanjem u grupnom odlučivanju, zaposleni koriste svoje sposobnosti i veštine i dižu nivo opšteg znanja na viši nivo. Jer se nedvosmisleno dokazalo u prethodnim poglavljima disertacije, da je zbir znanja članova tima po definiciji uvek veći od znanja bilo kog pojedinca u grupi.

Bitan faktor znanja je iskustvo. Tokom procesa rada zaposleni stiču iskustvo koje predstavlja praktičnu proveru teorijskih koncepata samih elemenata rada. Iskustvo vodi do ekspertize i ekspertskog znanja. Znanje predstavlja evolutivni proces i za organizaciju, u smislu rešenja problema koji se javlja u određenom kontekstu, važno je znanje koje je prošlo iskustvenu verifikaciju.

Vrste znanja

Brojne su podele vrsta znanja, ali među njima dve podele ističu suštinu prirode ovog pojma i mogu se koristiti za jasnu i sveobuhvatnu klasifikaciju znanja.

Prema prvoj podeli znanje može da bude: *proceduralno, deklarativno, semantičko i epizodno*, detaljno opisano u [34].

Proceduralno znanje predstavlja način, postupak na osnovu kojeg se radi određeni zadatak ili procedura. Karakteristika proceduralnog znanja je u tome što ono odgovara na pitanje kako nešto uraditi, ali ne i zašto se to radi. Ovakav oblik znanja koristi se u uslovima prilične izvesnosti u odlučivanju, odnosno kada su jasno definisani problem i kontekst na koje se ono odnosi.

Deklarativno znanje je znanje koje opisuje određenu pojavu. Ovo znanje poboljšava proceduralno, rutinsko znanje, tako što daje dodatne informacije, opisujući neku pojavu. Deklarativno znanje preporučuje se u ranim fazama procesa učenja i ono u sebi ne sadrži potrebnu akciju za rešavanje problema.

Semantičko znanje ukazuje na organizaciju sastavnih činilaca znanja. Kao i kod deklarativnog znanja, ne dostaje mu dinamički aspekt, odnosno okrenutost ka akciji. Sa druge strane, semantičko znanje pruža odgovor na pitanje zašto bi nešto trebalo uraditi.

Epizodno znanje je znanje sasavljeno iz iskustvenih zapisa, slučajeva, između kojih postoji određena organizacija. Iz njega se može saznati šta je posebno uraditi, ali i zašto to uraditi. Cilj je da svako znanje pređe put od proceduralnog i deskriptivnog do epizodnog da bi se na kraju problem koji zahteva epizodno znanje, mogao rešavati na proceduralan način. Za epizodno znanje može se reći da je najpraktičnije sa aspekta donošenja poslovnih odluka.

Japanski autori Nonaka, I., i Takeuchi, H., u jednoj od najpoznatijih podela znanja u oblasti menadžmenta znanja, znanje klasifikuju na eksplicitno i tacitno, videti [112].

Eksplisitno znanje je ono znanje koje organizacije čuvaju u bazama podataka u formi izveštaja, dokumenata, knjigama ili tabelama. Čuva se u obliku koji se može lako formalizovati i predstaviti. Predstavljeno je u vidu teoretskih koncepata i ponekad nema iskustvenu proveru.

Tacitno znanje je znanje koje je prošlo iskustvenu proveru, ali se ne može uvek formalizovati i predstaviti u pisanom obliku. Ovo znanje čuva se u glavama zaposlenih i predstavlja vrednost svake organizacije. Tacitno znanje je dokaz da je intelektualni kapital osnovno sredstvo ostvarivanja konkurentске prednosti i pokretačka snaga u borbi sa nepredvidivim promenama u okruženju. S obzirom da znanje zaposlenih predstavlja merilo kvaliteta organizacije, tacitno znanje omogućava da se za određeni problem koji se javlja u određenom kontekstu, pronađe pravo rešenje, tj. sprovede prava akcija za njegovo otklanjanje.

Odnos znanj i paterna

Znanje predstavlja dinamičku kategoriju. Posedovanje znanja je preduslov za donošenje ispravnih poslovnih odluka. Ono u sebi sadrži komponentu akcije. Da bi znanje bilo primenjivo, neophodno je da ima empirijsku proveru. Primena znanja kod poslovnog odlučivanja uvek je vezana za kontekst, odnosno za određeni problem koji u određenoj situaciji, zahteva određenu akciju za rešavanje istog. Trojka problem, rešenje i kontekst je poznata kao *patern*.

Patern je jedinica najmanje energije, koja za dati problem u datom kontekstu nudi rešenje, tj. omogućava akciju. Patern predstavlja optimalno znanje, a sam zahtev da za svaki problem u određenom kontekstu postoji gotovo rešenje je sasvim nerealan. Iz tog razloga postoje

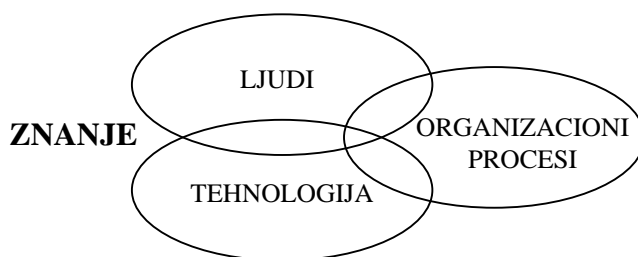
slučajevi. Oni predstavljaju zapise iskustva, koji mogu biti lični ili eksplicitno zabeleženi. Paterni se sastoje od opisa problema i odgovarajućeg rešenja koje je primenjeno u prošlosti.

Do paterna ili oblika bliskog paternu, može se doći upravljanjem slučajevima kroz životni ciklus znanja, a da se tako postigne efikasno donošenje odluka, odnosno upravljanje. Da bi se utvrdilo koliko je određeni slučaj koji je pronađen, a koji je zasnovan na prošlosti, blizu paterna datog problema, uvodi se proces *validacije*. Validacija omogućava otkrivanje paterna, tj. najprihvatljivijeg rešenja problema. Sa druge strane, proces validacije vrši ocenu određenog slučaja za novi problem.

S obzirom da znanje predstavlja osnovu za donošenje poslovnih odluka, u nastavku disertacije sledi sistemski pristup znanju, opisan kroz menadžment znanja.

7.2. Menadžment znanja

Menadžment znanja (MZ) je nov, interdisciplinarni model poslovanja koji se fokusira na znanje u okviru kompanije. *To je najveća konkurentska prednost kojom jedan poslovni sistem može da raspolaze*, detaljnije u [6]. Menadžment znanja uključuje ljude, tehnologiju i procese koji se međusobno prepliću, što je i prikazano na sledećoj slici.



Slika 7.1. Elementi menadžmenta znanja

Neke od osnovnih objašnjenja menadžmenta znanja, videti, [36] su:

- ✓ Menadžment znanja je proces sakupljanja i korišćenja kolektivnog organizacionog znanja, sačuvanog na bilo kom mestu u kompaniji: na papiru, u dokumentaciji, u bazama podataka (eksplicitno znanje) ili u ljudima (tacitno znanje) – i njena distribucija na mesta gde može proizvesti najveću korist.

- ✓ Menadžment znanja je novi, interdisciplinarni poslovni model koji se bavi svim aspektima znanja u kontekstu firme, uključujući stvaranje, kodifikaciju i razmenu znanja, i objašnjava kako te aktivnosti promovišu učenje i inovacije (koristeći tehnološke alate i organizacione rutine).
- ✓ Menadžment znanja zadovoljava kritična pitanja prilagođavanja, opstanka i konkurentnosti organizacije, koja se suočava sa rastuće diskontinualnim promenama u okruženju. U suštini, on obuhvata organizacione procese koji traže sinergetske kombinacije informacione tehnologije u obradi podataka i informacija, sa jedne strane i kreativnog i inovativnog kapaciteta ljudi sa druge strane.
- ✓ Menadžment znanja predstavlja razmenu informacija širom kompanije ili čak između poslovnih partnera. On stvara okruženje u kojem kompanija udružuje sve prednosti koje joj znanje obezbeđuje.
- ✓ Menadžment znanja može automatizovati klasifikaciju dokumenata uz korišćenje "mašinske" logike koja je sve sličnija ljudskoj logici, itd.

Na neki način, može se reći da je menadžment znanja vezan za opstanak u novom poslovnom svetu, tj. svetu takmičenja, u kome kompleksnost i neizvesnost svakodnevno rastu. To je svet u kome nema mesta za tradicionalan način poslovanja, jer ono što je juče funkcionisalo, danas možda neće. Kada se govori o menadžmentu znanja, može se reći da njegova suština nije samo u pronalaženju pravih odgovora, već u postavljanju pravih pitanja.

Kao što je već rečeno, menadžment znanja je proces sakupljanja i korišćenja kolektivnog organizacionog znanja, sačuvanog na bilo kom mestu u preduzeću: na papiru, u dokumentaciji, u bazama podataka (eksplicitno znanje) ili u ljudima (tacitno znanje).

Osnovni cilj koji organizacija treba da ostvari, jesta da vidi sve svoje procese kao procese znanja. Ovo uključuje kreiranje znanja, širenje, nadogradnju i primenu znanja kako bi preduzeće moglo da opstane. Ne postoji opšte pravilo ili algoritam koji garantuje uspešnu primenu menadžmenta znanja. Naime, to je jedan izuzetno kreativan, inovativan i kompleksan proces koji zavisi od velikog broja različitih faktora. Da bi ostvarila korist od znanja, organizacija treba da identifikuje, primeni i integriše znanje, na odgovarajući način koji ne može biti unapred propisan. Zato se za ovaj proces može reći da delom predstavlja nauku, delom umetnost a delom ponekad možda i čistu „sreću“.

Tradicionalno, postojao je unapred definisan recept za uspeh gde su kompanije bile vođene ustaljenim procedurama i starim poslovnim praksama koje su bile prihvaćene kao metode za uspeh. Ponavljajući uporno iste stvari, mnoge vodeće firme su uvidele da gube dodir sa poslovnim okruženjem koje se brzo menja i da njihov tržišni udeo opada. Jučerašnji uspeh više se nije mogao prevesti u sutrašnji uspeh. U savremenom poslovnom svetu funkcionisanje preduzeća, postalo je zavisno od stalnog unapređenja sistema u cilju zadržavanja konkurentnosti.

Opstanak na tržištu poslovnog sistema postao je uslovljen brzom reakcijom na promene u okruženju. Preduzeća se više nisu mogla osloniti na sadašnja pravila poslovanja jer se poslovno okruženje svakodnevno menja. Kao reakcija na ovako drastične promene u okruženju tokom 1980-ih i 1990-ih godina, javio se menadžment znanja. On je ustvari nastao kao najbolje rešenje za održanje konkurentnosti u novim, nepredvidivim uslovima. Da bi postale i ostale uspešne, kompanije danas moraju redefinisati i razmotriti svo svoje znanje, sadržano u korporativnim bazama podataka, istovremeno, stvarajući nove prakse kako bi se prilagodile poslovnom okruženju.

Informatička revolucija stavila je naglasak na razmenu velikih količina informacija koje su danas dostupne na Internetu. U vreme "e-svega", do informacija se može doći iz B2B (business-to-business, tj. poslovanje ka poslovanju), B2C (business-to-customer, tj. poslovanje ka kupcu) i C2C (customer-to-customer, tj. kupac ka kupcu), videti u [97].

Korporacije u SAD počele su da koriste ovu dostupnost informacija u svoju korist. Eksterni odnosi kao što je menadžment lanca snabdevanja uspešno su korišćeni za poboljšanje produktivnosti i fleksibilnosti zahvaljujući razmeni informacija između dobavljača i potrošača. Kompanije su preuzele ovu ideju razmene informacija preko menadžmenta znanja, kako bi je primenile i unutar preduzeća. Zahvaljujući tehnologiji, zaposleni danas mogu interno razmenjivati znanje nastojeći da učine preduzeća produktivnijim.

Slede neki od faktora koji su doprineli popularizaciji MZ, detaljno u [46]:

- ✓ *Stopa promene značajno se uvećala tokom prošle decenije.* Kompanije traže inovativne načine za pobeđivanje konkurencije. Inovacija je jedna od najznačajnijih stručnosti, potrebna svim organizacijama.

- ✓ *Globalizacija i geografska disperzija promenili su obim organizacije. Sve više organizacija pokušava da se osloni na godine iskustva, kako bi svojim globalnim obavezama upravljala na blagovremen i profitabilan način.*
- ✓ *Smanjenje broja zaposlenih i reinženjering doveli su do osipanja zaposlenih i "sužavanja znanja". To je podstaklo organizacije da procene svoje osnovno znanje i efektivnije ga koriste. Reinženjering je podrazumevao jednokratno rešenje za određene situacije. Time je stvoren začaran krug u kome su rešenja postala novi problemi, zbog čega nije bilo moguće ispratiti brze promene na savremenom tržištu.*
- ✓ *Umrežavanje i komunikacije olakšale su i ubrzale razmenu znanja. Razmena znanja putem tehnologije postaje najbolji način za distribuciju stručnosti između i unutar preduzeća. Sama tehnologija nije dovoljna.*
- ✓ *Rastuća dominacija znanja kao osnove za poboljšanje efikasnosti i efektivnosti podstakla je mnoge organizacije da počnu da koriste znanja koja su stekla tokom godina iskustva, itd.*

Ovi faktori jasno ukazuju na to, kako menadžment znanja doprinosi opstanku preduzeća. *Znanje postaje ključni faktor uspeha i osnovna komparativna prednost svakog preduzeća.* To je funkcija koja može i treba da bude uključena u svaki od poslovnih procesa – od novih proizvoda i usluga, preko novih kanala distribucije i novih marketing strategija, do novih polja i grana. Tehnologija je "kičma", a ljudske komponente su "mišići" neophodni za njeno korišćenje. Prema istraživanju koje je 2000. godine sproveda vodeća konsultantska firma iz SAD na 243 američke i međunarodne organizacije, čiji je cilj bio da oceni trenutni status menadžmenta znanja u poslovanju, preko 2/3 ispitanika izjavilo je da već imaju strategiju menadžmenta znanja. Najznačajniji rezultati ovog istraživanja, prikazano u [97] su:

- ✓ Menadžment znanja pruža stvarne koristi.
- ✓ Kompanije sa menadžmentom znanja su bolje od onih koje ga nemaju.
- ✓ Organizacije ne uspevaju da reše izazove menadžmenta znanja. Neuspeh da se menadžment znanja integriše u svakodnevni rad, nedostatak vremena da se uči i deli znanje, odsustvo obuke, mala lična korist i neuspesi u efektivnom korišćenju znanja i dalje predstavljaju problem.
- ✓ Organizacije se bore da shvate kulturološke implikacije menadžmenta znanja. One otkrivaju da menadžment znanja usložnjava opis radnog mesta.

- ✓ Kompanije i dalje vide menadžment znanja kao čisto tehnološko rešenje. Do promene u razmišljanju postepeno dolazi, ali tek treba da se uvidi značaj ljudskog faktora na uspeh MZ projekta.
- ✓ Mnogo toga treba da se uradi na usvajanju i integrisanju menadžmenta znanja, ali organizacije su spremne da ulože napor koji je potreban, itd.

Za kadrovske resurse se obično tvrdi da imaju vodeću ulogu u menadžmentu znanja. Kada je PricewaterhouseCoopers sproveo istraživanje u 90 kompanija otkriveno je da 32% zaposlenih veruje da su ljudski resursi ključni za menadžment znanja, a 25% zaposlenih izjasnilo se za informacionu tehnologiju. Iako su informaciona tehnologija i tehnologija uopšte značajni deo menadžmenta znanja, ljudi su i dalje njegova pokretačka snaga, opisano u [108].

Kompanije se opredeljuju za implementaciju MZ, najčešće iz dva osnovna razloga:

- ✓ radi lakšeg deljenja postojećeg znanja – koristeći mehanizme za lakše pokretanje znanja tamo gde je potrebno, i
- ✓ inovacija – skraćenje vremena između pojave ideja i njihove komercijalizacije.

Implementacije MZ u različitim organizacijama variraju od novih, tehnološki podržanih načina pristupa, kontrole i dostave informacija, do masivnih pokušaja promene korporativne kulture. Konkretna implementaciona rešenja razlikovaće se od kompanije do kompanije, a glavni faktori koji ih određuju su tip kompanije, njena kultura i potrebe kompanije.

Za uspešnu primenu MZ-a u organizacijama bilo koje vrste bitno je znati sledeće, videti u [109]:

- ✓ Kako sa zaposlenima?

Veoma je bitno da se uspostavi organizaciona kultura koja prepoznaje prećutno znanje i ohrabruje deljenje među zaposlenima, naravno u sredini u kojoj je znanje pojedinca cenjeno i nagrađivano. Detaljno objašnjenje koncepta MZ-a nikako ne sme da izostane. Često se dešavalo da se motivacija zaposlenih za primenu MZ rešavala različitim načinima nagrađivanja, ali tu uvek postoji opasnost da će zaposleni prihvatiti ovu novu inicijativu zbog samih nagrada. Zato je idealno da zaposleni shvate da je participacija u MZ nagrada sama za sebe. Drugim rečima, ako MZ ne olakšava rad zaposlenima, napori su osuđeni na propast.

- ✓ Ne dozvoliti tehnologiji da diktira MZ

MZ nije koncept baziran isključivo na tehnologiji. Tehnologija treba da pomogne MZ, a izbog odgovarajuće tehnologije zavisi od okolnosti u kojima se MZ primenjuje. Prvo treba odgovoriti na pitanja „ko” (ljudi), „šta” (znanje), „zašto” (poslovni ciljevi) a ostaviti „kako” (tehnologija) za kraj. Ne treba se nadati da će primena bilo kakvog softvera razvijenog u ove svrhe značiti uvođenje MZ u organizaciju.

- ✓ Treba imati određeni poslovni cilj

Program MZ mora biti povezan sa jasnim poslovnim ciljem, tj. mora biti implementiran sa jasnim razlogom. U suprotnom, može se shvatiti samo kao vežba koja ničemu nije poslužila.

- ✓ Menadžment znanja nije statičan

Kao i kod fizičkih dobara, vrednost znanja se gubi tokom vremena. Pošto znanje može da izgubi svoju vrednost prilično brzo, sadržaj MZ programa mora biti konstantno ažuriran, dopunjavan i brisan. Takođe, značaj znanja se menja u toku vremena, kao što se menjaju i same veštine zaposlenih. Tako da treba shvatiti da ne postoji krajnja tačka MZ programa. Kao i istraživanje i razvoj, MZ se konstantno unapređuje i menja. Takođe, značaj znanja se menja u toku vremena, kao što se menjaju i same veštine zaposlenih. Tako da treba shvatiti da ne postoji krajnja tačka programa MZ.

- ✓ Nisu sve informacije znanje

Organizacije treba da budu oprezne kad je u pitanju poplava informacija. Kvalitet retko znači i „pravi” kvalitet, pa ni MZ tu nije izuzetak. Poenta MZ programa i jeste da identifikuje znanje iz obilja informacija.

Izvori znanja u najgrubljoj podeli su unutrašnji (interni) i spoljni (eksterni), kao pandam prethodno opisanom petom poglavlju. Unutrašnji, ali i često najmanje dostupni izvori znanja u organizaciji su eksperti. Eksperata ima relativno malo u poređenju sa veličinom cele organizacije i njihovo je znanje veoma cenjeno.

Kada se govori o izvorima znanja koji se nalaze izvan poslovne organizacije, jedan od najčešćih i najkorisnijih izvora jesu svakako kupci, odnosno korisnici. Poznavanje kupaca svakako je jedan od bitnih faktora uspeha svake kompanije. Teškoće koje se ovde javljaju (a pogotovo kod velikih kompanija) odnose se na rasutost znanja o kupcima u okvirima kompanije. Mnogo različitih poslovnih funkcija susreće se sa kupcima, počev od marketinga, prodaje, preko servisa, logistike, pa čak i do finansijske funkcije. Svaka oblast ima svoja interesovanja za informacije o kupcima, različite načine čuvanja onoga što je naučeno. Međutim, kod znanja o korisnicima, pored znanja proizašlog iz podataka, postoji i

jedan „ljudski” oblik znanja, znanje koje je proisteklo iz interakcije između ljudi. Može se istaći da uključuje eksperimentalna zapažanja, komentare, naučene lekcije, zaključke i kvalitativne činjenice. U nastavku sledi opis koraka projektovanja sistema menadžmenta znanja.

7.3. Koraci projektovanja sistema menadžmenta znanja

Metodologija projektovanja sistema menadžmenta znanja (SMZ)

Da bi se uveo sistem menadžmenta znanja, koriste se sledeći koraci, videti [6]:

- ✓ evaluacija postojeće infrastrukture,
- ✓ formiranje tima MZ,
- ✓ snimanje znanja,
- ✓ projektovanje SMZ,
- ✓ verifikovanje i validacija SMZ,
- ✓ implementacija SMZ,
- ✓ upravljanje promenama, i
- ✓ evaluacija uvedenog sistema.

Evaluacija postojeće infrastrukture

U ovoj fazi potrebno je odrediti gde se poslovni sistem trenutno nalazi i gde želi da se nađe nakon uvođenja novog SMZ. Preporuka je da se počne sa malim projektom, a da ako on pokaže dobre rezultate, da se nastavi sa daljom izgradnjom sistema. Polazi se od ocene opravdanosti sistema, utvrđivanja granice sistema, zatim se radi studija izvodljivosti, da bi se na kraju odredilo da li će se graditi novi SMZ sistem, kupiti postojeći ili uraditi outsorsing.

Formiranje tima MZ

Pošto je doneta odluka o sprovođenju MZ projekta i nakon završene analize postojećeg stanja organizacije prelazi se na fazu formiranja MZ tima. Ovaj tim će biti zadužen za uvođenje sistema MZ.

U sastav tima ulaze eksperti iz onih organizacionih jedinica za koje će se i projektovati sistem menadžmenta znanja. Pored projektanta sistema MZ (projektanta znanja), u sastav MZ tima ulaze i vođa projekta koji je zadužen normalno odvijanje celog procesa. Broj učesnika u timu je fleksibilan i varira u zavisnosti od veličine projekta.

Projektant znanja identifikuje problemski domen, beleži znanje, piše i testira heuristike koje predstavljaju znanje i kordinira ceo projekat od početka do završetka. Bitno je da ta osoba

ima izvanredne komunikacione veštine, poznavanje alata za snimanje znanja, kao i odlično poznavanje informacionih tehnologija.

Snimanje znanja

Za proces snimanja znanja neophodna je pre svega identifikacija znanja, kao i stvaranje znanja. Stvaranje znanja predstavlja proces dolaženja do znanja iskustvenim putem, učenjem ili otkrićem. Zanimljiva podela stvaranja znanja je Nonakina podela, prikazana u [112]. On navodi da postoje 4 vrste konverzije znanja koje dovode do stvaranja znanja:

1. **Socijalizacija** (transformacija tacitnog u tacitno znanje) se dešava svaki put kada ljudi sa znanjem komuniciraju i razmenjuju svoja iskustva i ideje i tada se za prikupljanje znanja koristi tehnika naučnog posmatranja.

2. **Eksternalizacija** (transformacija tacitnog u eksplicitno znanje) predstavlja objašnjavanje ili pojašnjavanje tacitnog znanja preko analogija, modela ili metafora. Koristi se kao metod za prenošenje tacitnog znanja u pogodnom formatu na osobe bez prisustva eksperta od koji je izvor tacitnog znanja. Jedna od tehnika eksternalizacije je breinstorming.

3. **Internalizacija** (transformacija eksplicitnog u tacitno znanje) predstavlja složen proces dobijanja tacitnog znanja iz eksplicitnog, pri čemu se koriste tehnike rudarenja po podacima.

4. **Komunikacija** (transformacija eksplicitnog u eksplicitno znanje) predstavlja znanje koje se dobija kombinovanjem, reorganizovanjem ili sortiranjem različitih eksplicitnih znanja čijom reorganizacijom se dobija novo znanje. Ovo znanje se čuva u bazama znanja i sa njim je relativno jednostavno upravljati.

	TACITNO ZNANJE	EKSPPLICITNO ZNANJE
TACITNO ZNANJE	<p>SOCIJALIZACIJA</p> <p>(Sastanci i diskusija)</p>	<p>EKSTERNALIZACIJA</p> <p>(Dijalog, odgovaranje na pitanja)</p>
EKSPPLICITNO ZNANJE	<p>INTERNACIONALIZACIJA</p> <p>(Učenje iz izveštaja)</p>	<p>KOMUNIKACIJA</p> <p>(Slanje izveštaja)</p>

Slika 7.2. Konverzija znanja, videti [112]

Sva četiri načina stvaranja znanja, pokazuju kako znanje nastaje u kompaniji i predstavljaju fazu **identifikacije** znanja.

Nakon utvrđivanja procesa nastajanja znanja tj. u kojim poslovnim procesima se koristi i gde je potrebno, moguće je početi sa prikupljanjem identifikovanog znanja. Pod **snimanjem** znanja se podrazumeva proces u kome se ekspertske misli i iskustva beleže.

Sastoji se od 3 koraka:

- ✓ izbor odgovarajuće metode za prikupljanje znanja i njena primena,
- ✓ tumačenje prikupljenog znanja i
- ✓ oblikovanje prikupljenog znanja u oblik pogodan za čuvanje u bazi znanja.

Postoje brojne metode za snimanje znanja. Najčešće su:

- ✓ Intervju;
- ✓ Posmatranje;
- ✓ Brejnstorming i elektronski brejnstorming;
- ✓ Nominalna grupna tehnika;
- ✓ Metoda uporednog poređenja;
- ✓ Metoda sortiranja karata;

- ✓ Panel metoda;
- ✓ Delfi;
- ✓ Metoda ekspanzije, kontrakcije i ukrštanja;
- ✓ Analiza protokola;
- ✓ Skala ocena, i
- ✓ Konceptualno mapiranje.

Intervju je najčešće korišćena tehnika za dobijanje znanja od eksperata i treba ga dobro pripremiti. Može biti visoko ili slabo strukturiran u zavisnosti od vrste znanja koje se prikuplja.

Metoda posmatranja prati kako ekspert rešava problem i iz ovako prikupljenog znanja se pokušava doći do znanja.

Breinstorming je metoda u kojoj učestvuje grupa eksperata u rešavanju problema. Sastoji se u generisanju ideja od strane svih eksperata, gde se sve ideje ravnopravno beleže a do konačnog rešenja se dolazi biranjem najprihvatljivijih predloga.

Nominalna grupna tehnika ima sličnu proceduru rada kao breinstorming, pri čemu je razlika da sve ideje (rešenja) eksperti beleže tajno a na kraju procesa projektant znanja sumira sve alternative i njihove pozitivne i negativne strane daje ekspertima, pri čemu se glasanjem bira najprihvatljivija alternativa.

Panel metoda je tehnika snimanja znanja gde više eksperata na panel tabli ispisuje rešenje problema i na taj način formiraju bazu znanja.

Delfi metoda se sastoji u prikupljanju rešenja od svih eksperata. Sadrži niz sukcesivnih koraka gde u svakoj narednoj iteraciji eksperti imaju uvid u prethodno doneto rešenje i pokušavaju ponovo da reše problem, sve dok se ne usaglase mišljenja.

Analiza protokola se koristi kada treba zabeležiti jedan dijagnostički proces, pri čemu se ekspert posmatra dok radi dijagnozu i učestvuje u rešavanju problema.

Skala ocena služi za organizaciju i ocenu prikupljenih zapisa prošlog iskustva (slučajeva).

Konceptualno mapiranje se koristi kada rešavanje određenog problema zahteva da se problem predstavi u vidu semantičke mreže, gde se preko skupa objekata i relacija među njima može doći do rešenja.

Projektovanje SMZ

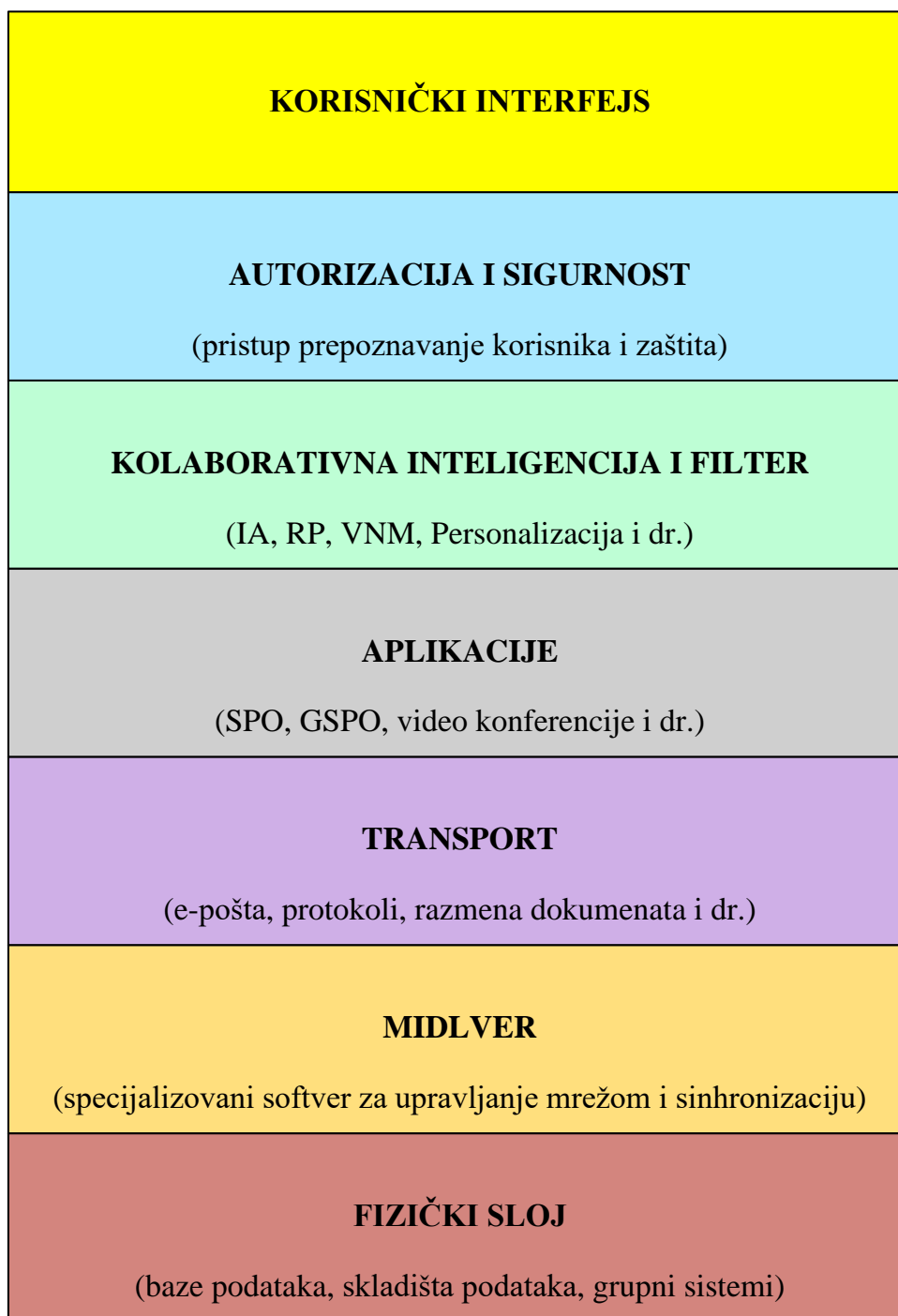
Poseban akcenat stavlja se na fazu projektovanja SMZ iz razloga što u ovoj fazi dolazi do integracije IT infrastrukture i znanja koje se želi koristiti. Potrebno je uskladiti SMZ sa postojećim informacionim sistemom kao i odrediti systemske komponente SMZ koji će predstavljati tehničku dokumentaciju glavnog projekta SMZ.

Preko informacionih i komunikacionih tehnologija moguće je ostvariti kvalitetan SMZ. Potrebno je usmeriti se na tacitna znanja i izgraditi SMZ koji omogućava da se dođe ne samo do znanja, već i do izvora znanja. Neophodno je realno sagledati postojeće stanje sistema i projektovati SMZ koji će biti okrenut ka saradnji i grupnom radu.

Arhitektura SMZ

SMZ se sastoji iz sledećih slojeva (naredna slika), videti u [112]:

1. Korisnički interfejs (sredstvo preko koga korisnik upravlja sa SMZ).
2. Sloj za autorizaciju i sigurnost (obezbeđuje pristup SMZ i osigurava da na određeno radno mesto može samo da uđe ovlašćeni korisnik).
3. Kolaborativno-inteligentni sloj (centralni deo SMZ i u njemu se obezbeđuje validacija znanja).
4. Aplikacioni sloj (sadrži aplikacije, između ostalih i aplikacije poslovne inteligencije koje su usklađene sa poslovnim procesom i deo su njega).
5. Transportni sloj (obezbeđuje komunikaciju i razmenu informacija kroz SMZ).
6. Midlver, sve aplikacije koje omogućavaju integraciju celog SMZ, kao i već postojećeg informacionog sistema.
7. Fizički sloj, svi uređaji koji čuvaju podatke, informacije i znanja, što mogu biti razni dokumenti, video i audio zapisi, baze podataka, itd.



Slika 7.3. Arhitektura SMZ

Organizovanje znanja

Snimljeno i zabeleženo znanje potrebno je transformisati u oblik koji dozvoljava da se zabeleženo znanje čuva u bazi znanja i tako bude dostupno i ostalim zainteresovanim u organizaciji. Da bi znanje moglo da se ubaci u bazu znanja potrebno je da se uobliči u određenu formu. Znanje može da se organizuje preko sledećih koncepata:

- ✓ mape znanja (semantičke mreže),
- ✓ tabela odlučivanja (ramovi znanja),
- ✓ drvo odlučivanja,
- ✓ ako-tada pravila (produkciona pravila),
- ✓ fazi pravila,
- ✓ slučajevi i
- ✓ inteligentni softverski agenti (IA).

Mape znanja

Mape znanja predstavljaju način vizualizacije tacitnog znanja. One treba da prikažu znanje preko skupa objekata i relacija među njima.

Tabele odlučivanja

Preko tabele odlučivanja znanje se može prikazati u tabelarnom obliku. Redovi tabele odlučivanja predstavljaju alternative, a kolone predstavljaju kriterijume, pri čemu se uz pomoć metoda višekriterijumskog odlučivanja može odrediti najprihvatljivija alternativa.

Drvo odlučivanja

Drvo odlučivanja predstavlja grafičku interpretaciju tabele odlučivanja, ali i druge koncepte (ako-tada pravila, analizu odlučivanja). Drvo odlučivanja je alat koji vizuelno predstavlja znanje i lakši pregled tog znanja. Na samom grafu kvadrati predstavljaju trenutke odlučivanja, dok krugovi predstavljaju čvorove mogućnosti.

Ako-tada pravila

Ako-tada pravila omogućavaju da se izradi sistem pravila i da konkretan problem odlučivanja pokreće mehanizme zaključivanja, koji prolazeći kroz sva, za problem bitna pravila, predlaže rešenje.

Fazi pravila i fazi sistemi

Fazi pravila se koriste kada se javi potreba za predstavljanje pojmova koji su po svojoj prirodi rasplinuti ili nejasni.

Slučajevi

Slučajevi su zapisi iskustva koje može da bude eksplicitno i tacitno. Predstavljaju zapise prošlog iskustva na osnovu kojih je donošena određena odluka i sastoje se iz dva dela: opis problema i rešenje problema. Slučajevi se predstavljaju preko skupa kriterijuma koji imaju određene vrednosti.

Inteligentni agenti

Predstavljaju sisteme koji se nalaze u informacionom sistemu, a sposobni su da samostalno deluju. Oni nisu namenjeni da zamene ljudsku inteligenciju već da je podrže. IA su ciljno orjentisani i izvode akcije za koje su namenjeni. Imaju sposobnost da uče o ponašanju određenog korisnika i da na osnovu toga sami počnu da obavljaju rutinske poslove. Imaju pristup bazi znanja i mogu da reše automatizovane radnje u smislu poboljšanja baze znanja. Kada znanje od epizodnog pređe u proceduralno znanje, tada IA može da automatizuje poslovni proces. IA služe da olakšaju i ubrzaju proces donošenja odluka korisnika.

Testiranje SMZ

Testiranje se sastoji iz 2 procedure: verifikacija sistema i validacija sistema.

Verifikaciona procedura obezbeđuje ispravnost sistema, tj. proverava da li je ispunjena sama svrha zbog koje su programi napravljeni. Proveravaju se dakle tehničke performanse sistema. Verifikacionom procedurom se proverava da li baza znanja ima grešaka (redundantnost, višeznačnost, nepotpunost). Redundantnost označava da se u bazi znanja nalazi više znanja nego što je potrebno. To je nekorisno znanje, odnosno duplikati znanja i opširno zabeležena znanja koja ne koriste pri rešavanju problema.

Konzistentnost pretpostavlja, da isti problem treba da se reši na isti način. Potpuno govori da znanje treba da bude funkcionalno, tj. da može da se koristi. Verifikacija se vrši na osnovu test primera, kako u graničnim vrednostima, tako i u nemogućim vrednostima. Verifikacija SMZ obuhvata dva načina testiranja: verifikovanje forme baze znanja i verifikovanje funkcionalnosti baze znanja.

Validacija proverava u kojoj meri sistem odgovara zahtevima korisnika. Validacija obezbeđuje da SMZ daje približno dobre rezultate kao i ljudski ekspert. Za razliku od verifikacije, validacija je dugotrajan proces. Nakon uspešno izvršenog logičkog testiranja sledi testiranje zadovoljstva korisnika. Cilj validacije je provera ponašanja sistema u realnim uslovima.

Implementacija SMZ

Ovaj korak predstavlja preslikavanje projekta u realnost. Postoji više načina za realizaciju SMZ. Može se implementirati sopstveni sistem, može se kupiti gotov sistem i nadograditi sa zahtevima iz projekta ili se ceo posao može poveriti trećoj strani (outsourcing). Da bi se

upoznali sa SMZ i naučili kako sistem radi, potrebno je da većina korisnika prođe određenu obuku. Obuka je neophodna iz razloga da bi se korisnici upoznali sa radom ovog sistema i počeli da ga prihvataju. U toku procesa obuke korisnici se upoznaju sa korisničkim interfejsom SMZ, sa osobinama SMZ kao i sa tim kako se SMZ uklapa sa ciljevima organizacije.

Održavanje SMZ

Da bi se SMZ jednostavno održavao treba da bude projektovan na način koji to omogućava. Sistem treba da bude izgrađen tako da omogućava laku izmenu i nadogradnju sistema. Nadogradnja predstavlja dodavanje novih pogodnosti sistema. Svaki SMZ treba da ima i određenu dokumentaciju iz koje može da se vidi na koji način se najlakše može održavati i nadograđivati sistem. Kada se projektuje i implementira SMZ, koristi se dokumentacija organizacione strukture i informacionih sistema.

Primena SMZ

SMZ se koristi za rešavanje sledećih problema:

- ✓ **Dijagnoza:** koristi se ekspertsko znanje za dijagnosticiranje određenog stanja poslovnog procesa. Dijagnoza funkcioniše tako što se pamte prethodni zapisi i iskustva koji vode do određene dijagnoze.
- ✓ **Trening:** sistemi za trening treba da pomognu mlađim saradnicima da se upoznaju sa prirodom posla i da uče na osnovu iskustva starijih kolega. Ovi sistemi omogućavaju da se vrši šta-ako analiza, kako bi se vršila simulacija poslovnih situacija.
- ✓ **Planiranje:** omogućava da se pre preduzimanja akcija sprovedu koraci koji će potvrditi da akcija koju treba izvršiti jeste najbolje rešenje.
- ✓ **Predviđanje:** je najteža oblast u poslovanju. Sistemi za predviđanje mogu na osnovu ulaznih parametara problema, da sa određenom verovatnoćom (validacija) predvide, koliko je verovatno da će se desiti određena situacija koja zahteva određeno rešenje.

Prenos i deljenje znanja

Da bi SMZ zaživeo i da bi organizacija uspešnije poslovala u organizaciji treba podsticati procese prenosa i deljenje znanja. Jednom izgrađena baza znanja je podložna promenama. Cilj svakog SMZ jeste da podstiče deljenje i prenos znanja. Treba da bude jasno da transfer

znanja može samo da prenese deo ljudske ekspertize (tacitnog znanja). U svakom slučaju prenošenje znanja omogućava deljenje znanja.

Transfer (prenos) znanja je termin koji se odnosi na mehanička svojstva transfera znanja. Tehnologija, dakle, nije preduslov za menadžment znanja jer ljudi predstavljaju ključni faktor. Znanje može da se deli na tri načina, videti [6]:

- ✓ grupni sekvencijalni transfer,
- ✓ eksplicitni međutimski transfer i
- ✓ tacitni prenos znanja.

Grupni sekvencijalni transfer označava prenos znanja među članovima tima. Tokom izvršenja svojih radnih zadataka, zaposleni stiču iskustvo i specijalizuju se u obavljanju posla tako da imaju mogućnost da vrše evaluaciju stečenog znanja.

Eksplicitni međutimski transfer vrši deljenje i prenos znanja među različitim timovima, koji obavljaju isti posao. Znanje i iskustvo koje stiče jedan tim tokom radnog procesa, može se na ovaj način preneti drugim timovima i zaposlenima u preduzeću.

Tacitni prenos znanja je najsloženiji, zato što zahteva da se znanje za rešavanje određenog problema primeni u potpuno drugačijem kontekstu. Ovakvo znanje teško može biti sačuvano u bazama znanja. Da bi se rešio problem u novom kontekstu potrebno je posedovati viziju i iskustvo.

Sve tri vrste deljenja i prenosa znanja treba potpomognuti sa informacionim tehnologijama, među kojima su:

- ✓ Intranet,
- ✓ Internet,
- ✓ Ekstranet, i
- ✓ GSPO.

Intranet je privatna računarska mreža jednog poslovnog sistema, koja koristi internetske standardne protokole kako bi zaposlenima omogućila jednostavnu komunikaciju kao i pristup informacijama firme. Svi zaposleni imaju pristup bazi znanja iz koje mogu da crpe korisno znanje, ali isto tako mogu da doprinose da baza znanja postane bolja. Komunikacija među članovima organizacije je omogućena uz pomoć alata kao što su:

- ✓ elektronska pošta,
- ✓ videokonferencija i
- ✓ aplikacija za elektronske sastanke itd.

Intranet je mreža koja spaja definisani skup klijenata koji koriste standardne Internet protokole (npr. TCP/IP i HTTP) ili se može interpretirati kao mreža priključaka iza firewalla (može i iza nekoliko firewalla) spojenih bezbednim (verovatno virtualnim) mrežama. Implementacija jednog softverskog sistema kao što je Intranet aplikacija omogućava njeno brzo i kvalitetno korišćenje. Osnovne funkcije Intraneta su elektronska pošta, zajedničko korišćenje fajlova, upravljanje pristupom informacija, pretraživanje informacija i upravljanje mrežom. Intranet ima brojnu primenu u organizaciji. Omogućava razmenu informacija među zaposlenima, daje podršku saradnji, odlučivanju i upravljanju kao i generisanje i dostavu poslovne dokumentacije.

Internet predstavlja svetsku računarsku mrežu, koja se sastoji od miliona računara raspoređenih širom sveta. On omogućava razmenu podataka između distribuiranih aplikacija. Razlika između Interneta i Intraneta je u tome što je Intranet privatna mreža u vlasništvu neke organizacije, dok Internet nije u vlasništvu ni jedne firme niti osobe. Internet omogućava da organizacija dobija znanja i van svog Intraneta. Internet omogućava kontakt sa različitim izvorima znanja. Ovim putem organizacija saznaje mišljenje korisnika o svojim proizvodima i uslugama, kao i da dobije predloge za poboljšanje i eventualne zahteve za nove korisnike. Internet pomaže da se ostvari koncept B2C (poslovanje okrenuto ka kupcu), SCM 102 (Supply Chain Management – koncept koji zahteva da se poseduje pravi proizvod na pravom mestu, u prvo vreme, po pravoj ceni i u pravom stanju) i CRM 103 (Customer Relationship Management – koncept povezan sa SCM koji zahteva da se poboljšaju odnosi sa dobavljačima i kupcima radi dolaženja do boljnog znanja).

Sigurnost i privatnost sve su značajniji aspekti koji se moraju razmetrati pri dizajnu mreže unutar preduzeća. Različite mogućnosti zloupotrebe i neovlašćenih upada mogu dovesti do gubitka informacija u elektronskom obliku, izvršenja različitih krivičnih dela kao i izazivanja nefunkcionalnosti delova informacionog sistema i oštećenja računarske opreme.

Ekstranet je prošireni Intranet u kome učestvuju i organizacije sa kojima matična firma saraduje (dobavljači, distributeri...). Pošto te organizacije imaju neke zajedničke ciljeve onda bi trebalo i da dele znanje. Na primer Tošiba i njen informacioni sistem (FYI, For Your

Information) u mreži Ekstranet sa preko 350 severnoameričkih distributera i 2000 uslužnih predstavnika, samo potvrđuju konstatacije o ulozi informacione tehnologije u menadžment znanja.

GSPO, kao što je objašnjeno u poglavlju o poslovnoj inteligenciji, omogućavaju da se ostvari saradnja, deljenje znanja i kvalitetnije odlučivanje za probleme udaljenosti članova tima i različitih vremenskih zona.

Pretvaranje podataka i informacija u znanja i širenje tog znanja kroz organizaciju je kontinualan proces čija je pretpostavka razvoj informacione infrastrukture i informacionog menadžmenta. Prethodno opisan koncept menadžmenta znanja, jasno ukazuje na njegov sve veći značaj u poslovanju i sticanju konkurentne prednosti preduzeća.

Intelektualni kapital i procesi osvajanja, čuvanja, evaluacije i prenosa znanja dokazuju da je menadžment znanja važan element procesa rada, kome se sve više daje značaj u poslovnom odlučivanju. Integrisane baze podataka, Internet, virtuelne organizacije i “on line” poslovne transakcije samo su neki od oblika sinergije informacionih tehnologija sa menadžmentom znanja.

U nastavku disertacije daje se akcenat na izabrane tehnike modelovanja znanja.

7.4. Tehnike modelovanja znanja

Postoje mnoge tehnike za modelovanje znanje, videti [2], najčešće su:

- ✓ CommonKADS
- ✓ Protégé 2010
- ✓ Višeperspektivno modelovanje
- ✓ Unified Modelling Language (UML)

CommonKADS

CommonKADS obezbeđuje alate za korporativno upravljanje znanjem i uključuje metode koje rade detaljne analize intenzivnih zadataka znanja i procesa. Paket modela predstavlja srž metodologije CommonKADS-a. Paket obezbeđuje modelovanje organizacije, zadatke koji se izvršavaju, agente koji su odgovorni za izvršavanje zadataka, samo znanje i dizajn sistema upravljanja znanjem.

Pod *organizacionim modelom* se smatra studija slučaja za sistem znanja. Studija je sprovedena na osnovu problema i prilika, ona može da se fokusira na oblasti kao što su struktura, procesi, ljudi, kultura, ljudska moć, resursi, kvarovi u procesima i imovina znanja. Organizacioni model nudi tri glavne svrhe: identifikacija oblasti u organizaciji gde aplikacije zasnovane na znanju mogu da budu implementirane, identifikacija uticaja aplikacije zasnovane na znanju, na organizaciju kada se implementira i na kraju on pruža sistemskim programerima osećaj gde će u organizaciji ta aplikacija biti primenjena.

Svrha modela agenta je razumevanje uloga igranih od strane različitih agenata, kada se izvodi zadatak. Agenti mogu biti ljudi, kompjuteri ili bilo koji drugi entiteti koji mogu da obavljaju zadatak. Model agenta specificira karakteristike agenata, njihov autoritet za izvođenje zadatka i ograničenja.

Svrha modela zadatka je, pružanje uvida u mogući uticaj koji će imati sistem znanja na organizaciju. Model zadatka se odnosi na karakteristike poslovnog procesa kao što su: ulazi i izlazi, preduslovi, performanse i kvalitet. Funkcija agenata koji će vršiti procesiranje, strukturno uklapanje ovih agenata, tokovi znanja između agenata, njihova ukupna kontrola, znanje i stručnost agenata i resursa dostupnih za izvršavanje poslovnih procesa.

Model znanja se koristi za opisivanje znanja vezanog za aplikacije koje se koriste za izvođenje zadataka i uloga znanja u rešavanju problema. Model znanja u CommonKADSu ima tri kategorije znanja: znanje zadataka koje opisuje redosled, izvršavanja koraka, znanje zaključivanja koje opisuje korake koji se izvršavaju korišćenjem domenskog znanja. I naravno domensko znanje koje samo sadrži svoja svojstva, koncepte, veze i drugo u aplikacionom domenu.

Komunikacioni model opisuje komunikaciju između agenata koja je neophodna za izvršavanje zadataka.

Model dizajna je tehnička specifikacija sistema koja se odnosi na njegovu arhitekturu, platformu, module, konstrukcije i računске mehanizme. On spaja sve ostale modele. CommonKADS spaja objektno orjentisani razvojni proces i koristi UML notaciju kao što su dijagrami klasa, slučajevi korišćenja, dijagrami aktivnosti i dijagrami stanja. CommonKADS takođe ima svoju grafičku notaciju za dekompoziciju zadataka, strukture zaključivanja i stvaranje domenskih šema.

Protégé 2010

Pre 13 godina je kreirana prva verzija Protége. Originalni alat je bio mala aplikacija čiji je cilj bio izgradnja alata za prikupljanje znanja za nekoliko specijalizovanih programa. Najnovija verzija, Protégé-2010, obuhvata "Open Knowledge Base Connectivity" model znanja, može da se izvršava na različitim platformama, podržava proširenja koja imaju prilagođen korisnički interfejs, i koristi je preko 300 pojedinaca i istraživačkih grupa. Protégé aplikacije predstavljaju grupu alata koji su razvijani preko deset godina, počevši od jednostavnog programa koji je pomagao pri konstrukciji specijalizovanih baza znanja, do grupe alata za kreiranje i održavanje opštih baza znanja. Protégé nije ni ekspertni sistem ni program koji gradi ekspertni sistem direktno. On je alat koji pomaže korisnicima da izgrade druge alate, koji su prilagođeni za pomaganje u prikupljanju znanja za ekspertne sisteme u specifičnim oblastima primene. Cilj Protégé 2010 je izgradnja baze znanja za višekratnu upotrebu u različitim okvirima za modelovanje usvajanjem standardnog jezika za prikaz i postavljanje temelja za rešavanje pitanja skalabilnosti u inženjerstvu znanja, videti [143].

Poslednja verzija, Protégé 2000, je razvijena od strane Musen-a i kolega iz Stanford Medical Informatics. Protégé ontologija (koja modeluje domen) ima klase, slotove, aspekte i aksiome. Klase predstavljaju reprezentaciju domen koncepata. Podklasa može da ima sva pojavljivanja klase. Dozvoljeno je višeklasno nasleđivanje: klasa može da ima dve ili više super klase; podržava i meta-klas koncept. Slotovi su svojstva ili atributi klase. Postoje dva oblika slotova. Sopstveni slotovi definišu unutrašnja svojstva klase ili individualna pojavljivanja okvira. Šablon slotovi su prikačeni za okvir klase kako bi definisali attribute njihove instance, što za uzvrat definiše specifične vrednosti za slotove. Slotovi su objekti prve klase i mogu biti korišćeni globalno ili lokalno. Aspekti su svojstva ili atributi slotova i koriste se za određivanje ograničenja na vrednostima slotova. Ograničenja uključuju kardinalnost slotova (broj vrednosti koje slot može da ima), tip vrednosti (kao što je integer, string...) i minimalne i maksimalne vrednosti za slot. Aksiome definišu dodatna ograničenja za okvire; oni mogu da ukazuju na vrednosti zajedno.

Informacije slučajeva se dobijaju korišćenjem „on-line“ formi. Sastoje se od grupe grafičkih ulaznih polja i pružaju korisnički interfejs koji je lak za korišćenje. To automatski pruža formu za dobijanje instance klase, kada korisnik definiše klasu i pridoda šablon slot tome. Korisnik može da prilagodi formu menjanjem rasporeda ili labela na formi i može da izabere različite načine za prikaz i dobijanje vrednosti slotova. Proces sticanja znanja u Protégé 2010 se sastoji od tri koraka. Prvo, klase i njihovi šablon slotovi moraju da budu definisani. Drugo, forma za dobijanje instanci klase mora da bude izložena. Na kraju se dobijaju instance klase. Svaka klasa ima svoju formu koja se koristi za dobijanje instance klase.

Baza znanja u Protégé se razvija u sekvencama. Prvo se definišu koncepti i njihove veze. Drugo, domen eksperti unose svoje znanje domenske oblasti korišćenjem specifičnog domenskog alata za dobijanje znanja. Na kraju se koriste tehnike rešavanja problema za dobijanje odgovora na pitanja i probleme domena korišćenjem baze znanja.

Višeperspektivno modelovanje

Višeperspektivno modelovanje omogućava korišćenje više tehnika zajedno, gde je svaka od tehnika odgovarajuća za modelovanje određenog dela znanja. Znanje organizacije je veoma kompleksno i heterogeno i ne postoji jedan metod koji može da modeluje sve ovo tačno i na odgovarajući način. Ova tehnika modelovanja se koristi za pravljenje različitih modela iste stvari, kako bi se omogućila različita gledišta. Ona se koristi za prikupljanje zahteva za razvoj softverskih projekata.

Višeperspektivno modelovanje ima šest kategorija: „šta“, „kako“, „kada“, „ko“, „gde“ i „zašto“. „Šta“ se odnosi na resurse date u formi znanja o stvarima. Obuhvata koncepte, fizičke objekte i stanja. „Kako“ se odnosi na procese, tj. na znanje o akcijama i događajima. Uključuje znanje o akcijama koje su potrebne ako se dogodi određeni događaj; koje akcije će dostići određeno stanje; potreban ili poželjan redosled akcija. „Kada“ se odnosi na određivanje vremena i ograničenja. To je znanje o vremenu kada će se akcija ili događaj desiti, ili treba da se desi. „Ko“ se odnosi na agente (ljude ili automate). To je znanje o agentima koji obavljaju određene akcije, njihovim sposobnostima i ovlašćenjima za izvođenje određenih akcija. „Gde“ se odnosi na znanje o komunikacijama, gde je znanje potrebno i odakle dolazi, i kako staviti i pružiti informacije. „Zašto“ se odnosi na znanje o razlozima, argumentima, empiriskim studijama i opravdanjima za stvari koje su urađene i načinu na koji su urađene.

Različiti analitičari su uključeni u različite faze projekta i imaju različite perspektive o projektu. Različite perspektive zahtevaju različite nivoe apstrakcije. Na primer, u projektu za razvoja sistema, menadžer ima celokupni pogled na projekat; analitičar sistema se bavi zahtevima za predloženi sistem; projektant sistema se koncentriše na aspekte dizajna; programer se bavi izradom programskog koda za svaki modul.

Odgovarajuća tehnika za modelovanje za višeperspektivno modelovanje može biti odabrana od poslovnih menadžerskih tehnika, softverskih inženjerskih tehnika i inženjerskih tehnika znanja. Ipak, da bi se ostvario višeperspektivan prikaz znanja, postoje tri glavne metode, i to: CommonKADS, UML i IDEF. CommonKADS se koristi za dekompoziciju zadataka, strukturu zaključivanja i domen šemu. UML se koristi za prikaz dijagrama klasa, dijagrame slučajeva korišćenja, dijagram aktivnosti i dijagrama stanja. IDEF je odgovarajući za

funkcionalno modelovanje, IDEF1 se koristi za specifikaciju veza među objektima, IDEF1X se koristi za objektno orjentisano projektovanje i IDEF5 prikazuje opis ontologije.

Unified Modelling Language (UML)

UML zajedno sa Object Constraint jezikom (OCL) predstavlja standard za objektno modelovanje definisan od strane Object Management Group - (OMG). UML se koristi za vizualizovanje, specificiranje, izgradnju i dokumentovanje šeme softvera koji se koristi za različite tipove intenzivnih softverskih sistema. Neki projekti ukazuju na uspešnost UML modelovanja sistema i u drugim oblastima. Generalno, UML može da se koristi za pronalaženje klasifikacija koje predstavljaju osnovno znanje o različitim stvarima. Ove klasifikacije mogu da budu zasnovane na hijerarhiji podklasa, agregaciji, vezama asocijacije ili bilo kojoj kombinaciji od ovih. Takve klasifikacije mogu da se koriste za bolje razumevanje neke oblasti, za vođenje otkrivanja znanja, za pripremu bolje prezentacije o nekoj temi, uključujući Web prezentacije ili za podršku procesu učenja.

Većina znanja je u formi prirodnog jezika. Postoje mnogi materijali koji mogu da pomognu u razumevanju neke teme, ali obično ne pružaju kompletan okvir za opšte ideje prezentovane u prirodnom jeziku. Jedna od poteškoća koja se javlja prilikom pronalaženja i dobijanja znanja se odnosi na dvoznačnost prirodnog jezika. Iako je UML kreiran i korišćen za dizajn softvera, može se primeniti i u mnogim drugim oblastima. Ovaj tip modelovanja je objektno orjentisan što znači kad god se sistem modeluje, njegove komponente postaju apstraktni objekti koji imaju neke svoje karakteristike (atributi) i funkcije (odgovornosti). Klasa je kolekcija ovih apstraktnih objekata. Klasa predstavlja objekte koji imaju slične attribute, semantiku i operacije, opisano u [16].

Dijagrami su grafičke reprezentacije grupe elemenata koje se koriste za vizuelizaciju sistema iz različitih uglova. Postoje različiti UML dijagrami koji mogu biti korisni za modelovanje znanja, dijagrami klasa, dijagrami stanja, dijagrami sekvenci, komunikacije, dijagrami aktivnosti i drugi. UML ima semantička pravila koja se primenjuju na modele kako bi bili semantički ispravno formirani. Semantička pravila obuhvataju imena, opseg, vidljivost, integritet i izvršenje. Dijagrami klasa sadrže klase i veze. Klase mogu biti opisane njihovim imenom, karakteristikama, funkcijama. Grafički se predstavljaju kao pravougaonici. Linije ili strele između klasa predstavljaju veze. Najčešći tipovi veza su agregacija, generalizacija i imenovana asocijacija. Agregacija se tretira kao specijalna forma asocijacije.

UML može da se koristi za modelovanje znanja jer podržava objektivne koncepte kao što je Protégé 2010 koji je razvijen korišćenjem objektno orjentisanog jezika Java, i CommonKADS koji koristi UML dijagrame za proces modelovanja znanja. Ranije verzije UML-a nisu bile namenjene za podršku sistemu zasnovanom na pravilima, ali zahvaljujući

novim karakteristikama UML-a, OCL-u i objektno orjentisanim programskim jezicima, modelari znanja su počeli intenzivno da ga koriste. UML može da se koristi za modelovanje znanja u razvoju inteligentnih sistema, ekspertnih sistema i drugih sistema za upravljanje znanjem, videti u [2].

Poređenje tehnika

Među predhodno pomenutim tehnikama, CommonKADS je jedina tehnologija koja može da se smatra metodologijom inženjerstva znanja. Sve ove tehnike podržavaju objektno orjentisani pristup modelovanju aktivnosti i njihovi modeli su nezavisni od platforme. Protégé se, za razliku od CommonKADS, višeperspektivno modelovanja i UML-a, ne koristi za crtanje vizuelnih modela ili dijagrama, već predstavlja alat koji omogućava unos znanja u bazu znanja. Deo za modelovanje je već ugrađen u Protégé i kao takav nevidljiv za korisnike. UML je standard definisan od strane OMG, dok ostale tehnike nisu standardizovane na formalni način. Postoji obilje dokumentacije za sve ove tehnike u različitim formama. Većina tehnika je u razvoju. Ove tehnike su korisne za modelovanje različitih oblasti, uključujući medicinu, inženjerstvo, trgovinu, socijalne nauke i drugo, prema [2].

Tehnike Karakteristike	CommonKADS	Protégé 2000	Multi- perspective	UML
Metodologija inženjeringa znanja	✓			
Objektno- orjentisani pristup	✓	✓	✓	✓
Nezavisnost od platforme	✓	✓	✓	✓
Hibridni pristup	✓		✓	✓
Alat za izmene		✓		
Dokumentacija	✓	✓	✓	✓
Razvoj		✓	✓	✓
Domen	Medicina, pravo, inženjerstvo, i društvene nauke	Medicina, pravo, inženjerstvo , i društvene nauke	Medicina, pravo, inženjerstvo, i društvene nauke	Medicina, pravo, inženjerstvo, i društvene nauke

Tabela 7.1. Osnovne karakteristike predstavljenih tehnika modelovanja

7.5. Studija slučaja – uvođenje menadžmenta znanja

U ovom delu disertacije biće predstavljen realan primer korišćenja sistema menadžmenta znanja. Radi se o „Microsoft SharePointu“, koji predstavlja platformnu zasnovanu na Veb aplikacijama razvijenu od strane „Microsoft“-a. Prva verzija pojavila se 2001. godine. SharePoint se u početku poistvećivao sa intranetom i sistemom za upravljanje dokumentima ali su novije verzije donele mnogobrojne nove mogućnosti i unapredjenja. SharePoint poseduje interfejs sličan Microsoft Office setu programa i u velikoj meri je integrisan sa Office paketom. Veb alati u okviru SharePointa-a dizajnirani su tako da pružaju široke mogućnosti analitičarima koji ne poseduju velika tehnička znanja. Njegovo korišćenje omogućava kreiranje internet portala, upravljanje dokumentacijom, olakšava kolaboraciju zaposlenih, implementaciju društvenih mreža, ekstranet, pretraživanje sadržaja, upotrebu poslovne inteligencije, komunikaciju zaposlenih itd. Prema podacima iz Microsoft-a, SharePoint koristi 78% kompanija koje se nalaze na listi „Fortune 500“ svetski poznatog poslovnog magazina „Fortune“.

Glavne funkcionalnosti SharePoint sistema su sledeće:

- ✓ Veb portali (sites),
- ✓ Kolaboracija (Communities),
- ✓ Poslovna rešenja (Composites),
- ✓ Upravljanje sadržajem (Content),
- ✓ Poslovna inteligencija (Insights), i
- ✓ Pretraživanje baze znanja (Search).



Slika 7.4. Funkcionalnosti softverskog rešenja Microsoft SharePoint

- ✓ **Veb portali (Sites):** Odnosi se na celovito rešenje za poslovne veb portale koji sadrže sve alate koje bi potencijalni korisnik mogao koristiti za stvaranje bilo kakvog veb sajta. Osnovne funkcionalnosti su tako implementire da korisnik ne mora imati nikakva tehnička znanja kako bi pustio u rad svoj veb sajt. Uz samo nekoliko intuitivnih odgovora moguće je jednostavno dodavanje i uklanjanje delova veb sajta. Ovu funkcionalnost moguće je implementirati u lokalu ili hostovati on-line (npr na „cloud“ servisu).
- ✓ **Platforma za kolaboraciju (Communities):** Sadrži alate potrebne za razmenu ideja, pronalaženje saradnika i stručnih resursa kao i stvaranje kreiranje sadržaja. Sa poznatim alatima, koji se koriste relativno lako, moguće je stvoriti wikije, blogove, „news feed“-ove, timske veb stranice, lične profile, itd. Osim kreiranja konkretnog sadržaja pomaže u stvaranju veza i razmeni ideja i znanja.
- ✓ **Poslovna rešenja (Composites):** Sadrži sve elemente potrebne za sastavljanje, povezivanje i konfigurisanje poslovnih rešenja za saradnju. Korisniku se nudi mnogo sastavnih blokova, od kalendara i zadataka do grafikona i poslovnih podataka, koji su namenjeni brzom stvaranju rešenja za saradnju koja se prikazuju u okviru veb pretraživača. Takođe, omogućen je rad sa podacima iz drugih eksternih sistema kao da se oni nalaze u okviru sistema SharePoint.
- ✓ **Sadržaj (Content):** Objedinjuje tradicionalno upravljanje sadržajem, društvene mreže i pretraživanje. Omogućuje preciznije i efikasnije upravljanje sadržajem i usko je povezan sa paketom Microsoft Office što znači da nudi poznato korisničko okruženje većini zaposlenih. Takođe pojednostavljuje označavanje i klasifikaciju sadržaja i tako korisnicima olakšava pronalaženje, razmenu i korišćenje informacija. Obraduje različite vrste informacija nezavisno od toga da li se radi o dokumentima, web stranicama ili društvenom sadržaju. Ova funkcionalnost takođe omogućuje korišćenje višefaznih pravilnika za definisanje vremena čuvanja sadržaja. Na taj način moguće je upravljati vremenom kada je sadržaj potrebno arhivirati ili proglasiti zastarelim. Ti pravilnici omogućuju efikasno skladištenje sadržaja umesto da se isti čuva neograničeno dugo. Nevažan sadržaj se može premestiti izvan indeksa pretraživanja, kako bi rezultati pretraživanja sadržali isključivo relevantan sadržaj koji je tada moguće pronaći u što kraćem roku i na taj način povećati produktivnost.
- ✓ **Poslovna inteligencija (Insights):** Omogućuje svim korisnicima pristup poslovnim informacijama koje su im potrebne za donošenje boljih poslovnih odluka. U takve

poslovne odluke korisnici mogu biti sigurni jer su donešene uz pomoć skupa robusnih alata koji omogućuju pristup pravim informacijama u pravo vreme, njihovu analizu i jednostavno zajedničko korišćenje. SharePoint Online 2013 je ključna komponenta Microsoft-ove platforme za poslovnu inteligenciju koja donosi nekoliko prednosti kao što su povezivanje korisnika radi saradnje, smanjenje troškova pomoću objedinjene infrastrukture i brz odgovor na poslovne potrebe.

- ✓ **Pretraživanje (Search):** Funkcionalnost koja omogućuje pronalaženje potrebnih podataka koji su korisnicima potrebni za obavljanje posla bilo da se radi o pretraživanju intraneta, traženju osoba kako bi se brže i jednostavnije povezali i razmenjivali ideje ili nečemu trećem. Interaktivno i vizuelno okruženje za pretraživanje pomoći će u nastojanju da se mnoštvo podataka dovede u red pri čemu su najviše od koristi sužavanje rezultata pretraživanja na osnovu metapodataka, rangiranje relevantnosti prema broju klikova i slično. Okruženje za pretraživanje je moguće dodatno prilagoditi tj. dodati vlastiti vokabular, podesiti stepen relevantnosti i koristiti specifične podatke za svaku osobu kako bi korisnik mogao biti siguran da će dobiti očekivane rezultate.

Microsoft SharePoint obezbeđuje integrisani skup serverskih aplikacija koje se lako koriste i poboljšavaju efikasnost organizacije i interakciju ljudi, sadržaja, procesa i poslovnih aplikacija. Sledi opis najvažnijih prednosti koje proizilaze iz korišćenja SharePoint platforme:

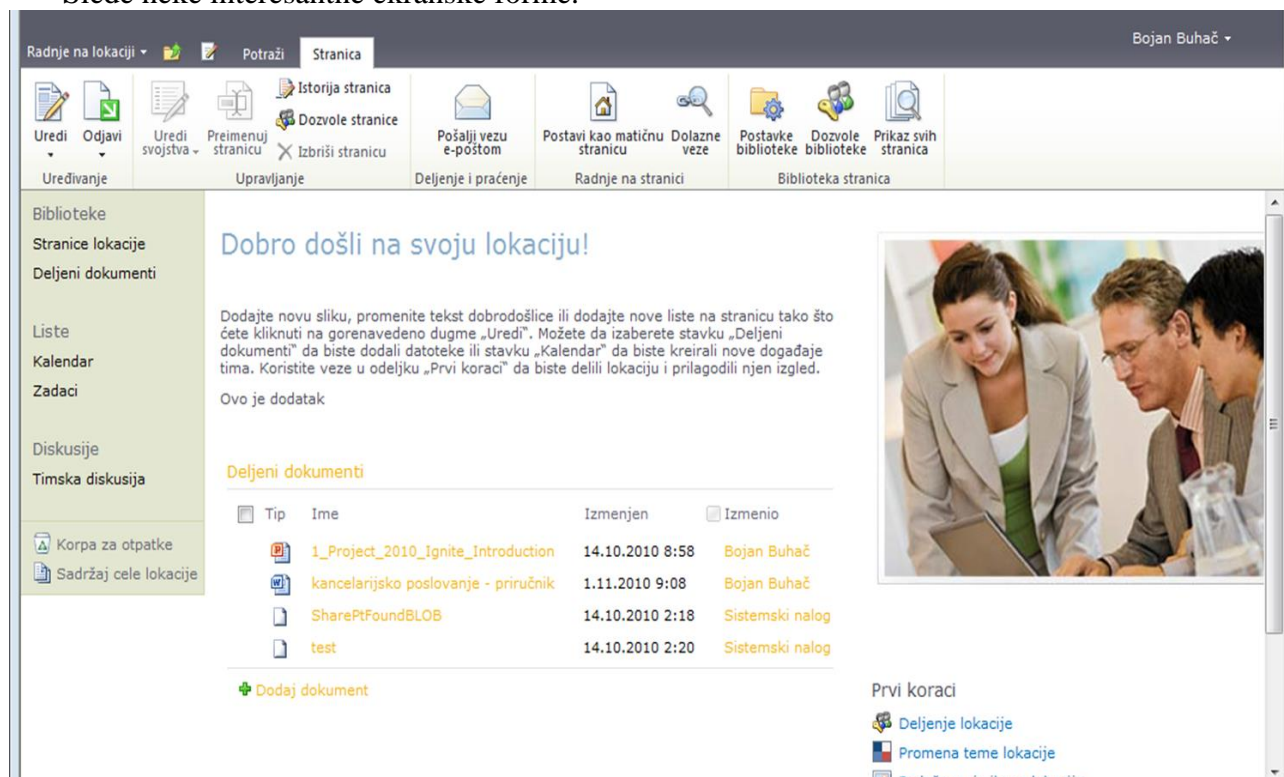
- ✓ *Obezbeđivanje jednostavnog, poznatog i doslednog korisničkog iskustva.* SharePoint je čvrsto integrisan sa poznatim klijentskim aplikacijama za stone računare, e-poštom i Web pregledačima da bi obezbedio dosledno korisničko iskustvo koje pojednostavljuje način na koji ljudi vrše interakciju sa sadržajem, procesima i poslovnim podacima.
- ✓ *Povećanje produktivnosti zaposlenih pojednostavlivanjem svakodnevnih poslovnih aktivnosti.* Automatizovane procesi omogućavaju pokretanje, praćenje i izveštavanje o uobičajenim poslovnim aktivnostima kao što su revizija i odobravanje dokumenata, pracenje izvršnih naredbi i prikupljanje potpisa.
- ✓ *Efikasno upravljanje sadržajem i promena njegove namene da bi se povećala poslovna vrednost.* Poslovni korisnici i autori sadržaja mogu da kreiraju i prosleduju sadržaj radi odobrenja i planiranog postavljanja na lokacije interne mreže ili Interneta. Upravljanje višejezičkim sadržajem pojednostavljeno je predlošcima

biblioteke dokumenata, koji su posebno dizajnirani da održe vezu između originalne verzije i različitih prevoda dokumenta

- ✓ *Pojednostavljeni pristup strukturiranim i nestrukturiranim informacijama u različitim sistemima širom organizacije.* Korisnici također mogu da kreiraju personalizovane prikaze i interakcije sa poslovnim sistemima preko pregledača, prevlačeći i otpuštajući unapred definisane pozadinske veze, koje mogu da se podešavaju. Poslovna skladišta dokumenata pomažu organizacijama da skladište i organizuju poslovne dokumente na jednoj centralnoj lokaciji.
- ✓ *Povezivanje ljudi sa informacijama i stručnim uputstvima.* SharePoint Enterprise Search uključuje poslovne podatke zajedno sa informacijama o dokumentima, ljudima i Web stranicama da bi se dobili sveobuhvatni, prikladni rezultati. Funkcije kao što su uklanjanje duplikata, ispravljanje pravopisa i upozorenja povećavaju prikladnost rezultata.
- ✓ *Ubrzani deljeni poslovni procesi izvan organizacije.* Omogućeno je korišćenje pametnih rešenja zasnovanih na elektronskim obrascima da bi se prikupile bitne informacije od klijenata, partnera i snabdevača preko Veb pregledača.
- ✓ *Deljenje poslovnih podataka bez otkrivanja osetljivih informacija.* Naprosto, postoji mogućnost separiranja i blokovskog deljenja znanja. Od šifre logovanog korisnika, definišu se stepeni prioriteta kao i nivoi mogućeg pristupa menadžera odnosno analitičara.
- ✓ *Omogućavanje donošenja odluka zasnovanih na informacijama predstavljanjem informacija koje su bitne za posao na jednoj centralnoj lokaciji.* SharePoint olakšava kreiranje direktnih, interaktivnih portala poslovnih informacija (BI) koji sakupljaju i prikazuju informacije bitne za posao iz različitih izvora, koristeći integrisane mogućnosti za poslovne informacije kao što su instrument table, Web komponente, kartice sa rezultatima, ključni indikatori uspešnosti (KPI), kao i tehnologije za povezivanje poslovnih podataka.

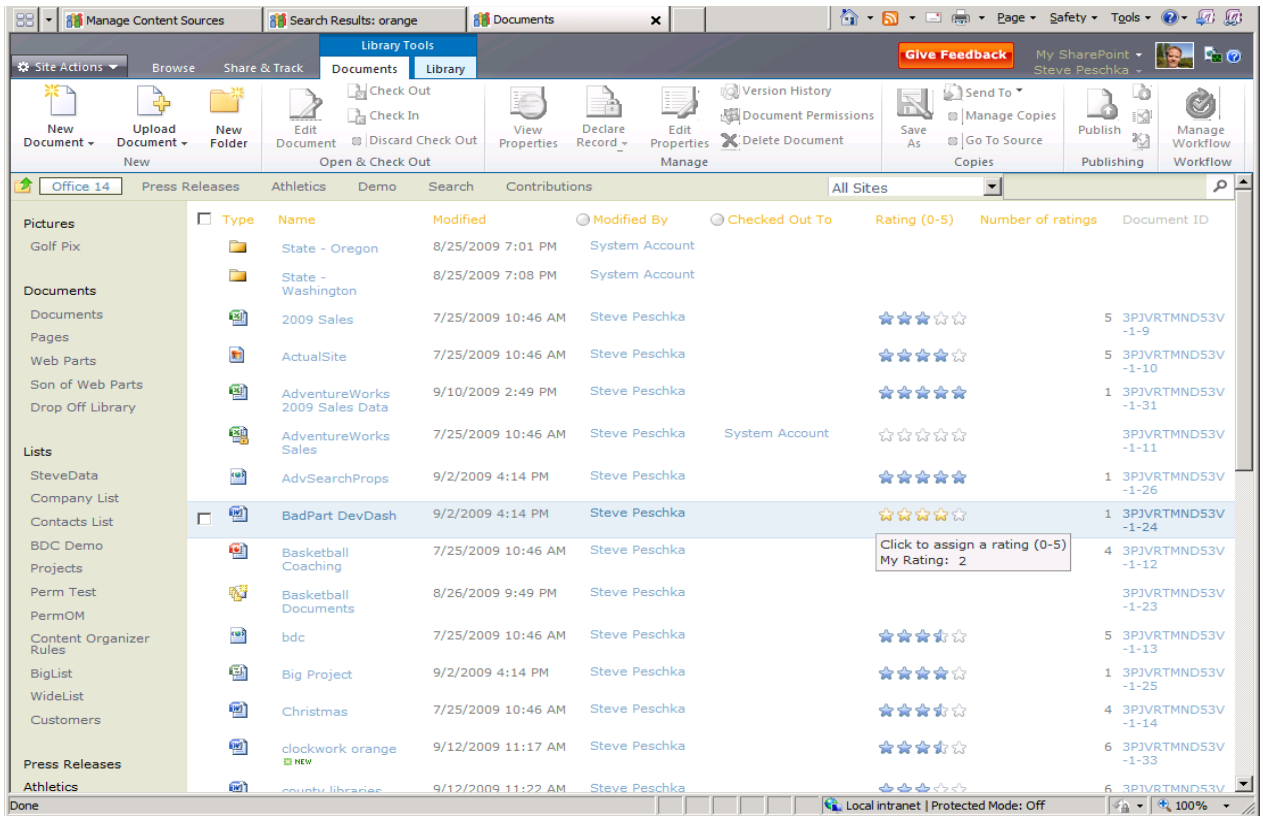
- ✓ *Obezbeđivanje jedinstvene, integrisane platforme za upravljanje aplikacijama na internoj mreži, eksternoj mreži i Internetu širom preduzeća. Server ima bogate, otvorene programske interfejsne aplikacija (API) i rukovaoce događajima za liste i dokumente.*

Slede neke interesantne ekranske forme.



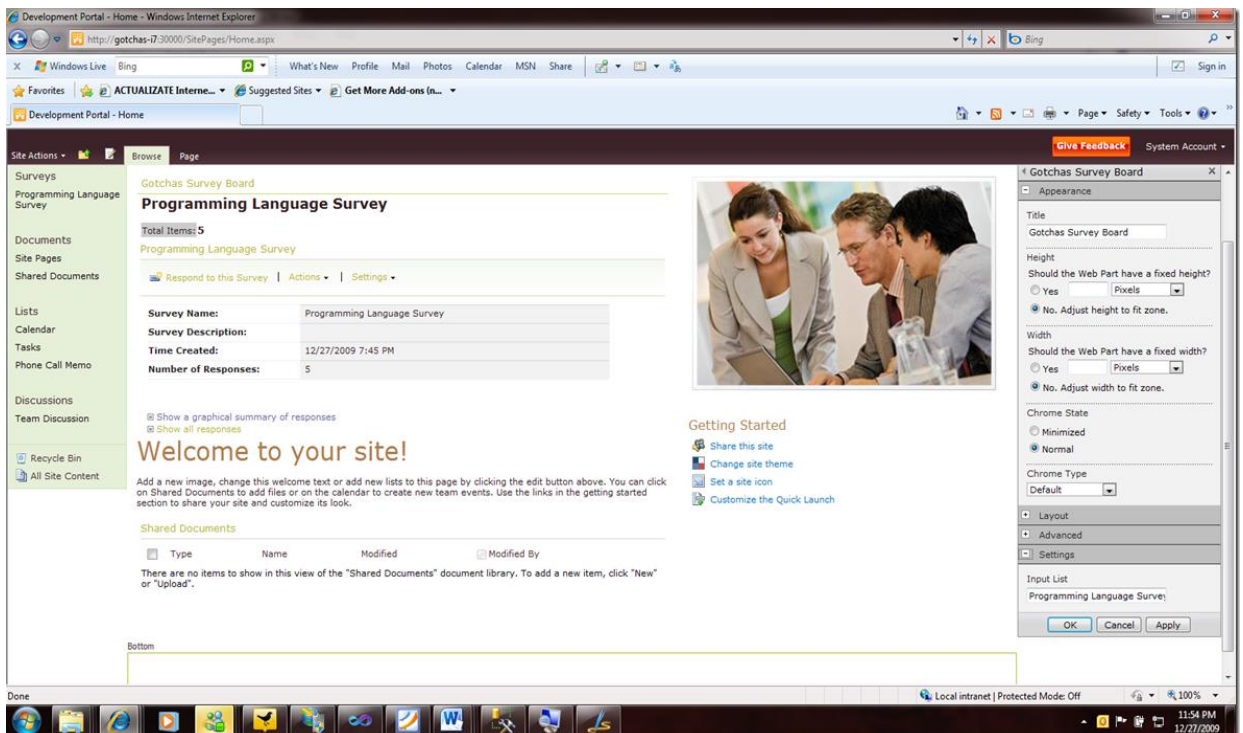
Slika 7.5. Izgled korisničkog panela Microsoft SharePoint 2010

Na prethodnoj slici vidi se veoma jednostavna i intuitivna forma rada sa prikazanim softverom.



Slika 7.6. Izgled panela za deljenje dokumenta u okviru Microsoft SharePoint 2010

Deljenje dokumenata prikazano je na prethodnoj slici, a na narednoj izgled kreiranja veb stranice.



Slika 7.7. Izgled panela za kreiranje veb stranica u okviru Microsoft SharePoint 2010

8.0. GRUPE I DRUŠTVENE MREŽE

Kao što je već istaknuto u prethodnim poglavljima, suština je da je *kvantum znanja grupe po definiciji uvek veći od znanja bilo kog pojedinca*. To je razlog da se u ovom poglavlju posebna pažnja posveti grupi ili u novije vreme sve češće nazvana kao *društvena mreža*. Prvo se ukazuje na formalizme predstavljanja društvene mreže, da bi se nakon tog opisa posebna pažnja posvetila aspektu *moći* člana tima u društvenoj mreži. Na samom početku ukazuje se na formalizme za predstavljanje društvenih mreža.

8.1. Formalizmi za predstavljanje društvenih mreža

Sama suština društvene mreže i njenog predstavljanja je vrlo jednostavna. Može da se definiše da je *društvena mreža* skup aktera (tačaka, čvorova, agenata) koji mogu biti međusobno povezani (veze, ivice, odnosi). Mreže mogu imati malo ili mnogo aktera i jednu ili više veza (relacija) među akterima u paru. Kako bi se razumela društvena mreža, za početak analize je potrebno u potpunosti opisati templejt društvenih odnosa. Odnosno, idealan slučaj je da je poznato sve o vezama koje postoje unutar svakog para aktera u populaciji, [11].

Jedan od razloga za korišćenje matematičkih i grafičkih tehnika za analizu društvene mreže je da bi se mreža opisala kompaktno i sistematično. Ovo takođe, omogućava korišćenje računara za brže i preciznije skladištenje i manipulaciju informacijama. Za male populacije aktera (npr. u kompaniji) se mogu opisati šabloni društvenih odnosa koji povezuju aktere prilično detaljno i efikasno koristeći reči. Kako bi se osigurao kompletan opis, moguće je napraviti listu svih logičkih parova aktera i opisati svaku moguću vezu za svaki par. Ovo može biti vrlo zamorno ukoliko postoji mnogo aktera i/ili vrsta veza. Formalno predstavljanje obezbeđuje sistematsko predstavljanje potrebnih informacija i pruža pravila za predstavljanje na načine koji su efikasniji od lista.

Drugi razlog za korišćenje (posebno matematičkih) formalnih metoda za predstavljanje društvenih mreža je taj da matematičko predstavljanje omogućava primenu računara za analizu mrežnih podataka. Na primer, postoje informacije o trgovinskoj razmeni 50 različitih vrsta robe između 170 zemalja sveta tokom jedne godine. Pvih 170 zemalja se mogu posmatrati kao akteri ili čvorovi, a količina svake vrste robe koju svaka zemlja izvozi u ostalih 169 zemalja može se smatrati kao jačina usmerene veze koja polazi od zemlje koja se posmatra do drugih zemalja. Analitičar koji proučava društvene mreže, može analizirati

da li su „strukture“ trgovine mineralima međusobno sličnije od struktura trgovine povrćem. Kako bi se dao odgovor na ovo pitanje, potrebno je obraditi ogromnu količinu podataka. Bilo bi potrebno nekoliko godina da se ovi podaci obrade ručno. Uz pomoć računara ovo se može obaviti za nekoliko minuta.

Treći, poslednji razlog za korišćenje *formalnih* metoda (matematičkih i grafičkih) za predstavljanje podataka društvenih mreža je taj da sama matematička pravila i tehnike grafova mogu same po sebi da daju odgovore na pitanja koja nas interesuju – stvari koje se ne bi videle ukoliko bi podaci bili predstavljeni rečima. Na primer, daje se opis grupe prijatelja sastavljene od četiri osobe: Marko, Nataša, Nikola, Ivana. Ovo se jednostavno može opisati rečima. Neka Marko voli Natašu i Nikolu ali ne i Anu;

Nataša voli Nikolu ali ne Marka ni Ivanu; Nikola voli svo troje a Ivana voli samo Nikolu.

Ovaj templejt veza „dopadanja“ se može opisati sa matricom akter – akter gde redovi predstavljaju izbore svakog od aktera. Ukoliko se akteri međusobno dopadaju, evidentira se „1“, a u drugom slučaju se postavlja se „0“. Takva matrica ima sledeći izgled:

	Marko	Nataša	Nikola	Ivana
Marko		1	1	0
Nataša	0		1	0
Nikola	1	1		1
Ivana	0	0	1	

Tabela 8.1. Matrica „dopadanja“

Mnoge stvari se odmah uočavaju u matrici kada su podaci predstavljeni na ovaj način, a koji možda ne bi bili tako uočljivi čitajući tekstualni opis pravila veza. Na primer, ukoliko pogledom pređemo preko svakog reda, možemo primetiti da se Nikoli dopada više osoba nego Marku, Ivani i Nataši i tu možda postoji neki šablon. Da li su muškarci skloniji tome da iskrenije govore o „dopadanju“ od žena? Korišćenjem predstavljanja putem matrice, takođe, odmah se postavlja pitanje lokacije na glavnoj dijagonali, budući da je ona prazna. Da li je ovo realno ili bi trebalo u istraživanje uključiti i vrstu veze „samodopadanja“? Ne postoje pravi odgovori na ova pitanja. Poenta je u tome da korišćenje matricnog predstavljanja može dovesti do lakšeg uočavanja određenih šablona i do postavljanja novih pitanja do kojih se ne bi došlo klasičnim tekstualnim opisom.

Predstavljanje mreža pomoću grafova

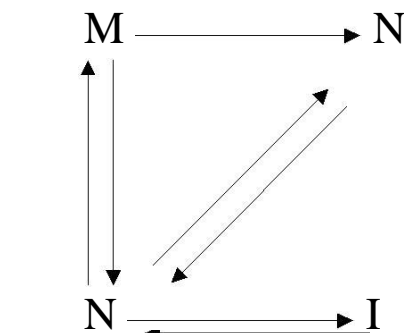
Analitičari društvenih mreža koriste dve vrste matematičkih alata za predstavljanje informacija o šablonima veza među akterima u društvu – *grafova* i *matrica*.

Postoji mnogo različitih vrsta grafova. Grafovi sa barovima (šipkama, pravougaonicima), isečcima u obliku pite, linijski grafovi trendova i mnogi drugi se jednim imenom nazivaju *grafovi* i/ili *grafici*. Analitičari mreža koriste najčešće jednu vrstu grafova koji se sastoji od tačaka (ili čvorova) kako bi se predstavili akteri, i linija (ili ivica) kako bi se predstavile veze, odnosi ili relacije. Pri korišćenju, sociolozi ovakve grafove nazivaju *sociogramima*. U matematici se koristi naziv *usmereni graf*, *graf sa znakovima* ili jednostavno *graf*. Postoje brojne varijacije na temu sociograma, ali svi koriste zajedničke odlike – svaki akter iz odabrane populacije je predstavljen krugom sa oznakom, a svaka veza koja postoji između dva aktera je predstavljena sa linijom. Pretpostavimo da nas zanima ko koga smatra prijateljem u jednoj grupi od četiri osobe (Marko, Nataša, Nikola, Ivana). Restavimo svakog aktera sa oznakom, nekad se čvorovi predstavljaju oznakama u krugovima, što je prikazano na sledećoj slici, videti u [115].



Slika 8.1. Akteri na grafu

Moderator prikuplja podatke tako što tražiti od svakog aktera (odvojeno) da se izjasni s kim je blizak prijatelj. Svaka od četiri osobe treba da odabere „bliskog prijatelja(e)“. U izmišljenom slučaju, Marko je odabrao Natašu i Nikolu ali ne i Ivanu, Nataša samo Nikola, Nikola je odabrao Marka, Natašu i Ivanu a Ivana samo Nikolu. Ovi podatci se mogu predstaviti strelicom, koja počinje od osobe koja je upravo intervjuisana a kraj bi joj bio osoba koju je ova prva odabrala.



Slika 8.2. Graf sa oznakama

Predstavljanje mreža pomoću matrica

Grafovi su veoma korisni za predstavljanje informacija o društvenim mrežama. Međutim, kada postoji mnogo aktera i/ili mnogo vrsta relacija, na njima je vrlo komplikovano uočiti pravila. Moguće je predstaviti podatke o društvenim mrežama u obliku matrica. Predstavljanje informacija na ovaj način takođe omogućava primenu matematičkih i računarskih alata za pronalaženje templejta. Analitičari društvenih mreža koriste matrice na razne načine, pa je potrebno razumeti neke osnovne stvari koje se tiču njih.

Matrica je u suštini samo skup podataka poređanih u obliku pravougaonika. Veličinu pravougaonika čine redovi i kolone elemenata. Matrica reda „m x n“ ima m redova i n kolona, pa tako matrica „3x6“ ima 3 reda i 6 kolona. Na sledećoj slici su prazne matrice reda 2x4 i 4x2.

1,1	1,2	1,3	1,4
2,1	2,2	2,3	2,4

Tabela 8.2. Prazna matrica reda 2x4

1,1	1,2
2,1	2,2
3,1	3,2
4,1	4,2

Tabela 8.3. Prazna matrica reda 4x2

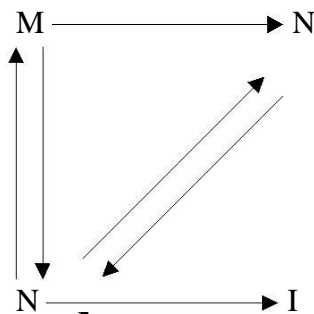
Elemente matrice čini njihova „adresa“. Element 1,1 je unos koji se nalazi u prvom redu u prvoj koloni, element 13,2 je u 13 redu, 2 kolone. Matrice su često predstavljene kao niz elemenata okruženih vertikalnim linijama ili ugaonim zagradama sa leve i desne strane. U HTML jeziku matrice se lakše predstavljaju tabelama. Sociolozi koji koriste matrice za društvene mreže jednostavno prikazuju podatke kao niz obeleženih redova i kolona. Oznake nisu deo matrice već služe za jasniju predstavu podataka. Sledeća matrica je tipa 4x4 sa dodatnim oznakama.

	Marko	Nataša	Nikola	Ivana
Marko		1	0	0
Nataša	1		1	0
Nikola	1	1		1
Ivana	0	0	1	

Tabela 8.4. Matrica reda 4x4

Najčešći oblik matrice u analizi društvenih mreža je onaj koji se sastoji od onoliko redova i

kolona koliko ima aktera u skupu podataka i gde elementi matrice predstavljaju veze među akterima. Najjednostavnija i najčešće korišćena matrica je binarna. Ukoliko veza postoji, unosi se „1“ u ćeliju, ukoliko ne postoji, unosi se „0“. Pva vrsta matrice je osnova za sve analize mreža i naziva se „matrica bliskosti“ jer govori o tome ko je pored koga ili ko je udaljen od koga u „društvenom prostoru“, koga čine relacije koje se mere. Po pravilu, na usmerenom grafu, osoba koja počinje vezu čini red, a ona na koju se cilja čini kolonu. Na primer, usmereni graf prijateljstva između Marka, Nataše, Nikole i Ivane na sledeći način:



Slika 8.3. Usmereni graf

Pošto se pretpostavi da su veze na nominalnom nivou (podaci su tipa binarnog izbora), mogu se predstaviti podatci pomoću matrice koja bi izgledala, kao na sledećoj tabeli.

	Marko	Nataša	Nikola	Miloš
Marko		1	1	0
Nataša	0		1	0
Nikola	1	1		1
Miloš	0	0	1	

Tabela 8.5. Matrica bliskosti

Pošto redovi predstavljaju izvor usmerenih veza, a kolone cilj veze, Marko ovde odabira Natašu, ali Nataša ne odabira Marka. Ovo je primer *asimetrične* matrice koja predstavlja usmerene veze (veze koje idu od pošiljalca ka primapcu), odnosno, element "m, n" ne mora da bude isti kao element "n, m". Ukoliko su veze predstavljene na matrici tipa „zdužene“ (na primer, veze koje predstavljaju relaciju „je poslovni partner od...“ ili „ko-pojavljivanja“ ili „ko-prisustvuje“ gde je odnos „u istom bordu direktora je sa ...“ matrica bi bila simetrična, odnosno element "m, n" bi bio isti kao element "n, m".

Podaci binarnih izbora su obično predstavljeni sa nulama i jedinicama koje ukazuju na prisustvo ili odsustvo, bilo koje logično moguće veze među parovima aktera. Znakovni grafovi se obično predstavljaju u matričnom obliku pomoću -1, 0 i +1 za označavanje negativnih, neutralnih i pozitivnih relacija. Kada se veze mere na rednom ili intervalnom

nivou, numerički iznos merenja veze se unosi kao element matrice. Mogući su i drugi oblici podataka - multi-kategorijski nominalni, redni sa više od 3 ranga, redni nominalni punog ranga, ali se oni retko koriste u sociološkim istraživanjima.

Za neke delove matrice se koristi određena terminologija. Ukoliko se posmatraju svi elementi reda (npr. koga je Marko odabrao kao prijatelja: 1,1,1,0) reč je o *vektoru reda*. Ukoliko se posmatraju svi oni koji su odabrali Marka kao prijatelja (prva kolona,1,0,1,0) reč je o *vektoru kolone*. Ponekad je korisno obavljati operacije na vektorima redova ili kolona. Na primer, ukoliko se sumiraju elementi vektora kolone u ovom primeru, dobila bi se mera „popularnosti“ svakog čvora u smislu koliko često su bili cilj usmerene prijateljske veze. U nastavku sledi opis nekih bitnih karakteristika društvenih mreža.

8.2. Važne osobine društvene mreže za poslovnu inteligenciju

U ovom delu disertacije, sledi opis nekih bitnih mera društvenih mreža za sistem poslovne inteligencije.

Distanca

Svojstva mreža koja su do sada bila razmatrana su uglavnom imala veze sa udaljenošću – direktnim vezama od jednog aktera ka drugom. Međutim, način na koji su ljudi *utkani* u mrežu je daleko složeniji. Dve osobe, A i B, imaju svaka po pet prijatelja. Pretpostavka je da ni jedan od prijatelja osobe A nema prijatelja osim njega, a svaki od prijatelja osobe B ima po 5 prijatelja. Podaci dostupni osobi B kao i potencijalni uticaj daleko su veći nego kod osobe A, odnosno, ponekad biti „prijatelj prijatelja“ može imati velike posledice. Kako bi se prikazao ovaj aspekt utkanosti pojedinaca, može se proučiti distanca (udaljenost) jednog aktera u odnosu na druge. Ukoliko su dva aktera udaljena, distanca među njima je jedan (signalu je potreban jedan korak kako bi stigao od izvora do cilja). Ukoliko osoba A kaže osobi B, a B kaže osobi V (ali A ne kaže osobi V), tada su akteri A i V na distanci veličine dva. Broj aktera koji se nalaze na distanci jedan od drugog, može biti važan za razumevanje razlika među akterima pri ograničenjima i mogućnostima koje poseduju, kao rezultat njihovog opažnja.

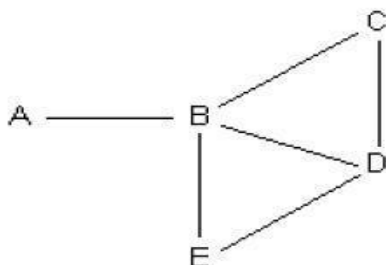
Ponekad može biti od značaja i podatak na koliko načina je moguće povezati dva aktera na određenoj distanci, odnosno, da li akter A može da dospe do aktera B na više od jednog načina. Višestruke veze mogu, jače od jedne veze ukazati na jaku vezu između dva aktera. Distance među akterima u mreži mogu biti važne makro-karakteristike mreže kao celine. Tako, gde su distance velike, potrebno je više vremena da se informacija proširi kroz populaciju. Može se desiti da neki akteri nisu svesni ili su pod uticajem drugih, čak i ako su

u principu dostupni, troškovi razmene mogu biti previše veliki. Variranje distance među akterima može biti podložno diferencijaciji i stratifikaciji. Oni akteri koji su bliži mnogim drugim akterima su u poziciji da iskoriste više moći od onih koji su udaljeni.

Hod

Postoji velika razlika između prostog i usmerenog grafa. Ako su A i B jedan do drugog na prostom grafu, oni imaju distancu veličine jedan. Na usmerenom grafu, A može biti odmah do B, ali B ne mora biti odmah do A. Distanca od A do B je jedan, ali ne postoji distanca od B do A. Usled ovih razlika, potrebno je opisati distance među akterima na grafu i digrafu.

Prosti graf. Najčešći oblik veze između dva aktera na grafu se naziva *hod*, videti u [11]. Hod je niz aktera i relacija koja počinje i završava se sa akterima. *Zatvoren hod* je onaj kod koga je početak i kraj hoda u jednom akteru. Hodovi nemaju ograničenja, oni mogu obuhvatiti jednog aktera ili jednu vezu nekoliko puta. Ciklus je posebno ograničen hod koji se često koristi u algoritmima koji proučavaju okruženja (tačke koje su odmah pored) aktera. *Ciklus* je zatvoren hod od 3 ili više aktera, gde su svi različiti, osim početnog i krajnjeg aktera. Dužina hoda predstavlja broj relacija koje ga čine, što je prikazano na sledećoj slici, videti [78]:



Slika 8.4. Graf sa hodovima

Na ovom grafu postoji mnogo hodova (čak beskonačno mnogo, ukoliko se u obzir uzmu hodovi bilo koje dužine). Kao primer, posmatrajmo samo neke, počevši od aktera A do aktera C. Postoji samo jedan hod dužine 2 (A,B,C), jedan hod dužine 3 (A,B,D,C), nekoliko hodova dužine 4 (A,B,E,D,C; A,B,D,B,C; A,B,E,B,C). Pošto su bez ograničenja, isti akteri i relacije se mogu koristiti više od jednog puta u jednom hodu. Ne postoje ciklusi koji počinju i koji se završavaju sa A. Postoje neki koji počinju i završavaju se sa B (B,D,C,B; B,E,D,B; B,C,D,E,B).

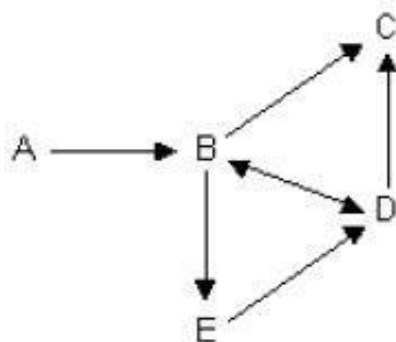
Pojam veze se može ograničiti na samo one hodove koji ne koriste relacije više puta. *Staza* između dva aktera je bilo koji hod koji koristi jednu relaciju ne više od jednog puta (međutim, isti akteri mogu biti deo jedne staze više puta). Dužina staze je broj relacija koje je čine. Sve staze su hodovi, ali nisu svi hodovi staze. Ako staza počinje i završava se u jednom akteru, naziva se *zatvorena staza*. U navedenom primeru, postoji više staza od A do

C. Putanje kao što je A,B,D,B,C se isključuju, jer su to hodovi, ali ne i staze jer je relacija BD korišćena više od jednog puta.

Najkorisnija definicija veze između dva aktera (ili aktera sa samim sobom) je *put*. Put je hod u kome se svaki akter i svaka relacija mogu koristiti samo jednom. Jedini izuzetak za ovo je *zatvoreni put*, koji počinje i završava se u istom akteru. Svi putevi su staze i hodovi, ali nisu sve staze i hodovi putevi. U navedenom primeru postoji ograničen broj puteva koji povezuju A i C: A,B,C; A,B,D,C; A,B,E,D,C.

Usmereni grafovi. Hodovi, staze i putevi se mogu definisati za usmerene grafove, ali postoje dve vrste, u zavisnosti od toga da li se u obzir uzima pravac.

Polu-hodovi, polu-staze i polu-putevi su isti kao i za podatke koji nisu usmereni. Pri definisanju ovih distanci, usmerenost veza se zanemaruje (lukovi – usmerene veze se tretiraju kao da su ivice – neusmerene veze). Dužina ovih distanci predstavlja broj relacija u hodu, stazi ili putu. Ukoliko se u obzir uzme usmerenost veza, mogu se definisati hodovi, staze i putevi na isti način, ali uz ograničenje da se ne može menjati smer dok se ide po relacijama od jednog aktera do drugog. Usmereni graf prikazan je na sledećoj slici, videti u [78]:



Slika 8.5. Usmereni novi graf

Na ovom di-grafu postoji nekoliko hodova od A do C. Međutim, ne postoji ni jedan hod od C do A. Neki hodovi od A do C su i staze (npr. A,B,E,D,B,C), ali postoje samo 3 puta od A do C. Jedan je dužine 2 (A,B,C), drugi dužine 3 (A,B,D,C) i jedan dužine 4 (A,B,E,D,C).

Razne vrste veza (hodovi, staze, putevi) omogućavaju mnogo različitih načina za posmatranje distanci među akterima. Glavni razlog zbog koga istraživači, koji proučavaju mreže, smatraju distance bitnim, jeste da one omogućavaju proučavanje jačine veza ili relacija. Akteri koji su povezani kratkim vezama ili distancama mogu imati jače veze, kao i akteri koji su povezani više puta. Njihova veza može biti manje podložna ometanju, pa je time stabilnija i pouzdanija.

Prečnik i geodezijska distanca

Jednu posebnu vrstu distance aktera unutar mreže pogotovo koriste algoritmi za određivanje složenijih osobina položaja pojedinaca i strukture mreže kao celine. Ona se naziva *geodezijska distanca*. Kako za usmerene, tako i za neusmerene podatke, geodezijska distanca je broj relacija na najkraćem hodu od jednog aktera do drugog.

Geodezijska distanca je u širokoj upotrebi u analizi mreža. Može postojati mnogo veza između dva aktera u mreži. Ukoliko se razmotri kako relacije među dva aktera mogu uticati i na šanse i na ograničenja, moguće je da ni jedna od ovih veza nije bitna. Na primer, ukoliko želimo da pošaljemo poruku akteru_1; pošto znamo adresu elektronske pošte, možemo je poslati direktno, što čini put dužine 1. Takođe, poznamo još jednog aktera_2 koji, zna adresu elektronske pošte prvog. Može se poslati poruka preko drugog aktera, što bi činilo put dužine 2. U ovom izboru, pre će se izabrati da se poruka pošaljeme direktno, jer je taj način brži i ne zavisi od posrednika. Geodezijska distanca je najčešće optimalna ili najefikasnija veza između dva aktera. Mnogi algoritmi u mrežnoj analizi pretpostavljaju da će akteri najpre da koriste geodezijske puteve kad god se pojavi izbor.

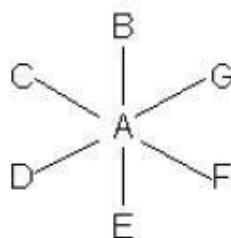
Centralnost i moć

Svi sociolozi bi se složili da je **moć** osnovno svojstvo društvenih struktura. Proučavanje mreža je u velikoj meri doprinelo shvatanju pojma društvene moći. Možda najbitnije, mrežni pristup naglašava da je moć nasledno relaciona. Pojedinaac poseduje moć pri dominaciji nad drugima, moć ega je zapravo zavisnost aktera, i obrnuto. Usled toga što je moć posledica šablona odnosa, količina moći u društvenim strukturama može da varira. Ukoliko je sistem labavo povezan (mala gustina), tu neće postojati mnogo moći; u gušćim sistemima postoji potencijalno više moći. Moć je i sistemsko (makro) i relaciono (mikro) svojstvo. Količina moći u sistemu i njena rasprostranjenost među akterima su povezani, ali ne predstavlja istu stvar. Dva sistema mogu imati istu količinu moći, ali ona može biti ravnomerno raspoređena u jednom, i neravnomerno u drugom. Moć u društvenim mrežama se može posmatrati ili kao mikro svojstvo (opisuje relacije među akterima) ili kao makro svojstvo (opisuje celu populaciju), kao i sa drugim sociološkim pojmovima, mikro i makro su blisko povezani u polju društvenih mreža.

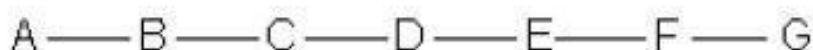
Analitičari mreža često opisuju način na koji je akter utkan u relacionu mrežu time što mu postavljaju ograničenja i pružaju mu šanse. Akteri sa manje ograničenja i koji imaju više šansi od drugih su na povoljnim strukturnim pozicijama. Ovo znači da akteri mogu dobiti povoljnije razmene i da će takav akter biti u centru pažnje za razliku od onih koji su na manje povoljnim pozicijama. Mrežna analiza je napravila bitan doprinos u definicijama i konkretnim merama za pristupe moći pozicija u strukturi društvenih odnosa.

Nekoliko aspekata moći

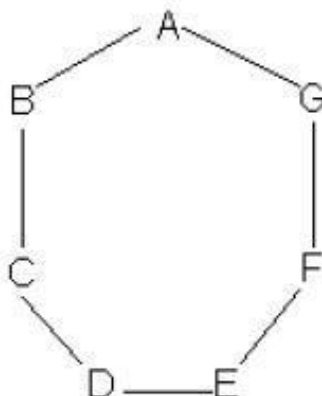
Za razumevanje pristupa koje koriste istraživači mreža pri proučavanju moći, posmatraćemo tri mreže, koje se mogu nazvati *zvezda*, *linija* i *krug*, videti [78]:



Slika 8.6. Mreža zvezda



Slika 8.7. Mreža linija



Slika 8.8. Mreža krug

Pri pogledu na mreže, vidi se da akter A ima jako povoljnu strukturnu poziciju u mreži zvezde, ukoliko mreža opisuje vezu tipa razmene resursa. Postavlja se pitanje zbog čega strukturna lokacija može biti povoljnija ili nepovoljnija za aktere. Sledi objašnjenje.

Stepen

Prvo, akter A ima više šansi i alternativa u odnosu na druge aktere. Ukoliko akter D odabere da uskrati resurs akteru A, akter A ima više alternativnih mesta gde bi mogao da ga pribavi. Međutim, ukoliko D odustane od razmene sa A, onda akter D neće imati sa kim da razmenjuje resurse. Što više veza akter poseduje, to više moći ima. U mreži zvezda, akter A ima stepen 6, ostali akteri imaju stepen 1. Ova logika nalaže da su mere centralnosti i moći zasnovane na *stepenu aktera*. Akteri koji imaju više veza imaju i više mogućnosti jer imaju više izbora. Ovo ih automatski čini manje zavisnim od bilo kog drugog aktera, pa time poseduju i više znanja, a samim tim i moći.

Posmatrajmo mrežu *krug* u smislu stepena. Svaki akter ima isti broj alternativnih aktera za razmenu (isti stepen), pa su sve pozicije podjednako povoljne ili nepovoljne. U mreži linija, stvari su složenije. Akteri na kraju linije (A i G) su zapravo u nepovoljnom položaju, ali su ostali praktično u istim pozicijama. U globalu, akteri koji su u centru strukture, u smislu da imaju veći stepen ili su bolje povezani, imaju povoljnije pozicije pa time poseduju više moći.

Blizina

Drugi razlog zbog koga akter A ima više moći od drugih u mreži zvezda je taj što je akter A *bliži* većem broju aktera od bilo kog drugog. Moć može proizlaziti direktnom razmenom, ali moć takođe dolazi i iz uloge *referentne tačke* po kojoj se drugi akteri porede i stajanjem u centru pažnje iz ugla većine aktera. Akteri koji mogu kraćim putanjama dospeti do drugih ili do kojih drugi mogu dospeti kraćim putanjama imaju bolje pozicije. Ova strukturna prednost se može prevesti u moć. U mreži zvezda, akter A je na geodezijskoj distanci veličine 1 od drugih aktera, a svi ostali na geodezijskoj distanci veličine 2 od drugih aktera. Ova logika strukturne prednosti nalaže da pristupi koji se zasnivaju na udaljenosti (blizini) i distanci zapravo predstavljaju izvor moći.

Posmatrajmo mrežu *krug* u smislu blizine aktera. Svaki akter se nalazi na različitoj dužini puta u odnosu na druge aktere, ali svi akteri imaju iste distribucije okoline i naizgled su jednaki u smislu njihovih strukturnih pozicija. U mreži linija, središnji akter D je bliži svim drugim akterima nego skup aktera C,E, skup B,F, i skup A,G. Još jednom su akteri na krajevima linije u nepovoljnim pozicijama.

Između (međučvor)

Treći razlog za povoljnu poziciju aktera A u mreži zvezda je taj što on leži *između* svih drugih parova aktera i ni jedan drugi akter se ne nalazi između A i drugih aktera, odnosno, on predstavlja međučvor. Ukoliko A želi da kontaktira F, on će to jednostavno učiniti. Ukoliko F želi da kontaktira B, on to mora da učini preko aktera A. Ovo daje akteru A priliku da se nađe u ulozi brokera (posrednika) prema svim drugim akterima – on može da naplaćuje svoje usluge, da izoluje aktere ili da spreči kontakte.

U mreži *krug*, svaki akter se nalazi između jednog para aktera. Postoje dva puta koja spajaju

svaki par aktera i svaki treći akter leži na jednom od njih ali ne i na nekom drugom putu. Opet, svi akteri su podjednako u povoljnom ili nepovoljnom položaju. U mreži linija, krajnje tačke A i G se ne nalaze između parova i time nemaju moć posrednika. Akteri koji su bliži sredini lanca leže na više puteva između parova i imaju povoljniji položaj.

Svaka od ove tri ideje ima svoju metodu. Geodezijski jedinični vektor se zasniva na principu bliskosti, pristup toka na ideji *između*, a Bonačićeva mera moći se koristi za princip stepena.

Bonačićev indeks moći

Poznati istraživač ove oblasti, Filip Bonačić, videti u [15] je predložio izmenu pristupa stepena centralnosti i ona je široko prihvaćena. Njegova ideja je vrlo jednostavna. Prvobitni pristup stepenu centralnosti tvrdi da akter koji ima više veza verovatno ima i više moći jer može direktno da utiče na ostale aktere. Ovo ima smisla, ali ako akteri imaju isti stepen, a to ne mora da znači da su akteri podjednako važni.

Pretpostavimo da dva učesnika, Marko i Nikola imaju po 5 bliskih prijatelja. Markovi prijatelji su, međutim, prilično izolovani članovi tima i nemaju mnogo drugih prijatelja osim njega. Svaki od Nikolinih prijatelja ima još mnogo prijatelja. Ko je onda više u centru? Prvo bi se pomislilo da je to Nikola, jer su članovi tima sa kojima je on povezan bolje povezani od Markovih direktnih saradnika. Bonačić zastupa ideju da je centralnost pojedinca izražena brojem njegovih veza i brojem veza koje imaju akteri u njegovom okruženju.

Dok je prvobitna teza bila da što su akteri više u centru, to poseduju više moći, Bonačić je doveo u pitanje ovu tezu. Ukoliko uporedimo Marka i Nikolu, videćemo da je Nikola više u centru, ali da li on zaista ima više moći? Jedan argument bi bio da je on verovatno uticajniji ako je više povezan, jer može brzo da dospe do drugih. Ali, ukoliko su akteri sa kojima je jedan povezan, dobro povezani sa ostalim, oni su prilično nezavisni od tog pojedinca. Sa druge strane, ukoliko članovi tima, sa kojima je pojedinac povezan, nisu dovoljno povezani, onda oni zavise od tog pojedinca. Bonačić zagovara teoriju da to što je pojedinac povezan sa drugima znači da je on više u centru, ali ne i da ima toliko jaku moć.

Bonačić je istakao da su i centralnost i moć u funkciji sa vezama koje akter ima u okruženju. Što akteri u okruženju imaju više veza, to pojedinac poseduje manje moći, i obrnuto. Bonačić je predložio da se svakom akteru odredi centralnost koja je jednaka njegovom stepenu i doda izmerena funkcija stepena aktera sa kojima je on povezan, a zatim se ceo postupak ponovi. Ponavljanjem postupka, relativne veličine rezultata svih aktera će biti iste.

Pojava ovakve „merljive“ društvene nauke vođene podacima je tekla vrlo polako, uglavnom

vođena nekolicinom računarskih naučnika, fizičara i sociologa. Kada bi pogledali vodeće stručne časopise iz ekonomije, sociologije i političkih nauka, našli bi malo dokaza o sve učestalijoj merljivoj društvenoj nauci koja je angažovana na prikupljanju i kvantitativom modeliranju ovih novih digitalnih tragova. Međutim, merljiva društvena nauka poprima velike razmere na mestima kao što je Gugl ili Jahu.

Nakon ukaza na važnost društvenih mreža u sistemima poslovne inteligencije i uloge moći kao člana tima u nastavku disertacije se ukazuje na mogućnosti ekstremizacije znanja.

9. MOGUĆNOSTI EKSTREMIZACIJE ZNANJA

Za donošenje različitih vrsta odluka, često je potreban različit pristup. Danas se koristi veliki broj metoda grupnog odlučivanja, kao što su Brainstorming, Delfi, Metoda nominalne grupne tehnike, itd., videti u [29]. Međutim, samo mali broj njih omogućava dobru kontrolu vremena, jer se akcenat stavlja na maksimizaciju kvaliteta odluke, dok se vremenski aspekt skoro u potpunosti zanemaruje.

Jedna od tradicionalno korišćenih metoda je i Delfi metoda, čiji model je više nego pogodan za odlučivanje u grupi, upravo jer daje dobru mogućnost kontrole vremena s obzirom na to da, svaki DO predviđanja vrši samostalno, što proces čini paralelnim, umesto sekvencijalnim. Samostalno predviđanje DO garantuje anonimnost i oslobađa ga potencijalnog pritiska grupe, čime podstiče objektivnost rezultata, detaljnije u [23]:

Ukratko, proces sprovođenja Delfi metode se može opisati u nekoliko sledećih koraka, videti [95]:

1. Definisavanje problema,
2. Odabir grupe DO,
3. Izrada i distribucija upitnika,
4. Popunjavanje upitnika od strane eksperata,
5. Statistička obrada popunjenih upitnika, i
6. Izrada konačnih rezultata ukoliko je postignut statistički značajno slaganje u predviđanju.

Kako je Delfi metoda iterativna (sprovodi se kroz nekoliko krugova, najčešće tri do četiri), koraci pod rednim brojem 4-6. se ponavljaju sve dok se ne postigne približan nivo saglasnosti među DO.

Ako se želi sistematizacija procesa u nekoliko logičkih celina, mogu se uočiti sledeće celine:

- ✓ Faza pripreme procesa (koraci 1,3)
- ✓ Faza odabira DO (korak 2)
- ✓ Faza predloga odluke (korak 4)
- ✓ Faza obrade rezultata (koraci 5,6)

Jasno je da je najkompleksnija faza donošenja odluke, i upravo u poboljšanju te faze se nalazi suština poboljšanja celog procesa. Jasno je da pripremnu i fazu obrade podataka IKT veoma pojednostavljuje i ubrzava, a da faze 1, 2 i 4, ne utiču na akumulaciju znanja, već samo na vremenski aspekt procesa i stoga nisu od suštinskog značaja za posmatrani problem.

Koristeći tehniku mrežnog planiranja, koja se zasniva na teoriji grafova mogu se modelirati struktura procesa i tok vremena u odlučivanju Delfi metodom, i to u jednoj njenoj iteraciji (Slika 9.1).

Formalno, graf G je određen parom $G = (N, L)$, pri čemu je $N = \{1, \dots, n\}$ i $L \subseteq \{(i, j) \mid i \in N, j \in N\}$. Elementi skupa N se zovu čvorovi ili temena, a elementi skupa L grane ili ivice, opisane u prethodnom poglavlju o društvenim mrežama. Grane grafa mogu biti usmerene ili neusmerene. Grani grafa se može pridružiti broj koji se naziva težina ili dužina grane, videti u [89]:

Put P između dva čvora, početnog s i krajnjeg n , je niz grana od kojih prva polazi iz početnog čvora, svaka sledeća polazi iz čvora u kome se prethodna završila, a poslednja se završava u krajnjem čvoru $P = ((s, i_1), (i_1, i_2), \dots, (i_{k-1}, i_k), (i_k, m))$. Alternativno, put se obeležava nizom čvorova $P = (s, i_1, i_2, \dots, i_{k-1}, i_k, m)$. Putu se pridružuje dužina koja je jednaka zbiru dužina grana koje ga sačinjavaju.

Najvažnija pravila konstruisanja mrežnog dijagrama, prema [89] su:

1. Aktivnost A predstavlja se usmerenom granom (i, j) pri čemu čvor i označava događaj početka aktivnosti, a čvor j događaj završetka aktivnosti A . Usmerenje grane je od početka ka završetku.
2. Mrežni plan mora imati samo jedan početni i samo jedan završni događaj.
3. Dve ili više aktivnosti ne smeju imati isti zajednički početni i završni događaj. Ovaj problem se lako rešava uvođenjem tzv. fiktivnih aktivnosti.
4. Radi očuvanja konzistentnosti pri crtanju mrežnog dijagrama, nekad se moraju uvesti dodatne grane koje se po konvenciji označavaju isprekidanom linijom. Ove aktivnosti se nazivaju *fiktivne* jer nemaju trajanje i ne troše resurse.

Primenom ovih pravila, dobija se sledeći mrežni dijagram (Slika 9.1.), gde su:

Čvorovi – događaji

P – priprema procesa,

T – testiranje DO,

S – početak donošenja odluka (događaj bez vremenskog trajanja),

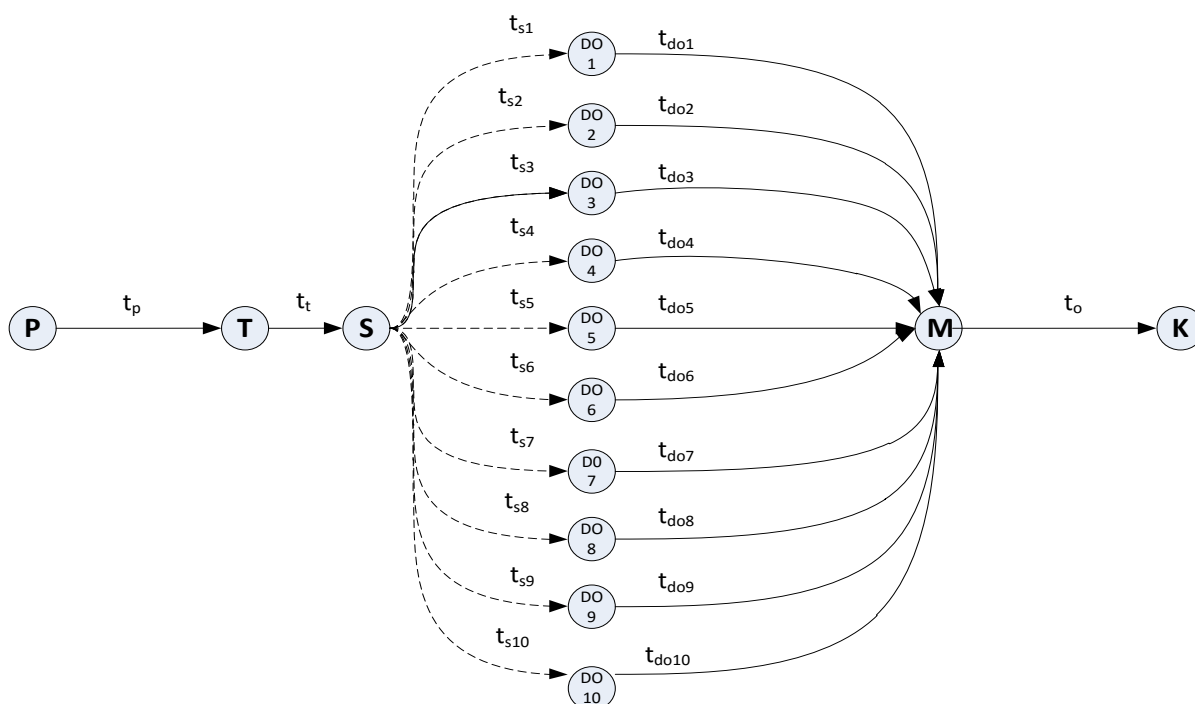
M – prikupljanje i obrada podataka od strane moderatora,

K – kraj, tj. objavljivanje rezultata,

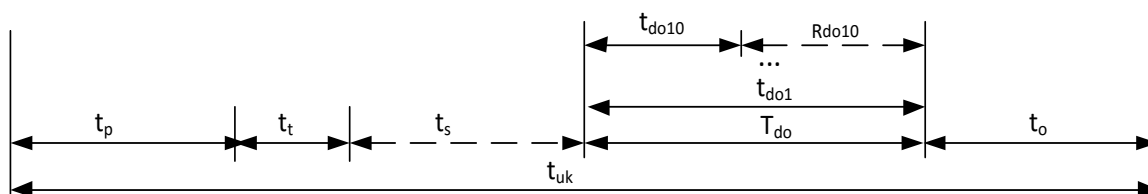
DO_i, $i = [1, 10]$ – Donošenje odluke od strane i -tog donosioca odluke.

Grane - Aktivnosti

- t_p – vreme strukturiranja problema i pripreme
- t_t – vreme testiranja
- t_s – vreme početka donošenja odluka (fiktivno vreme)
- t_{do}^i – vreme potrebno i-tom DO da donese odluku
- T_{do} – ukupno vreme faze „donošenje odluke“
- t_o – vreme obrade podataka
- t_{uk} – ukupno vreme trajanja procesa
- R_{do}^i – vremenska rezerva i-tog donosioca odluke



Slika 9.1. Prikaz strukture jednog ciklusa u Delfi metodi



Slika 9.1a. Prikaz vremenskog horizonta jednog ciklusa u Delfi metodi

Jasno je da događaj početak faze „priprema procesa“ (P), sa prethodne slike, ujedno i početni događaj celokupnog procesa odlučivanja. Završni događaj faze „pripreme procesa“ je ujedno i početni događaj faze „testiranje“ (T). Uvodi se 10 fiktivnih aktivnosti (jer postoji 10 donosilaca odluke) sa početnim događajem (S), koji je ujedno završni događaj „faze

testiranja“ (T), a sve u cilju nenarušavanja pravila br.3 za konstruisanje mrežnog dijagrama. Završni događaj za svaku od ovih fiktivnih aktivnosti predstavlja početni događaj „donošenje odluke“ svakog od donosilaca odluke (DO_i) gde je $i = [1,10]$. Završni događaj za svakog DO predstavlja početni događaj „prikupljanje i obrada podataka od strane moderatora“ (M), a završni događaj ove, krajnje, faze je „objavljivanje rezultata“, tj. „kraj“ (K).

Sa Slike 9.1., vidi se da ukupno vreme trajanja procesa (t_{uk}) čini zbir trajanja svake od faza u procesu, a da je trajanje faze „donošenje odluke“ (T_{do}) jednako „najsporijem“ DO, tj. onom DO čije je procenjeno vreme potrebno za donošenje odluke maksimalno u posmatranom skupu DO (na konkretnom primeru to je DO1 tj. $i = 1$). Samim tim, vremenska rezerva za svakog donosioca odluke R_{do}^i gde je $i = [1,10]$ predstavlja razliku ukupnog vremena za fazu donošenja odluke i vremena potrebnog i-tom DO da donese odluku tj. $R_{do}^i = T_{do} - t_{do}^i$ za $i = [1,10]$. Vreme (t_s) je fiktivno vreme, te je stoga isto za sve DO. Vreme pripreme procesa je t_p , vreme potrebno za testiranje DO - t_t , a vreme potrebno za obradu rezultata - t_o .

Iz prethodnog izvodimo formulu koja opisuje međusobne odnose vremena:

$$t_{uk} = t_p + t_t + t_s + T_{do} + t_o \quad (9.1.)$$

$$\text{gde je } T_{do} = \max(t_{do}^1, t_{do}^2, \dots, t_{do}^n) \quad (9.2.)$$

Kako su pretpostavke da se parametar T – potrebno vreme DO za donošenje odluke u datim uslovima može analitički odrediti, kao i parametar Z – količina znanja koju DO može da pruži u datim uslovima, to znači da se može odrediti i parametar E – efikasnost svakog DO koja se izražava kao odnos Z i T, i pokazuje količinu znanja koju svaki DO prenese u svakoj jedinici vremena:

$$E_i = \frac{Z_i}{T_i} \quad \frac{[\text{jedinica znanja}]}{[\text{vremenskih jedinica}]} \text{ gde je } i = [1,10] \quad (9.3.)$$

Studija slučaja za poboljšanje procesa odlučivanja

U ovom delu disertacije biće tretirana pristup kada su DO na istom hijerarhijskom nivou odlučivanja.

Kada je koeficijent efikasnosti - E poznat, cilj je napraviti takav redosled DO, da odnos ukupnog prosečnog znanja prenetog u svakoj jedinici vremena, na nivou celog procesa, bude maksimalno.

Ovaj problem sličan je problemu „trgovačkog putnika“ koji pripada klasi problema kombinatorne optimizacije na mrežama, a gde je cilj da se reši sledeći problem: Kojim redosledom trgovački putnik (u ovom slučaju informacija) treba da se kreće kroz mrežu gradova (u ovom slučaju donosilaca odluke), a da se pritom minimizuje ukupni pređeni put (u ovom slučaju vreme), i da pritom obiđe sve gradove (u ovom slučaju donosiocce odluka) samo jedanput.

Kako je problem „trgovačkog putnika“ kombinatorni problem, ukupan broj ruta (od kojih treba izabrati najkraću) iznosi $\frac{(n-1)!}{2}$ jer smer kretanja ne utiče na ukupnu dužinu rute, videti [130]. Za dimenziju $n = 10$ donosilaca odluke, bilo bi potrebno generisati 181 440 ruta, i od njih izabrati najbolju, što je resursno veoma zahtevno.

U opštem slučaju, problem kombinatorne optimizacije podrazumeva nalaženje minimuma funkcije f na skupu S , gde je S diskretan, konačan ili prebrojivo beskonačan skup. Dakle, model trgovačkog putnika ima sledeći oblik, [130]:

$$\min f(x) = \sum_{i=1}^m \sum_{j=1}^m d_{ij} X_{ij} \quad (9.4.)$$

pri ograničenjima:

$$\sum_{j=1}^m X_{ij} = 1 \quad i = 1, 2, \dots, m \quad (9.5.)$$

$$\sum_{i=1}^m X_{ij} = 1 \quad j = 1, 2, \dots, m \quad (9.6.)$$

$$\sum_{i \in S, j \in S} X_{ij} \leq |S| - 1 \quad S \subset \{1, 2, \dots, m\} \quad (9.7.)$$

gde je S bilo koji podskup gradova (donosilaca odluke), a X_{ij} binarna promenljiva definisana:

$$X_{ij} = \begin{cases} 1, & \text{ukoliko grana } (i, j) \text{ pripada optimalnoj ruti trgovačkog putnika} \\ 0, & \text{u suprotnom slučaju} \end{cases}$$

d_{ij} predstavlja rastojanje između dva grada (donosioca odluke, odnosno čvora na mrežnom dijagramu) i određuje se pomoću Euklidske metrike:

$$d(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (9.8.)$$

Pritom, prva dva ograničenja (9.5.) i (9.6.) vode računa o kretanju trgovačkog putnika (informacije) i ukazuju da iz svakog čvora koji poseti trgovački putnik (informacija), može da izlazi samo jedna grana. Ograničenje (9.7.) sprečava stvaranje ciklusa koji ne predstavljaju potpunu rutu trgovačkog putnika.

Iako navedeni egzaktni model sigurno vodi do optimalnog rešenja, njegova računska složenost je velika zbog čega troši mnogo resursa, stoga se preporučuje korišćenje tzv. „proždrljivih“ heuristika čija je računska složenost daleko manja, ali i rešenje koje se dobije ne mora nužno biti optimalno, ali će biti „dovoljno dobro“. Ističe se da za korišćenje heurističkih algoritama nije potrebno definisati egzaktni matematički model, što čini još jednu svojevrsnu uštedu resursa, detaljno u [130].

Jedna od heuristika pogodnih za rešavanje problema „trgovačkog putnika“ a koja koristi Euklidsku metriku, koja je izuzetno korisna za ovaj problem jer daje podjednaku važnost i vremenskoj dimenziji, i dimenziji znanja, zove se Algoritam „najbližeg suseda“ i glasi, [130]:

1. Odabrati početni čvor rute (u ovom slučaju će to biti čvor koji ima maksimalan parametar E, jer na taj način prednost za ulazak u rutu se daje donosiocima odluke koji su najefikasniji)
2. Pronaći čvor najbliži poslednjem čvoru uključenom u rutu. Ovaj najbliži čvor uključiti u rutu.
3. Ponavljati korak 2. sve dok se svi čvorovi ne uključe u rutu. Spojiti prvi i poslednji čvor rute.

Pretpostavimo da smo odredili sledeće ulazne parametre koji se mogu svesti na skalu od 0-100 za svaki parametar:

	DO1	DO2	DO3	DO4	DO5	DO6	DO7	DO8	DO9
Z – Znanje	55	47	26	87	22	41	78	54	13
T – vreme	45	10	25	10	85	96	35	14	35

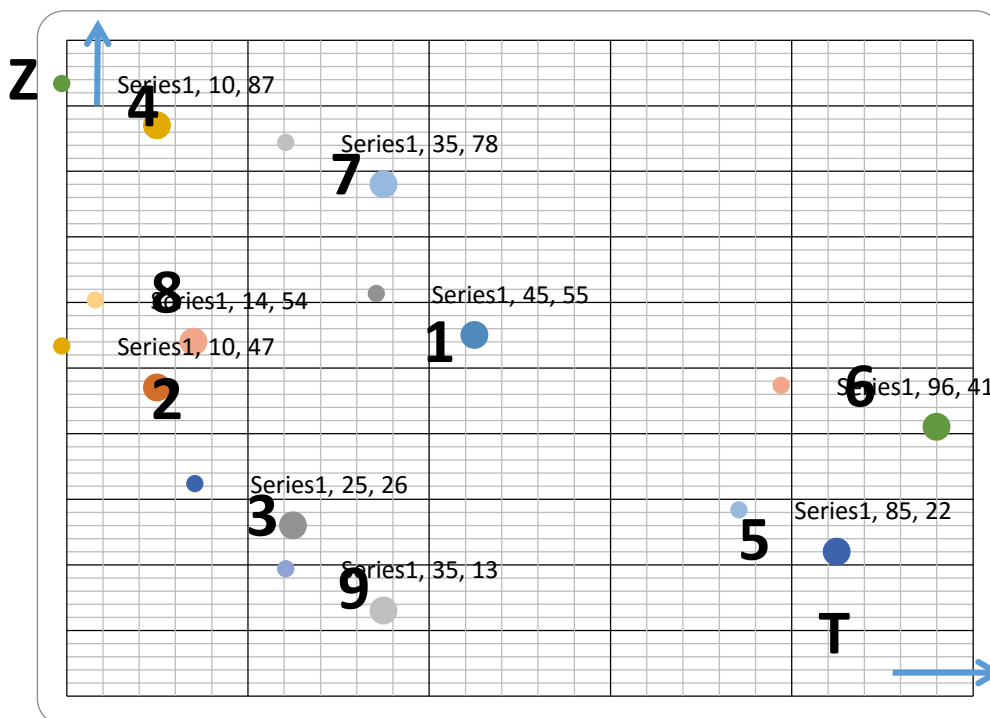
Tabela 9.1. Početne vrednosti procenjenog znanja i vremena svakog DO

Primenjujući relaciju (9.3.) za svakog DO, i zaokruživši brojeve na dve decimale, dobijamo sledeće vrednosti parametra efikasnosti – E:

E – efikasnost	1,22	4,70	1,04	8,70	0,26	0,43	2,23	3,86	0,37
-----------------------	------	------	------	-------------	------	------	------	------	------

Tabela 9.2. Izračunat koeficijent efikasnosti DO

Dakle, u startu imamo sledeći raspored čvorova (na x-osi se nalazi vreme, na y-osi znanje):



Slika 9.2. Početni raspored čvorova

Kako DO 4 ima najveći koeficijent efikasnosti, taj čvor se bira kao početni. Koristeći relaciju (9.8.), meri se udaljenost preostalih čvorova od čvora DO4 i bira minimalna:

$$D(4,1) = \sqrt{(10 - 45)^2 + (87 - 55)^2} = 47,42$$

$$D(4,2) = \sqrt{(10 - 10)^2 + (87 - 47)^2} = 40$$

$$D(4,3) = \sqrt{(10 - 25)^2 + (87 - 26)^2} = 62,82$$

$$D(4,5) = \sqrt{(10 - 85)^2 + (87 - 22)^2} = 99,25$$

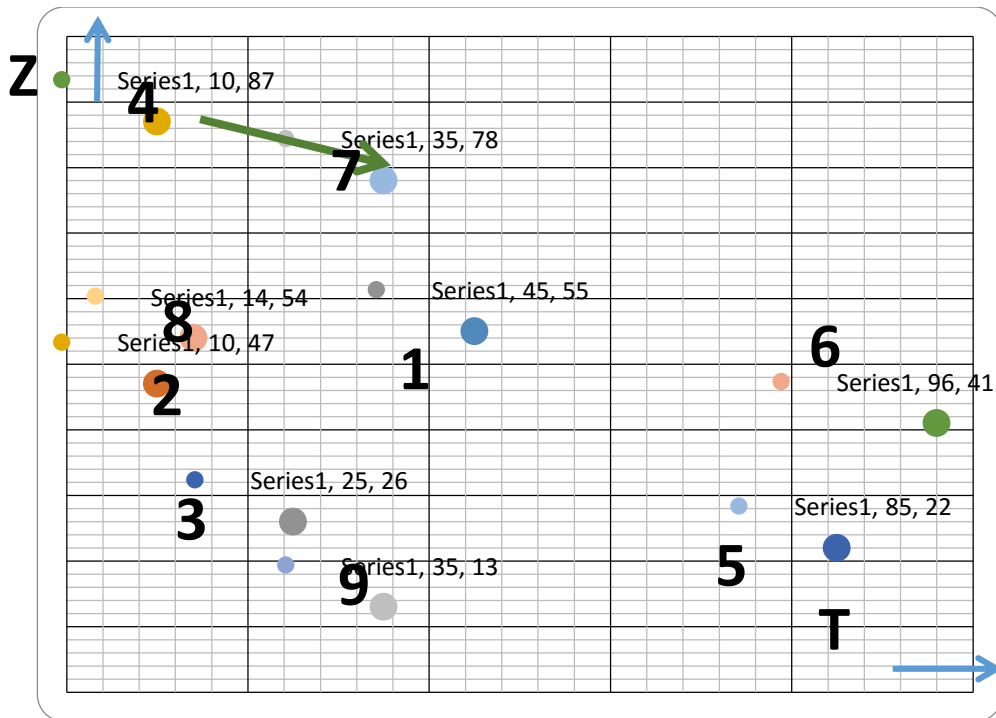
$$D(4,6) = \sqrt{(10 - 96)^2 + (87 - 41)^2} = 97,53$$

$$D(4,7) = \sqrt{(10 - 35)^2 + (87 - 78)^2} = 26,57 \rightarrow \text{minimalna vrednost}$$

$$D(4,8) = \sqrt{(10 - 14)^2 + (87 - 54)^2} = 33,24$$

$$D(4,9) = \sqrt{(10 - 35)^2 + (87 - 13)^2} = 78,11$$

S obzirom na to da je najmanja udaljenost $D(4,7)$ do čvora 7, sledeći čvor biće čvor DO7. Sada parcijalna ruta (DO4 - DO7) ima sledeći oblik:



Slika 9.3. Parcijalna ruta DO4-DO7

Pošto je parcijalna ruta došla do čvora DO7, meri se udaljenosti od čvora 7 do ostalih čvorova koji nisu uključeni u rutu, i bira minimalna:

$$D(7,1) = \sqrt{(35 - 45)^2 + (78 - 55)^2} = 25,08$$

$$D(7,2) = \sqrt{(35 - 10)^2 + (78 - 47)^2} = 39,82$$

$$D(7,3) = \sqrt{(35 - 25)^2 + (78 - 26)^2} = 52,95$$

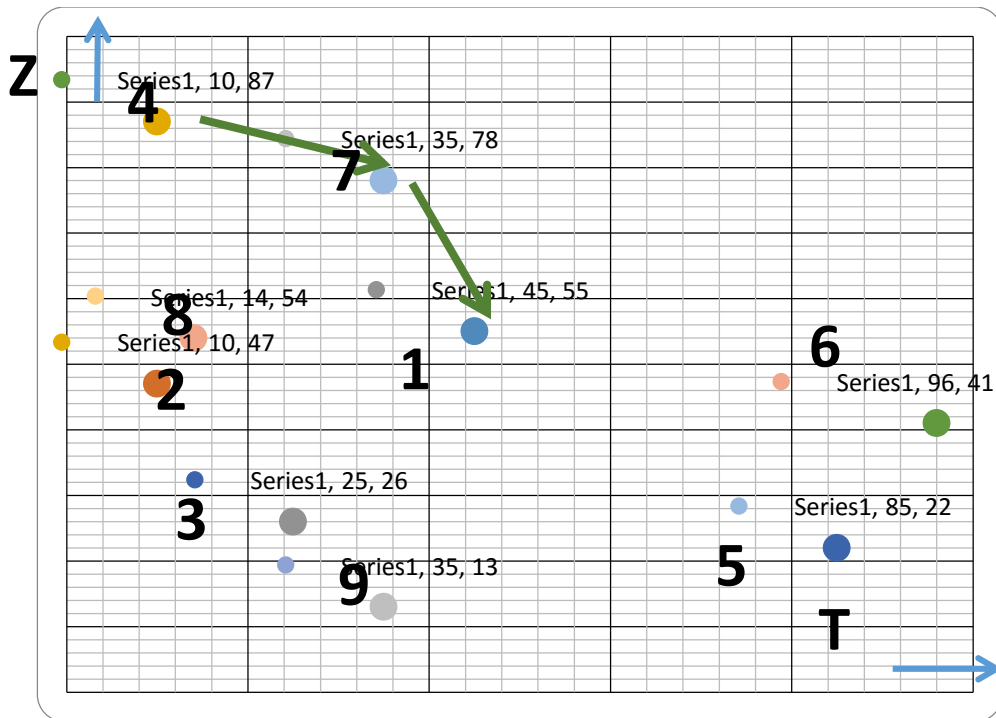
$$D(7,5) = \sqrt{(35 - 85)^2 + (78 - 22)^2} = 75,07$$

$$D(7,6) = \sqrt{(35 - 96)^2 + (78 - 41)^2} = 71,34$$

$$D(7,8) = \sqrt{(35 - 14)^2 + (78 - 54)^2} = 31,89$$

$$D(7,9) = \sqrt{(35 - 35)^2 + (78 - 13)^2} = 65$$

S obzirom na to da je najmanja udaljenost $D(7,1)$ do čvora 1, sledeći čvor biće čvor DO1. Sada parcijalna ruta (DO4 – DO7 – DO1) ima sledeći oblik:



Slika 9.4. Parcijalna ruta DO4-DO7-DO1

Pošto je parcijalna ruta došla do čvora DO1, meri se udaljenosti od čvora 1 do ostalih čvorova koji nisu uključeni u rutu, i bira minimalna:

$$D(1,2) = \sqrt{(45 - 10)^2 + (55 - 47)^2} = 35,9$$

$$D(1,3) = \sqrt{(45 - 25)^2 + (55 - 26)^2} = 35,23$$

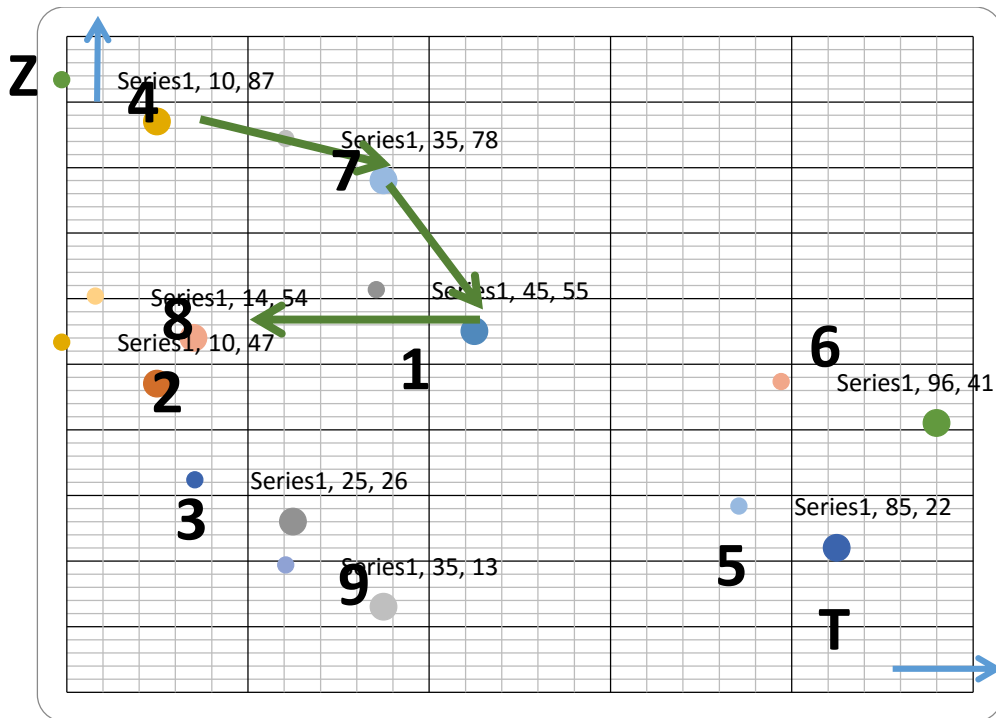
$$D(1,5) = \sqrt{(45 - 85)^2 + (55 - 22)^2} = 51,85$$

$$D(1,6) = \sqrt{(45 - 96)^2 + (55 - 41)^2} = 52,89$$

$$\mathbf{D(1,8) = \sqrt{(45 - 14)^2 + (55 - 54)^2} = 31,02}$$

$$D(1,9) = \sqrt{(45 - 35)^2 + (55 - 13)^2} = 43,17$$

S obzirom na to da je najmanja udaljenost $D(1,8)$ do čvora 8, sledeći čvor biće čvor DO8. Sada parcijalna ruta (DO4 – DO7 – DO1 – DO8) ima sledeći oblik:



Slika 9.5. Parcijalna ruta DO4-DO7-DO1-DO8

Pošto je parcijalna ruta stigla do čvora DO8, meri se udaljenosti od čvora 8 do ostalih čvorova koji nisu uključeni u rutu, i bira minimalna:

$$D(8,2) = \sqrt{(14 - 10)^2 + (54 - 47)^2} = 8,06$$

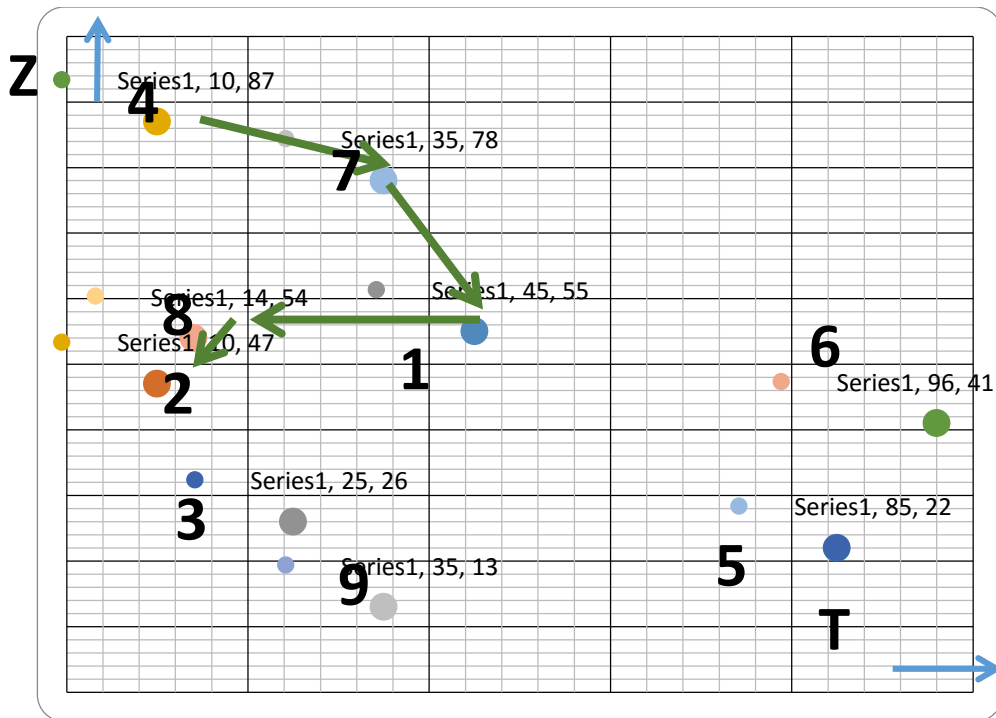
$$D(8,3) = \sqrt{(14 - 25)^2 + (54 - 26)^2} = 30,08$$

$$D(8,5) = \sqrt{(14 - 85)^2 + (54 - 22)^2} = 77,88$$

$$D(8,6) = \sqrt{(14 - 96)^2 + (54 - 41)^2} = 83,02$$

$$D(8,9) = \sqrt{(14 - 35)^2 + (54 - 13)^2} = 46,06$$

S obzirom na to da je najmanja udaljenost $D(8,2)$ do čvora 2, sledeći čvor biće čvor DO2. Sada parcijalna ruta (DO4 – DO7 – DO1 – DO8 – DO2) ima sledeći oblik:



Slika 9.6. Parcijalna ruta DO4-DO7-DO1-DO8-DO2

Pošto je parcijalna ruta stigla do čvora DO2, meri se udaljenosti od čvora 2 do ostalih čvorova koji nisu uključeni u rutu, i bira minimalna:

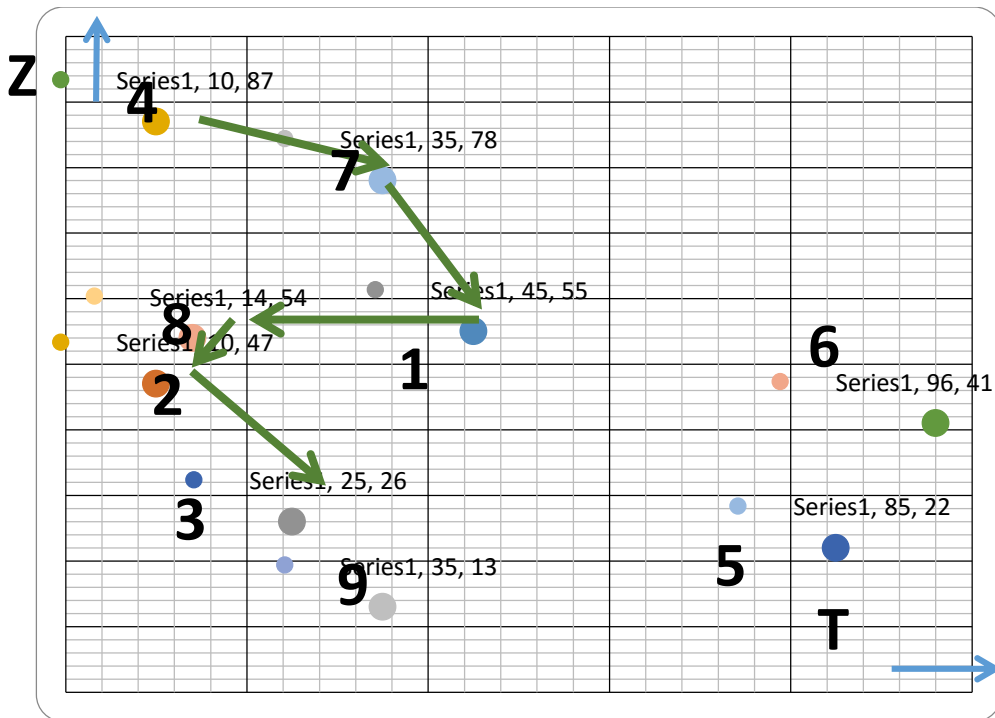
$$D(2,3) = \sqrt{(10 - 25)^2 + (47 - 26)^2} = 25,81$$

$$D(2,5) = \sqrt{(10 - 85)^2 + (47 - 22)^2} = 79,06$$

$$D(2,6) = \sqrt{(10 - 96)^2 + (47 - 41)^2} = 86,21$$

$$D(2,9) = \sqrt{(10 - 35)^2 + (47 - 13)^2} = 42,2$$

S obzirom na to da je najmanja udaljenost $D(2,3)$ do čvora 3, sledeći čvor biće čvor DO3. Sada parcijalna ruta (DO4 – DO7 – DO1 – DO8 – DO2 – DO3) ima sledeći oblik:



Slika 9.7. Parcijalna ruta DO4-DO7-DO1-DO8-DO2-DO3

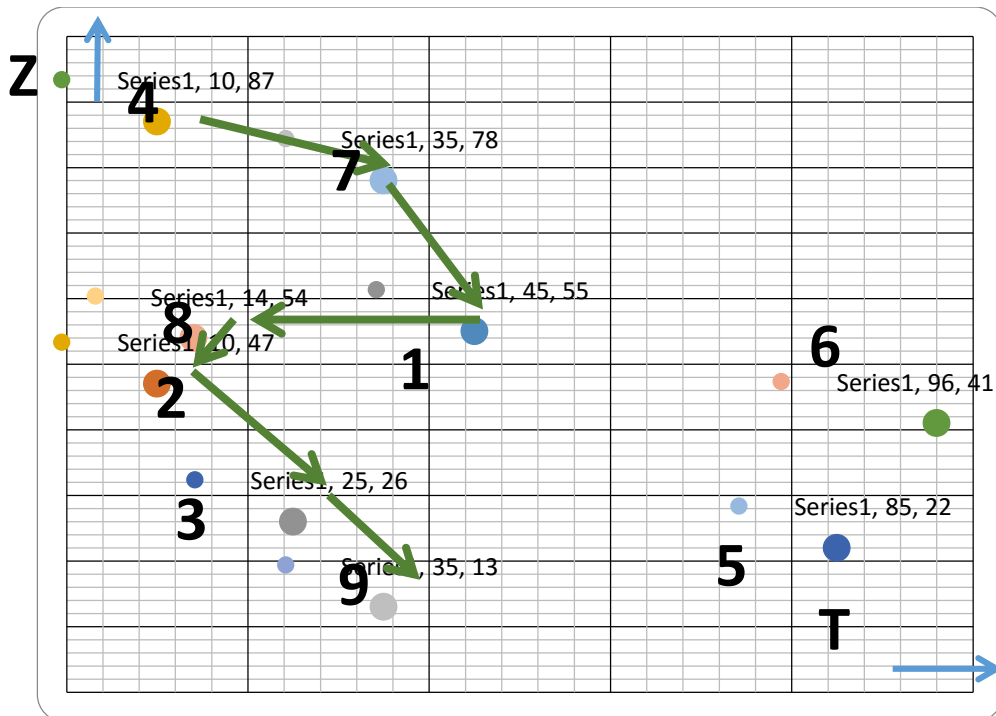
Pošto je parcijalna ruta stigla do čvora DO3, meri se udaljenosti od čvora 3 do ostalih čvorova koji nisu uključeni u rutu, i bira minimalna:

$$D(3,5) = \sqrt{(25 - 85)^2 + (26 - 22)^2} = 60,13$$

$$D(3,6) = \sqrt{(25 - 96)^2 + (26 - 41)^2} = 72,57$$

$$D(3,9) = \sqrt{(25 - 35)^2 + (26 - 13)^2} = 16,4$$

S obzirom na to da je najmanja udaljenost $D(3,9)$ do čvora 9, sledeći čvor biće čvor DO9. Sada parcijalna ruta (DO4 – DO7 – DO1 – DO8 – DO2 – DO3 – DO9) ima sledeći oblik:



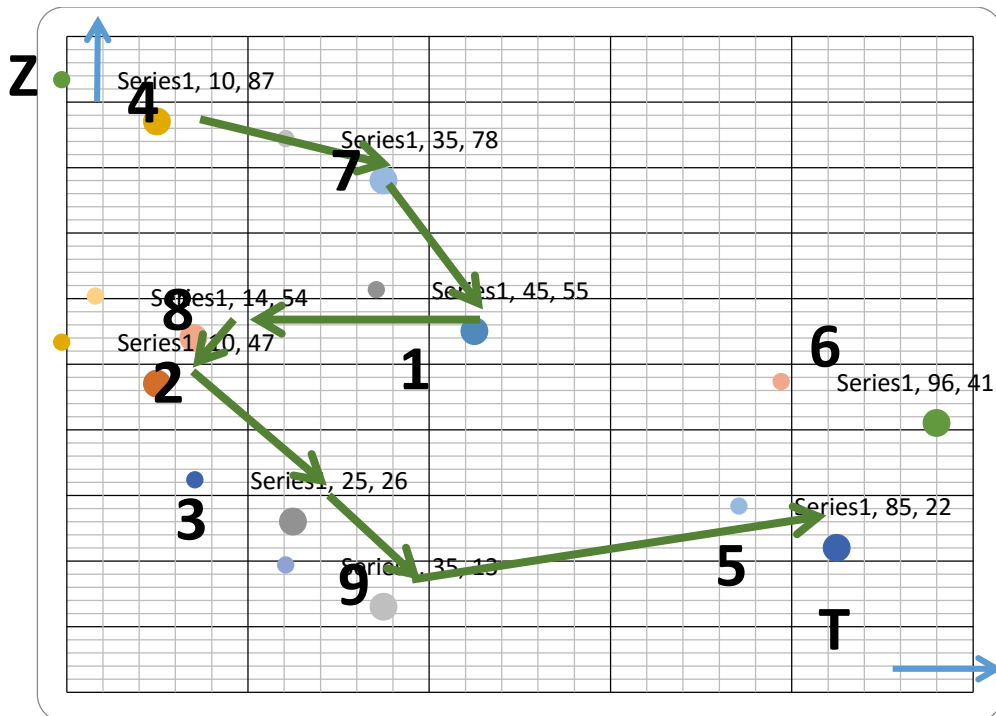
Slika 9.8. Parcijalna ruta DO4-DO7-DO1-DO8-DO2-DO3-DO9

Pošto je parcijalna ruta došla do čvora DO9, meri se udaljenosti od čvora 9 do ostalih čvorova koji nisu uključeni u rutu i bira se najmanja:

$$D(9,5) = \sqrt{(35 - 85)^2 + (13 - 22)^2} = 50,8$$

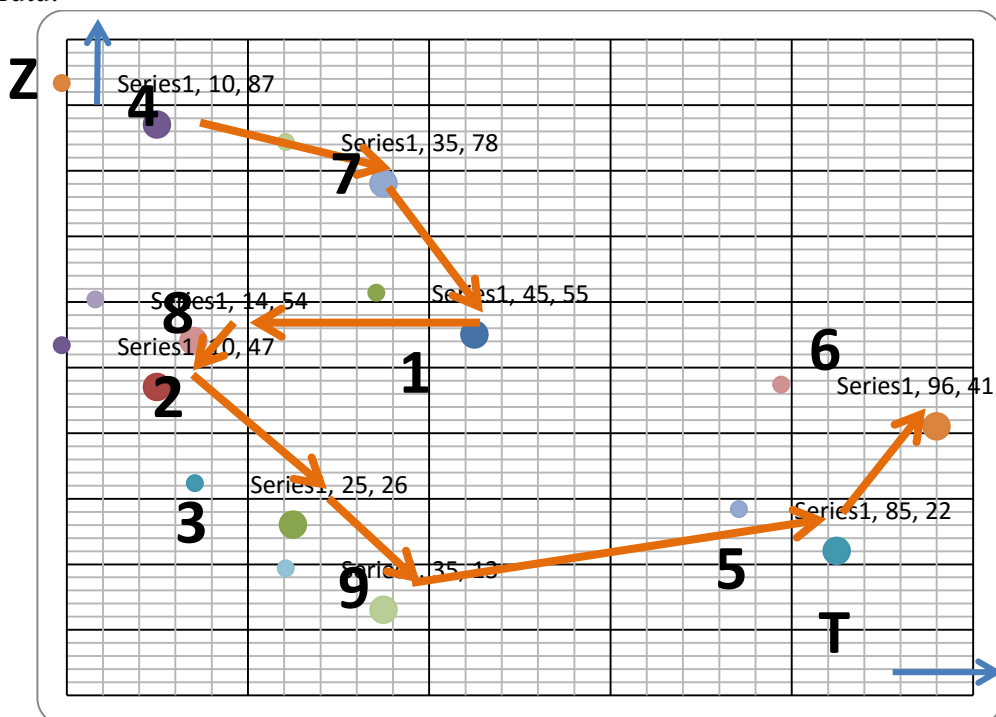
$$D(9,6) = \sqrt{(35 - 96)^2 + (13 - 41)^2} = 67,12$$

S obzirom na to da je najmanja udaljenost $D(9,5)$ do čvora 5, sledeći čvor biće čvor DO5. Sada parcijalna ruta (DO4 – DO7 – DO1 – DO8 – DO2 – DO3 – DO9 – DO5) ima sledeći oblik:



Slika 9.9. Parcijalna ruta DO4-DO7-DO1-DO8-DO2-DO3-DO9-DO5

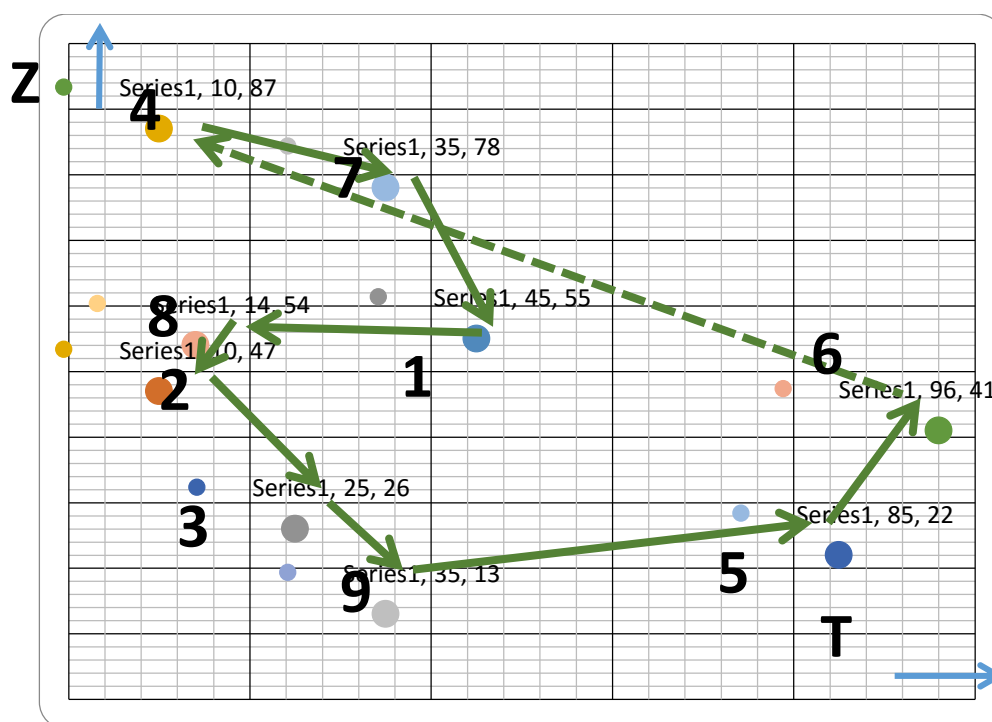
Pošto je parcijalna ruta stigla do čvora DO5, a jedini preostao čvor je čvor 6, uključujemo ga u rutu:



Slika 9.10. Parcijalna ruta DO4-DO7-DO1-DO8-DO2-DO3-DO9-DO5-DO6

Zatvaramo krug i zaustavljamo algoritam. Dobijamo kompletnu rutu: DO4 – DO7 – DO1 – DO8 – DO2 – DO3 – DO9 – DO5 – DO6 – DO4, koja pokazuje kojim redosledom treba

konsultovati donosioce odluka kako bi se ostvario maksimalan protok znanja za minimalno vreme. Početni i krajnji čvor rute (DO4) predstavlja najboljeg moderatora za dati slučaj.



Slika 9.11. Kompletna ruta DO4-DO7-DO1-DO8-DO2-DO3-DO9-DO5-DO6-DO4

Neke od prednosti ovog algoritma:

- jednostavnost,
- laka implementacija u IKT,
- metrika koja daje podjednaku značajnost brzini i znanju.

Nedostaci algoritma:

- za veći broj donosilaca odluke, računaska složenost geometrijski raste.

Na ovaj način prikazan je put saradnje među članovima tima, tokom prenosa i deljenja znanja, radi predloga odluke kao najprihvatljivijeg rešenja. Upravo sa aspekta, kako se procesom grupnog odlučivanja može unaprediti sistem poslovne inteligencije.

U nastavku disertacije slede zaključna razmatranja, kao i korišćena literatura.

10. ZAKLJUČAK

Uvođenje sistema poslovne inteligencije je projekat, za koji se može reći da nema kraja. Kako konkurencija postaje agresivnija, okolina nestabilnija i budućnost neizvesnija, zahtevi pred sistemima analize i prognoze postaju sve složeniji. Postavlja se pitanje cene ovakvih sistema i troškova informacija. U ovom veku, eri informacija, cena znanja jednaka je ceni opstanka na tržištu. Tome ide na ruku činjenica, da je zbir znanja članova tima, egzaktno uvek veći od znanja bilo kog pojedinca, što je ujedno i glavni predmet istraživanja same doktorske disertacije.

U uslovima ogromnih promena na tržištu i velikih izazova s kojima se suočava globalna ekonomija, kompanije treba da fokusiraju svoje investicije u projekte koji ispunjavaju poslovne zahteve, koji nastaju u uslovima velikih pritisaka, odnosno na one projekte koji korisnicima brzo obezbeđuju povratnu vrednost. To prouzrokuje uvođenje sistema za upravljanje poslovnim informacijama, jer je očigledno isplativa investicija, zato što knjigovodstvo ne beleži oportunitetne troškove loših poslovnih odluka na osnovu nedostajućih informacija. Takve promašaje beleže tržište i konkurencija. Ta dva *merna instrumenta* nepogrešivo mogu kazniti slabosti poslovanja.

S obzirom da su u globalnoj ekonomiji kreditne linije zategnute i da se potrošnja usporava, kompanijama je potrebno da identifikuju poslovne šanse i modernizuju aktivnosti s ciljem ne samo da održe na minimumu poverenje klijenata, već i da ga podignu na viši nivo. Bolji, pravovremeni uvid u poslovne operacije može da pomogne da se identifikuju segmenti poslova kojima su potrebna unapređenja, a čiji je rezultat smanjenje troškova. Identifikovanjem mogućnosti za poboljšanje uvida u tokove novca i likvidnost, pravljenjem efikasnog rasporeda radne snage, još bržim pristupom ključnim podacima o poslovanju preduzeća, modernizacijom lanca nabavke, brzim uštedama u nabavci i smanjenjem troškova, dugoročno posmatrano može se napraviti ključna prednost u performansama preduzeća. Sve ovo samo ako se poseduje resurs *znanje*.

Jer, prema [127], znanje je uvek vezano za kontekst, a to znači da znanje koje važi u određenom kontekstu, ne mora da važi u drugom kontekstu. Kada se javi određeni problem koji treba da se reši, bitan je kontekst u kome se javlja, jer ono zajedno sa problemom određuje rešenje koje se traži. *Znanje predstavlja trojku problema, konteksta i rešenja*. Zapravo, znanje je model koji predlaže rešenje određenog problema u određenom kontekstu.

U realnim primenama znanja u sistemima poslovne inteligencije, dodaje se dodatni zahtev da znanje treba da prođe iskustvenu proveru, tj. potrebno je vremena dok se ne prihvati da je određeno rešenje znanje, tj. da može da se koristi za donošenje upravljačkih odluka.

Za znanje važi da:

- ✓ Posedovanje znanja je preduslov donošenja ispravnih upravljačkih poslovnih odluka;
- ✓ Znanje u sebi sadrži komponentu akcije (rešenja);
- ✓ Znanje je vezano za kontekst;
- ✓ Znanje može da se predstavi trojkom: problem, kontekst i rešenje;
- ✓ Znanje predstavlja prihvatljivo rešenje određenog problema, koje najčešće nije optimalno, jer u realnim problemima višekriterijumskog odlučivanja nema savršenog rešenja, i
- ✓ Svako znanje traži određenu empirijsku potvrdu.

Tako se procesi znanja prvenstveno analiziraju pomoću tradicionalnih poslovnih metoda za modelovanje procesa. Potrebe i zahtevi za znanjem se modeluju zajedno sa intenzivnim procesima znanja.

Ovakvom kombinacijom i komparativnom analizom, koriste se prednosti svakog od metoda a time se poboljšava kompletan menadžment znanja.

Disertacija je realizovana po sledećim poglavljima.

Nakon *uvoda*, u kome se daje kratak prikaz sveobuhvatnog poduhvata i svih izazova istraživanja, sledi *drugo poglavlje* u kome se definiše *predmet i cilj istraživanja*. U ovom poglavlju pored predmeta i cilja istraživanja, daje se pregled dosadašnjih rezultata u oblasti istraživanja, zatim polazne hipoteze, metode iztraživanja, kao i očekivani naučni i stručni doprinosi.

Treće poglavlje opisuje fenomen odlučivanja, kroz određivanje mesta odlučivanja u savremenom menadžmentu, njegovog interdisciplinarnog okvira, zatim vreste odlučivanja i prelaza u poslovnu inteligenciju. Posebno se ukazuje na objedinjavanje fenomena odlučivanja u savremenu naučnu disciplinu pod nazivom *teorija odlučivanja*. Dat je kratak istorijski pregled fenomena odlučivanja, kao i modela odlučivanja. Istaknuta je jaka veza trojnosti kvalitativnog, kvantitativnog i informatičkog pristupa u odlučivanju. Pored toga u okviru ovog poglavlja, imajući u vidu naslov teze, posebno se tretira grupno odlučivanje i njegovi najvažniji elementi, kao što su *grupa, proces grupnog odlučivanja* i rezultat, odnosno *grupna odluka*.

Sledeće, *četvoro poglavlje* tretira važnost i ulogu *podatka* u poslovnoj inteligenciji. Razlog tome jeste činjenica da je podatak nosilac svih informacija a samim tim i znanja. Poseban akcenat je stavljen na kvalitet podataka, jer od toga naprosto zavisi i informacija, odnosno dobijeno znanje, na osnovu koga treba doneti ispravnu upravljačku poslovnu odluku. Jer dobre odluke, po pravilu omogućavaju ostvarenje dobrih rezultata. U nastavku ovoga poglavlja ukazano na jedan gorući problem, a to je probleme šta raditi sa nedostajućim podacima. Za ovakav realan problem, predstavljena su dva pristupa, a to su: jedan tradicionalni pristup i jedan pristup predviđanja nedostajućih podataka.

Peto poglavlje ukazuje na celovit osvrt na sisteme poslovne inteligencije. Implementirani model poslovne inteligencije proširuje se novim alatima za podršku odlučivanju, sa posebnim osvrtom na grupu, odnosno članove tima koji u procesu odlučivanja, treba da predoče najprihvatljivije rešenje. Na početku ovog poglavlja, dat je osnovni koncept poslovne inteligencije. Zatim se isti sistemski proširuje i kompletira sistem poslovne inteligencije. Naravno osnova za svaki sistem poslovne inteligencije jeste bogata baza podataka ili skladište podataka. To je dovoljan razlog da je u ovom poglavlju posebna pažnja posvećena osnovnim koracima za projektovanje i razvoj skladišta podataka. Kao veza u odnosu na prethodno poglavlje, posebno se ukazuje na proces ekstrakcije, transformacije i učitavanja podataka. Na samom kraju ovoga poglavlja, ukazano je na osnove otkrivanja znanja u samim podacima.

U okviru *šestog poglavlja*, dat je prikaz aplikativnog dela sistema poslovne inteligencije. Ukazano je se kako je moguće upotrebiti sistem poslovne inteligencije ako su podaci i sama baza podataka, atributi koji opisuju saobraćajne nezgode. Konkretno, radi se o delu baze podataka koja tretira saobraćajne nezgode na teritoriji grada Beograda. Podaci su transformisani i u bezličnoj su formi, tako da je akcenat na aplikativnoj i akademskoj formi sistema poslovne inteligencije. U prvom delu ovog poglavlja prikazan je sam razvoj skladišta podataka, sa posebnim osvrtom na triplet, nezgoda, vozilo i lice. U drugom delu ovog poglavlja prikazan je sam rad sa analitičkom bazom, i to kroz dva softverska alata. Na samom kraju ovoga poglavlja ukazano je na primer rada sa analitičkom bazom podataka.

Sledeće *sedmo poglavlje* predstavlja logički nastavak prethodnog, sa posebnim osvrtom na menadžment znanja i to preko celina modelovanja i upravljanja znanjem. Jer je kroz disertaciju i sva poglavlja, vezivno tkivo upravo je znanje. Zato se u ovom poglavlju na početku ukazuje na sam fenomen znanja. Nakon toga, a sa aspekta savremenog poslovanju kompanije, uveden koncept menadžmenta znanja, kao i osnovni koraci za projektovanje

sistema menadžmenta znanja. Na kraju ovoga poglavlja, ukazano je na tehnike modelovanja znanja, kao i na samu studiju slučaja o mogućnostima uvođenja menadžmenta znanja.

Osmo poglavlje daje akcenat na grupe i društvene mreže, jer se radi o unapređenju sistema poslovne inteligencije, radom grupnog odlučivanja. Upravo zbog samog naslova disertacije, na početku ovog poglavlja prvo su predstavljani neki osnovni formalizmi za prikaz i predstavljanje društvene mreže, jer grupa u procesu grupnog odlučivanja, naprosto je jedan vid društvene mreže. Pored toga u ovom poglavlju, poseban akcenat je stavljen na prikaz važnih osobina društvene mreže sa aspekta poslovne inteligencije.

Deveto poglavlje ukazuje na pokušaj da se modelima i metodama ekstremizacije na grafovima, pokuša doći do odgovora: Koji je to najkraći put u grupi za prenos i distribuciju kvantuma znanja. Uveden je jedan jasan matematički formalizam, kao pandam problemu trgovačkog putnika. S tim što resurs za distribuciju, nije nista drugo nego upravo znanje, a umesto lokacija ili gradova prolaska, kretanje se realizuje od jednog do drugog donosioca odluke u društvenoj mreži. Prikazana je studija slučaja, sa jednim hipotetičkim pristupom sa aspekta koeficijenta efikasnosti u odnosu na znanje koje poseduje svaki donosilac odluke i potrebno vreme za donošenje odluke po svakom učesniku sesije.

Na samom kraju disertacije u nastavku se ukazuje na korištenu raspoloživu literaturu i posećene internet sajtove.

Što se tiče daljih pravaca istraživanja, oni se kreću u dva pravca. *Prvi* se odnosi na mogućnosti projektovanja matematičkih modela za ekstremizaciju protoka znanja u sesiji tokom grupnog odlučivanja. *Drugi* pravac istraživanja se odnosi na mogućnosti definisanja koeficijenta važnosti (vektora težinskih koeficijenata) ili ekspertize znanja svakog člana tima.

11. LITERATURA

- [1]. Adèr, H.J., Chapter 13: *Missing data*. In Adèr, H.J., & Mellenbergh, G.J. (Eds.) (with contributions by Hand, D.J.), *Advising on Research Methods: A consultant's companion*, 2008.
- [2]. Abdullah, M.S., Benest, A. E., and Kimble, C., *Knowledge Modelling Techniques for Developing Knowledge Management Systems*. 3rd European Conference on Knowledge Management, Dublin, Ireland. http://www.imamu.edu.sa/DContent/IT_Topics/Knowledge%20Modelling%20Techniques%20for%20Developing%20Knowledge.pdf jul 2011.].
- [3]. Albescu, F., Pugna, I, Paraschiv, D., *Business Intelligence & Knowledge Management – Technological Support for Strategic Management in the Knowledge Based Economy*, Academy of Economic Studies, Bucharest, 2008.
- [4]. Allison, P. D. *Missing data*. Thousand Oaks, CA: Sage Publications, 2001.
- [5]. Ana Candida Cruz Natali, Ricardo de Almeida Falbo., *Knowledge Management in Software Engineering Environments*. Proc. of the 16 th Brazilian Symposium on Software Engineering, 202.
- [6]. Awad E. & Ghayiri H, *Knowledge Management*, Prentice Hall, 2004.
- [7]. Awad, M. M. I., Abdullah, M. S., & Ali, A. B. M., *Extending ETL framework using service oriented architecture*. Procedia Computer Science, 3, 2011. Elsevier. Dostupno na: <http://www.sciencedirect.com/science/article/B9865-527GFKD-92/2/0bf38d931ef19898694a8d5aa1e83399>]:
- [8]. Ballou, D., Tayi, G., *Enhancing Data Quality in Data Warehouse Environments*, COMMUNICATIONS OF THE ACM, 1999.
- [9]. Ballou, D.P. and G.K. Tayi, *Enhancing Data Quality in Data Warehouse Environments*. COMMUNICATIONS OF THE ACM, 1999.
- [10]. Barry, D., *Data Warehouse from Architecture to implementation*, Addison-Wesley, 1997.
- [11]. Breiger, Ronald L, "The Analysis of Social Networks", University of Arizona, Tucson, 2004.
- [12]. Bertossi, L., *Data Quality in Information Integration and Business Intelligence*, Carleton University School of Computer Science Ottawa, Canada, 2010.

- [13]. Bhavani, T., *Data Mining: Technologies, Techniques, Tools and Trends*, 2009.
- [14]. Bing, L., *Web Data Mining*, Springer-Verlag Berlin Heidelberg, 2007.
- [15]. Bonacich, F., Power and centrality: a family of measures. *American Journal of Sociology*, 1987.
- [16]. Bogdan D. C., Rudolph, L.M., Kenneth, M., *The Impact of UML Class Diagrams on Knowledge Modeling, Discovery and Presentations*. *Journal of Information Technology Impact*, USA, 2003.
- [17]. Böhnlein, M., Ulbrich-vom Ende, A., *Business process oriented development of data warehouse structures*, 2000, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.83.884>, november 2011.
- [18]. Böhnlein, M., Ulbrich-vom Ende, A., *Deriving Initial Data Warehouse Structures from the Conceptual Data Models of the Underlying Operational Information Systems*, 1999, <http://www.wi-inf.uni-duisburg-essen.de/MobisPortal/pages/rundbrief/pdf/BoUI99.pdf>, november 2011.
- [19]. Bonifati, A., Cattaneo, F., Ceri, S., Fuggetta, A., Paraboschi, S., *Designing data marts for data warehouses*, 2001, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.104.2850>, november 2011.
- [20]. Brownem, D., Desmeijter, B., *IBM Cognos Business Intelligence*, V10.1 Handbook, Redbook, 2010.
- [21]. Borgatti, S. P., & Carboni, I., On Measuring Individual Knowledge in Organizations. *Organizational Research Methods*, 10(3 449-462), (2007).
- [22]. Bruckner, R., List, B., Schiefer, J., *Developing requirements for data warehouse systems with use cases*, 2001, http://www.ifs.tuwien.ac.at/~bruckner/pubs/amcis2001_requirements.pdf, november 2011.
- [23]. Bui, T.X., “*Cooperative Co-oP group decision support systems for multiple criteria group decision making*”, Springer-Verlag, Berlin Heidelberg, 1987.
- [24]. Cabibbo, L., Torlone, R., *A Logical Approach to Multidimensional Databases*, 1998, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.102.8903>, november 2011.
- [25]. Casati, F., Palpanas, S., *Managing Data Quality in Business Intelligence Applications*, University of Trento, 2010.

- [26]. Chau K.W., Cao Y., Anson M., and Zhang J., *Application of data warehouse and Decision Support System in construction management*, Automation in Construction, Vol.12, 2002.
- [27]. Chaudhuri S., and Dayal U., *An Overview of Data Warehousing and OLAP Technology*, ACM Sigmod Record, 1997.
- [28]. Ćirić B., *Poslovna inteligencija*, Data Status, Beograd, 2006.
- [29]. Čupić, M., Suknović, M., *Odlučivanje*, FON, Beograd, 2008.
- [30]. Crager, J., Lemons, D., *Measuring the Impact of Knowledge Management*, American Productivity and Quality Center, 2003.
- [31]. Curko, K., Vuksic, V. B., Loncar, A., *The Role of Business Process Management Systems and Business Intelligence Systems in Knowledge Management*, International Journal of Computers and Communications, 2009.
- [32]. Daniel, F., et al. *Managing Data Quality in Business Intelligence Applications. in VLDB Endowment*, Auckland, New Zealand: ACM, 2008.
- [33]. David C. Howell, *Treatment of missing data* (http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data/Missing.html) mart 2010.
- [34]. Delibašić B, *Projektovanje i iplementacija sistema menadžmenta znanja*, Magistarska teza, FON, Beograd, 2004.
- [35]. Delibašić, B., Suknović, M., Jovanović, M., *Algoritmi mašinskog učenja za otkrivanje zakonitosti u podacima*, FON, Beograd, 2009.
- [36]. Despres, C., Chavel, D., *A thematic analysis of the thinking in knowledge management*, Knowledge Horizons: The Present and the Promise of Knowledge Management, Butterworth – Heinemann, Oxford, 2000.
- [37]. DeSanctis, G., Gallupe, B., "A Foundation For The Study of Group Decision Support Systems", Management Science Vol. 33 No. 5, 1987.
- [38]. English, L., *Improving Data Warehouse and Business Information Quality*, John Wiley & Sons, 1999.
- [39]. Freitas, G.M., Laender, A.H.F., Campos, M.L., *MD2: Getting users involved in the development of data warehouse application*, 2001, <http://ceur-ws.org/Vol-58/laender.pdf>, novembar 2011.
- [40]. Gao, J., S. Lin, and A. Koronios. Data Quality in Engineering Asset Management Organisations – Current Picture in Australia. in 1st World Congress on Engineering Asset Management (WCEAM). 2006. Gold Coast, Australia.

- [41]. Giorgini, P., Rizzi, S., Garzetti, M., *Goal-oriented requirements analysis for data warehouse design*, 2005, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.79.4326>, november 2011.
- [42]. Giorgini, P., Rizzi, S., Garzetti, M., *GRAnD: A goal-oriented approach to requirement analysis in data warehouses*, 2007, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.78.3431>, november 2011.
- [43]. Golfarelli, M., Maio, D., Rizzi, S., *Conceptual design of data warehouses from E/R schemes*, 1998, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.48.5342>, november 2011.
- [44]. Golfarelli, M., Rizzi, S., *A methodological framework for data warehouse design*, 1998, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.48.461>, november 2011.
- [45]. Golfarelli, M., *From User Requirements to Conceptual Design in Data Warehouse Design – a Survey*, 2009, <http://bias.csr.unibo.it/golfarelli/papers/surveygolfarelli08.pdf>, november 2011.
- [46]. Gronau, N., Weber, E., *Management of Knowledge Intensive Business Processes*, Business Process Management, Springer, Heidelberg, 2004.
- [47]. Gronau, N., Korf, R., Müller, C., *KMDL - Capturing, Analysing and Improving Knowledge-Intensive Business Processes*, Journal of Universal Computer Science, JUCS, 2005.
- [48]. Guo, Y., Tang, S., Tong, Y., & Yang, D., *Triple-driven data modeling methodology in data warehousing: a case study*, 2006, http://www.fing.edu.uy/inco/grupos/csi/esp/Cursos/cursos_act/2007/DataWarehouse/Material/p59-guo-DOLAP06.pdf, november 2011.
- [49]. Hibbard, J., *Knowing What We Know*, Information Week, October, 1997.
- [50]. http://ac.els-cdn.com/S0167923606000303/1-s2.0-S0167923606000303-main.pdf?_tid=e8eee920-0be3-11e2-a877-00000aab0f6b&acdnat=1349108606_ed3c2e45378342930e866193d0a1f334, oktober 2012.
- [51]. http://ac.els-cdn.com/S0169023X03000351/1-s2.0-S0169023X03000351-main.pdf?_tid=e8d3ba14-0bdf-11e2-ace6-00000aab0f01&acdnat=1349106888_39a5f308dcbc6df4b79fe52da23d368c, oktober 2012.

- [52]. http://ac.els-cdn.com/S0926580506000550/1-s2.0-S0926580506000550-main.pdf?_tid=e1b19132-0be1-11e2-999b-00000aab0f26&acdnat=1349107736_c6a4174d46feef0c0eef731dd3019450 ,
oktobar 2012.
- [53]. http://ac.els-cdn.com/S1877705811065672/1-s2.0-S1877705811065672-main.pdf?_tid=927e7ecc-0be1-11e2-bec0-00000aab0f6b&acdnat=1349107602_09b04b481b8e6e3fd198710a69ea6b02,
oktobar 2012.
- [54]. http://asbbs.org/files/2011/JBEL_Vol_3_2011.pdf, novembar 2012.
- [55]. http://books.google.com/books?hl=sr&lr=&id=F75MqHCmkXsC&oi=fnd&pg=PA163&ots=BxnSz9q4WC&sig=W6_K4-YhcNZJ5T82QQIS-YdSFfw#v=onepage&q&f=false
- [56]. <http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.31.699> jul 2011.
- [57]. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.85.4340> jul 2011.
- [58]. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.86.2465> jul 2011.
- [59]. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.97.680> jul 2011.
- [60]. http://help.sap.com/saphelp_nw70/helpdata/en/b2/e50138fede083de10000009b38f8cf/frameset.htm februar 2011.
- [61]. <http://i-know.tugraz.at/blog/2004/11/defining-an-infrastructure-for-knowledge-intensive-business-processes> jul 2011.
- [62]. <http://is2.lse.ac.uk/asp/aspecis/20080189.pdf>
- [63]. http://link-to-results.com/index.php?option=com_content&view=article&id=48:social-network-analysis&catid=48:sna-and-neadership-networks&Itemid=56 jul 2011.
- [64]. http://media.wiley.com/product_data , mart 2010.
- [65]. <http://myweb.polyu.edu.hk/~cekwchau/AIC1.pdf>, oktobar 2012.
- [66]. <http://research.microsoft.com/pubs/76058/sigrecord.pdf>, novembar 2012.
- [67]. <http://web.efzg.hr/dok/pds/statistickemetode/> , april 2010.
- [68]. http://www.ef.uns.ac.rs/Download/metodologija_nir/16%20obrada%20i%20analiza%20podataka.pdf, mart 2010.
- [69]. <http://www.business-intelligence-secrets.com/2011/04/14/knowledge-management-and-business-intelligence/> "Knowledge Management and Business Intelligence", oktobar 2011.

- [70]. <http://www.business-software.com/offer/top-10-business-intelligence-software/> "Top 10 Business Intelligence Software" 3.2
- [71]. <http://www.cin.ufpe.br/~ejvm/OLAP/p9-niemi.pdf>, decembar 2012.
- [72]. <http://www.elegantjbi.com/resources/bi-whitepapers/open-source-bi-and-proprietary-bi.htm>, septembar 2012
- [73]. <http://www.elegantjbi.com/resources/bi-whitepapers/operational-bi.htm>, septembar 2012
- [74]. <http://www.ijcttjournal.org/volume-2/issue-2/IJCTT-V2I2P133.pdf>, decembar 2012.
- [75]. http://www.imamu.edu.sa/DContent/IT_Topics/Knowledge%20Modelling%20Techniques%20for%20Developing%20Knowledge.pdf jul 2011.
- [76]. http://www.jucs.org/jucs_14_4/analyzing_wiki_based_networks/jucs_14_4_0526_0545_mueller.pdf
- [77]. <http://www.microstrategy.com/download/files/whitepapers/open/Gartner-Magic-Quadrant-for-BI-Platforms-2012.pdf> oktobar 2012.
- [78]. Hanneman, Robert A. and Riddle, Mark, *Introduction to social network methods*, University of California, Riverside, 2005.
- [79]. Hurtado. C. A., Mendelzon A. O., *OLAP Dimension Constraints*, University of Toronto, 2010.
- [80]. Hüsemann, B., Lechtenbörger, J., Vossen, G., *Conceptual Data Warehouse Design*, 2000, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.33.670>, novembar 2011.
- [81]. Imhoff, C., Galembo, N., Geiger, J. G., *Mastering Data Warehouse Design: Relational and Dimensional Techniques*, Wiley Publishing, Inc., 2003.
- [82]. Inmon, W., *Building the Data Warehouse*, John Wiley & Sons, Inc., 2002.
- [83]. Iščlamović, S., *Mogućnosti primene poslovne inteligencije za analizu i predviđanje uspeha studiranja*, Doktorska disertacija, FON, Beograd, 20132.
- [84]. Jiwei, H., Micheline, K., *Data Mining: Concepts and Techniques*, Simon Fraser University, 2001.
- [85]. Jones, P., *SAP Business Information Warehouse Reporting*, McGraw-Hill, United States, 2008.
- [86]. Kasabov, N. K., *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*. MIT press, (1996).

- [87]. Kevin, D., *Barriers to Effective Use of Knowledge Management Systems in Software Engineering*. Communications of the ACM, Vol. 46, No. 1, 2003. [http://gcc.uni-paderborn.de/www/wi/wi2/wi2_lit.nsf/663247270b635985c1256bc900519bef/4570465d846fd0ecc1256cb20050757b/\\$FILE/p99-desouza.pdf](http://gcc.uni-paderborn.de/www/wi/wi2/wi2_lit.nsf/663247270b635985c1256bc900519bef/4570465d846fd0ecc1256cb20050757b/$FILE/p99-desouza.pdf)
- [88]. Kimball R., M. Ross, W. Thornthwaite, J. Mundy and B. Becker, *The data warehouse lifecycle toolkit: practical techniques for building data warehouse and business intelligence systems*, Wiley Computer Publishing, New York, 2008.
- [89]. Krčevinac, S., Čangalović, M., i drugi ., *Operaciona istraživanja 2*. Beograd: Fakultet organizacionih nauka, 2010.
- [90]. Larissa T. Moss, Shaku Atre, *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*, Addison Wesley, 2003.
- [91]. List, B., Bruckner, R., Machaczek, K., Schiefer, J., *A comparison of data warehouse development methodologies: case study of the process warehouse*, 2002, <http://www.springerlink.com/content/fc30yqj49749h54g/>, novembar 2011.
- [92]. List, B., Schiefer, J., Tjoa, A. M., *Process-oriented requirement analysis supporting the data warehouse design process: A use case driven approach*, 2000, <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=78AC122ECE33E6FD64896AD4DBBCD15C?doi=10.1.1.71.9751&rep=rep1&type=pdf>, novembar 2011.
- [93]. Little, R. J. A. and D. B. Rubin., *Statistical analysis with missing data*, measurement. Amsterdam: Elsevier, 1997.
- [94]. Luetić A., 2008, „*Primjena koncepta poslovne inteligencije u procesu strategijskog i financijskog upravljanja*“, magistarski rad, Ekonomski fakultet, Split.
- [95]. Levi-Jakšić, M., Marinković, S., & Petković, J., *Menadžment inovacija i tehnološkog razvoja*. Beograd: Fakultet organizacionih nauka, 2012.
- [96]. Luján-Mora, S., Trujillo, J., *A Comprehensive Method for Data Warehouse Design*, 2003, <http://dss.trilog.com/DSS/documents/A%20Comprehensive%20Method-for%20DW%20Design.pdf>, novembar 2011.
- [97]. Malhotra, J., *Knowledge Management for the New World of Business*, www.brint.com/km/whatis.htm, August 2002.
- [98]. Malinowski, E., Zimanyi, E., *Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications*, Springer, 2008.

- [99]. Mazon, J., Pardillo, J., Trujillo, J., *A Model-Driven Goal-Oriented Requirement Engineering Approach for Data Warehouses*, 2007, <http://www.springerlink.com/content/978-3-540-76291-1/#section=362739&page=1&locus=16>, november 2011.
- [100]. Mazon, J., Trujillo, J., Serrano, M., Piattini, M., *Designing data warehouses: From business requirement analysis to multidimensional modeling*, 2005, <http://www.citeulike.org/group/14571/article/8855085>, november 2011.
- [101]. McDonald, K., Wilmsmeier, A., Dixon D.C., Imon W.H., *Mastering the SAP Business Information Warehouse: Leveraging the Business Intelligence Capabilities of SAP NetWeaver*, 2nd Edition" Wiley, Canada, 2006.
- [102]. Michael, O., *BI Means Business*, System iNEWS, 201.
- [103]. Moody, D., Kortink, M., *From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design*, 2000, http://ssdi.di.fct.unl.pt/bddw/material_apoio/artigos/files/2000-Moody.pdf, november 2011.
- [104]. Moon S. W., Kim J. S., Kwon K. N., *Effectiveness of OLAP-based cost data management in construction cost estimate*, Automation in Construction, Vol.16, 2007.
- [105]. Mueller, C., Gronau, N., Lembcke, R., *Application of Social Network Analysis in Knowledge Processes*, In 16th European Conference on Information Systems (Golden W, Acton T, Conboy K, van der Heijden H, Tuunainen VK eds.), 2196-2208, Galway, Ireland, 2008.
- [106]. Müller, C., Meuthrath, B., Baumgra, A., *Analyzing Wiki-based Networks to Improve Knowledge Processes in Organizations*. Journal of Universal Computer Science, vol. 14, no. 4, 2008.
- [107]. Mundy, J., Thornthwaite, W., Kimball, R., *The Microsoft® Data Warehouse Toolkit: With SQL Server™ 2005 and the Microsoft® Business Intelligence Toolset*, Wiley Publishing, Inc., 2006
- [108]. Murray, E. J., *Advances in Knowledge Management Development Methodologies*, Information Science Reference, 2007.
- [109]. Newman, B., *Managing Knowledge: The HR Role*, HR Executive Review, vol. 6, no. 4, 1999, 2007.
- [110]. Nichols, D.M., et al., A lightweight metadata quality tool, in Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries. 2008, ACM: Pittsburgh PA, USA. p. 385-388.

- [111]. Niemi T., Nummenmaa J., and Thanisch P., *Constructing OLAP Cubes Based on Queries*, ACM Fourth International Workshop on Data Warehousing and OLAP, pp 9–11, 2001.
- [112]. Nonaka, I., Takeuchi, H., *The Knowledge-Creating Company: How Japanese Companies*, Tokyo, 2001.
- [113]. Norbert, G., Edzard, W., *Defining an Infrastructure for Knowledge Intensive Business Processes*, I-KNOW '04, 2004.
- [114]. Olivier, S., *Social network analysis*, 2004.
<http://www.adb.org/Documents/Information/Knowledge-Solutions/Social-Network-Analysis.pdf> jul 2011.
- [115]. Paolillo, John C, *Social Coherence in Computer-Mediated Communication*, Indiana University, Bloomington, 1996.
- [116]. Paim, F., Carvalho, A., Castro, J., *Towards a methodology for requirements analysis of data warehouse systems*, 2002,
<http://www.lbd.dcc.ufmg.br/colecoes/sbes/2002/008.pdf>, novembar 2011.
- [117]. Paim, F.R.S., Castro, J.F.B., *DWARF: An approach for requirements definition and management of data warehouse systems*, 2003,
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.9692>, novembar 2011.
- [118]. Phipps, C., Davis, K. C., *Automating Data Warehouse Conceptual Schema Design and Evaluation*, 2001, <http://ceur-ws.org/Vol-58/hipps-davis.pdf>, novembar 2011.
- [119]. Pickles, A., *Missing data*, problems and solutions, 2005.
- [120]. Rubin, D.B., *Multiple imputation for nonresponse in surveys*, John Wiley & Sons, New York, 1997.
- [121]. Sabherwal, R., *Knowledge Management and Business Intelligence: Resource Competitors or Synergistic Forces?*, University of Missouri Curators, 2009.
- [122]. Schafer, J.L. *Analysis of incomplete multivariate data*, Chapman & Hall, London. 1997.
- [123]. Schafer, J.L., Olsen, K. *Multiple imputation for multivariate missing-data problems: A data analyst's perspective*, *Multivariate Behavioral Research*. 33, 1997.
- [124]. Schiefer, J., List, B., Bruckner, R., *A holistic approach for managing requirements of data warehouse systems*, 2002,
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.20.783>, novembar 2011.
- [125]. Shen L., Liu, S., Chen, S., Wang, X., *The Application Research of OLAP in Police Intelligence Decision System*, *Procedia Engineering*, Vol.29, pp 397-402, 2012.

- [126]. Singh K., Upadhyaya S., and Cantt M., *Constructing the OLAP Cube from Relational Databases / Flat Files*, International Journal of Computer Trends and Technology, Vol. 2, No. 2, 2011.
- [127]. Suknović, M., Delibašić, B., *Poslovna inteligencija i sistemi za podršku odlučivanju*, FON, Beograd, 2010.
- [128]. Suknović, M., Krulj, D., Čupić, M., Martić, M., *Projektovanje i razvoj skladišta podataka studentske službe FONa*, SYMORG, Zlatibor, 2002.
- [129]. Suknović, M., *Razvoj metodologije podrške grupnom odlučivanju*, Doktorska disertacija, FON, Beograd, 2001.
- [130]. Teodorović, D. *Transportne mreže*,. Beograd: Saobraćajni fakultet, 2007.
- [131]. Thierauf, R. J., *Knowledge management systems for business*, Quorum Books, strana 23, 1999.
- [132]. Trembley, M. C., Fuller, R., Berndt, D., *Doing more with more information: Changing healthcare planning with OLAP tools*, Decision Support Systems, Vol. 43, pp 1305-1320, 2007.
- [133]. Trujillo, J., Palomar, M., Gomez, J., Song, I., *Designing data warehouses with OO conceptual models*. *IEEE Computer*, 2001, <http://www.dlsi.ua.es/~mpalomar/gold.pdf>, novembar 2011.
- [134]. Vadim, V., Ayuyev, J., Philip, W., Obradovic, Z., *Dynamic Clustering-Based Estimation of Missing Values in Mixed Type Data*, TempI, Filadelfija, 2002.
- [135]. Vaisman, A., *Data quality-based requirements elicitation for decision support systems*, Idea group publishing, 2007, <http://www.irma-international.org/viewtitle/7616/>, novembar 2011.
- [136]. Vidette, Poe., *Building a Data Warehouse for Decision Support*, Prentice Hall, 1996.
- [137]. Vinekar, V., Teng, J.T.C., Chennamaneni, A., *The Interaction of Business Intelligence and Knowledge*, Journal of International Technology and Information Management, strana 155, 2009.
- [138]. Vuleta, J., *Metode ekstremzacije na grafovima*, Naučna knjiga, Beograd, 1985.
- [139]. Verbitskiy, Y., *Data quality management un a busuness untelligence envuronment: from the lens of metadata*, University of South Australia, Australia, 2011.
- [140]. Vuković, S., Delibašić, B., Uzelac, A., Suknović, M., *A case-based reasoning model that uses preference theory functions for credit scoring*, Expert Systems With Applications, DOI: <http://dx.doi.org/10.1016/j.eswa.2012.01.181>, 2012.

- [141]. Vuleta, J., *Metode ekstremizacije na grafovima*, Naučna knjiga, Beograd, 1985.
- [142]. Winter, R., Strauch, B. *A method for demand-driven information requirements analysis in data warehousing*, 2003, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.10.8678>, novembar 2011.
- [143]. William E. Grosso, Henrik Eriksson, Ray W. Ferguson, John H. Gennari, Samson W. Tu, Mark A. Musen, 1999. *Knowledge Modeling at the Millennium* (The Design and Evolution of Protégé-2010). <http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.31.699> jul 2011.
- [144]. Wnag, J., Kourik, J. L., Maher, P. E., *Identifying Characteristics and Roles of OLAP in Business Decision Support Systems*, Journal of Business and Educational Leadership, Vol. 3, No. 1, 2011.
- [145]. www.psychwiki.com/wiki/Dealing_with_Missing_Data, april 2010.

INDEKS SLIKA

	Strana
Slika 3.1. Konceptualni okvir faza grupnog odlučivanja	18
Slika 4.1. Logička hijerarhija podataka, informacija i znanja	25
Slika 4.2. Taksonomija rudarenja podataka	29
Slika 4.3. Proces rudarenja podataka	33
Slika 4.4. Priprema podataka za proces izgradnje skladišta podataka - <i>ETL</i> proces	38
Slika 4.5. Aspekti kvaliteta podataka	42
Slika 5.1. Evolucija baza podataka	59
Slika 5.2. Rešenje je u OLAPu	61
Slika 5.3. Razvoj skladišta podataka	63
Slika 5.4. Sinonimi za otkrivanje znanja u podacima	64
Slika 5.5. Uticaj raznih disciplina na DM	65
Slika 5.6. Evolucija algoritama za otkrivanje znanja	66
Slika 5.7. Pravci razvoja DM	67
Slika 5.8. Proces otkrivanja znanja	67
Slika 5.9. Sistem poslovne inteligencije	69
Slika 5.10. Veza između konteksta i razumevanja znanja	71
Slika 5.11. Piramida znanja	72
Slika 5.12. Zahtevi u skladištu podataka	79
Slika 5.13. Nivoi apstrakcije zahteva u skladištu podataka	80
Slika 5.14. Faze u dizajnu skladišta podataka	85
Slika 5.15. Model ETL procesa	94
Slika 5.16. Tradicionalna ETL arhitektura	97
Slika 5.17. Distribuirana arhitektura ETL procesa	98
Slika 5.18. Tehnike data mininga	104
Slika 6.1. Sirovi podaci preuzeti iz MUPa Beograd	107
Slika 6.2. Dijagram objekti veze baze SQL Server	109
Slika 6.3. Procedura učitavanja podataka u SQL Server	110
Slika 6.4. Koraci transformacije podataka na tri definisana fajla	111
Slika 6.5. Učitavanje tekstualnih fajlova Nezgoda, Lice i Vozilo u SQL Server	112
Slika 6.6. Izgled ekrana za definisanje polja, strukture i tipa podataka fajla Nezgoda	115
Slika 6.7. Napunjena baza SQL Servera izvornim podacima	118
Slika 6.8. Veza analitičke baze sa SQL Serverom	119
Slika 6.9. Struktura analitičke baze podataka	120
Slika 6.10. Procesiranje baze podataka	120
Slika 6.11. Naredba za pristup podacima analitičke baze	121
Slika 6.12. Pregled ukupnog broja nezgoda, vozila i lica razvrstanih po posledicama lica	122

Slika 6.13. Pregled distribucije nezgoda, vozila i lica po času nezgode	123
Slika 6.14. Distribucija nezgoda, vozila i lica po času i svojstvu lica	124
Slika 6.15. Pristup podacima o konkretnoj nezgodi	125
Slika 6.16. Pregled zavisnosti nezgoda i lica i pristup podacima lica	125
Slika 6.17. Naredba za pokretanje alata za rad sa analitičkom bazom	126
Slika 6.18. Izbor izvora podataka i načina rada sa analitičkom bazom	127
Slika 6.19. Formiranje konekcije sa analitičkom bazaom podataka	127
Slika 6.20. Kreiranje veze sa analitičkom bazom	128
Slika 6.21. Definisane naziva i servisa za pristup analitičkoj bazi podataka	128
Slika 6.22. Povezivanje sa serverom	129
Slika 6.23. Izgled ekranske forme Excela, spremnog za rad sa analitičkom bazom	129
Slika 6.24. Pregled nezgoda i lica po posledicama lica	131
Slika 6.25. Pregled nezgoda i lica po mesecima u godini.	132
Slika 6.26. Pregled nezgoda i vozila po vrsti vozila i posledicama lica	132
Slika 6.27. Pregled nezgoda, lica i vozila po polu i kategoriji vozačke dozvole	133
Slika 6.28. Pregled nezgoda i lica po danu rođenja lica.	134
Slika 7.1. Elementi menadžmenta znanja	139
Slika 7.2. Konverzija znanja	147
Slika 7.3. Arhitektura SMZ	150
Slika 7.4. Funkcionalnosti softverskog rešenja Microsoft SharePoint	162
Slika 7.5. Izgled korisničkog panela Microsoft SharePoint 2010	166
Slika 7.6. Izgled panela za deljenje dokumenta u okviru Microsoft SharePoint 2010	167
Slika 7.7. Izgled panela za kreiranje veb stranica u okviru Microsoft SharePoint 2010	167
Slika 8.1. Akteri na grafu	170
Slika 8.2. Graf sa oznakama	170
Slika 8.3. Usmereni graf	172
Slika 8.4. Graf sa hodovima	174
Slika 8.5. Usmereni novi graf	175
Slika 8.6. Mreža zvezda	177
Slika 8.7. Mreža linija	177
Slika 8.8. Mreža krug	177
Slika 9.1. Prikaz strukture jednog ciklusa u Delfi metodi	183
Slika 9.1a. Prikaz vremenskog horizonta jednog ciklusa u Delfi metodi	183
Slika 9.2. Početni raspored čvorova	187
Slika 9.3. Parcijalna ruta DO4-DO7	188
Slika 9.4. Parcijalna ruta DO4-DO7-DO1	189
Slika 9.5. Parcijalna ruta DO4-DO7-DO1-DO8	190
Slika 9.6. Parcijalna ruta DO4-DO7-DO1-DO8-DO2	191

Slika 9.7. Parcijalna ruta DO4-DO7-DO1-DO8-DO2-DO3	192
Slika 9.8. Parcijalna ruta DO4-DO7-DO1-DO8-DO2-DO3-DO9	193
Slika 9.9. Parcijalna ruta DO4-DO7-DO1-DO8-DO2-DO3-DO9-DO5	194
Slika 9.10. Parcijalna ruta DO4-DO7-DO1-DO8-DO2-DO3-DO9-DO5-DO6	194
Slika 9.11. Kompletna ruta DO4-DO7-DO1-DO8-DO2-DO3-DO9-DO5-DO6-DO4	195

INDEKS TABELA

	Strana
Tabela 4.1. Mere varijacije pre i posle zamene srednjom vrednošću	50
Tabela 5.1. Poslovni zahtevi	81
Tabela 5.2. Korisnički zahtevi	82
Tabela 5.3. Implementacioni zahtevi	83
Tabela 5.4. Atributi implementacionih zahteva	84
Tabela 5.5. Poređenje osnovnih pristupa	92
Tabela 5.6. Poređenje Statističke analize i Data Mininga	100
Tabela 6.1. Atributi zapisa nezgoda	113
Tabela 6.2. Atributi zapisa vozilo	114
Tabela 6.3. Atributi zapisa lice	114
Tabela 7.1. Osnovne karakteristike predstavljenih tehnika modelovanja	161
Tabela 8.1. Matrica „dopadanja“	169
Tabela 8.2. Prazna matrica reda 2x4	171
Tabela 8.3. Prazna matrica reda 4x2	171
Tabela 8.4. Matrica reda 4x4	171
Tabela 8.5. Matrica bliskosti	172
Tabela 9.1. Početne vrednosti procenjenog znanja i vremena svakog DO	186
Tabela 9.2. Izračunat koeficijent efikasnosti DO	186

BIOGRAFIJA AUTORA

Srđan N. Lalić je rođen 30.12.1981.godine u Sarajevu. U Beogradu je završio osnovnu školu i XII Beogradsku gimnaziju opšteg smera. Fakultet organizacionih nauka Univerziteta u Beogradu upisao je 2000/2001 školske godine, a diplomirao 2006. godine sa temom „Menadžment znanja: Put od podatka do odluke“ i stekao zvanje diplomirani inženjer organizacionih nauka - odsek za menadžment.

2006. godine je upisao master studije na Fakultetu organizacionih nauka Univerziteta u Beogradu, smer menadžment i položio sve nastavnim planom predviđene ispite sa prosečnom ocenom 10. U oktobru 2008. godine odbranio je master rad na temu „Primena poslovne inteligencije u operativnom odlucivanju“ i stekao akademski naziv master inženjer organizacionih nauka.

Od 2008. godine zaposlen je u Republičkom zavodu za statistiku Srbije na mestu savetnika u Odseku za evropske integracije i upravljanje projektima. 2009. godine postavljen je za rukovodioca Grupe za evropske integracije i preuzima poslove koji se odnose na harmonizaciju statističkog sistema Republike Srbije sa statističkim sistemom Evropske Unije.

U septembru 2010. godine postavljen je od Vlade Republike Srbije za zamenika sekretara Pregovaračkog poglavlja 18 koje se odnosi na statistiku. Kao zamenik sekretara Pregovaračkog poglavlja 18 rukovodi radom Pregovaračkog poglavlja 18 ali učestvuje i u radu još 16 Pregovaračkih poglavlja u kojima statistika ima bitnu ulogu.

U martu 2011. godine rukovodi Light Peer Review-om (ocenjivanje statističkog sistema) koji je organizovan od strane Evrostata u kome pored predstavnika Zavoda za statistiku učestvuju i brojni korisnici statistike (mediji, fakulteti, naučno-istraživački instituti, privredne komore) kao i proizvođači statistike.

2012. godine postaje rukovodilac bilateralne saradnje između statistike Švedske i Zavoda za statistiku Republike Srbije. Kroz ovu saradnju radi na implementaciji aktivnosti neophodnih za harmonizaciju statistike Srbije sa evropskim statističkim sistemom u oblasti statistike životne sredine, statističkih metodologija, uzorka, nacionalnih računa.

Od 2012. godine član je Odbora koji se bavi saradnjom i razmenom iskustava statističkih institucija zemalja u regionu gde predstavlja Republiku Srbiju.

2016. godine postaje rukovodilac bilateralne saradnje između statistike Francuske i Zavoda za statistiku Republike Srbije.

Прилог 1.

Изјава о ауторству

Потписани-а Срђан Лалић

број индекса _____

Изјављујем

да је докторска дисертација под насловом

Унапређење система пословне интелигенције процесом групног одлучивања

- резултат сопственог истраживачког рада,
- да предложена дисертација у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, _____

Прилог 2.

**Изјава о истоветности штампане и електронске верзије
докторског рада**

Име и презиме аутора Срђан Лалић

Број индекса _____

Студијски програм _____

Наслов рада Унапређење система пословне интелигенције процесом групног одлучивања

Ментор др Милија Сукновић, ред. проф. Факултета организационих наука

Потписани/а _____

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла за објављивање на порталу **Дигиталног репозиторијума Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис докторанда

У Београду, _____

Прилог 3.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Унапређење система пословне интелигенције процесом групног одлучивања

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигитални репозиторијум Универзитета у Београду могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство
2. Ауторство - некомерцијално
3. Ауторство – некомерцијално – без прераде
4. Ауторство – некомерцијално – делити под истим условима
5. Ауторство – без прераде
6. Ауторство – делити под истим условима

(Молимо да заокружите само једну од шест понуђених лиценци, кратак опис лиценци дат је на полеђини листа).

Потпис докторанда

У Београду, _____

1. Ауторство - Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.
2. Ауторство – некомерцијално. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.
3. Ауторство - некомерцијално – без прераде. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.
4. Ауторство - некомерцијално – делити под истим условима. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.
5. Ауторство – без прераде. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.
6. Ауторство - делити под истим условима. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.