

**UNIVERZA V LJUBLJANI  
EKONOMSKA FAKULTETA**

**SPECIALISTIČNO DELO**

**ROBERT LALIĆ**



**UNIVERZA V LJUBLJANI  
EKONOMSKA FAKULTETA**

## **SPECIALISTIČNO DELO**

**VPELJAVA SISTEMA OLAP  
V PODORO OBVEŠČEVALNEMU IN ANALITIČNEMU  
DELU POLICIJE NA PODROČJU PREPOVEDANIH DROG**

**Ljubljana, maj 2005**

**ROBERT LALIĆ**

## IZJAVA

Študent Robert LALIĆ izjavljam, da sem avtor tega specialističnega dela, ki sem ga napisal pod mentorstvom doc. dr. Aleša GROZNIKA ter skladno s 1. odstavkom 21. člena Zakona o avtorskih in sorodnih pravicah dovolim objavo specialističnega dela na fakultetnih spletnih straneh.

V Ljubljani, dne 17.05.2005

Podpis: \_\_\_\_\_

## KAZALO

1. UVOD .....	1
1.1. Problematika.....	1
1.2. Namen specialističnega dela .....	3
1.3. Cilj specialističnega dela.....	3
1.4. Metode dela.....	4
2. SISTEMI ZA PODPORO ODLOČANJU.....	6
3. OSNOVNI GRADNIKI SISTEMOV ZA PODPORO ODLOČANJU .....	8
3.1. Podatkovna skladišča .....	8
3.1.1. Osnovne značilnosti in opredelitev podatkovnih skladišč .....	8
3.1.2. Arhitekture podatkovnih skladišč .....	11
3.2. Večdimenzionalno modeliranje podatkov.....	13
3.2.1. Zakaj večdimenzionalno modeliranje podatkov?.....	13
3.2.2. Osnove večdimenzionalnega modeliranja podatkov.....	14
3.2.3. Različne predstave o večdimenzionalnem modeliranju podatkov .....	16
3.3. Sprotna analitična obdelava podatkov (OLAP).....	18
3.3.1. Lastnosti sistema OLAP v primerjavi s transakcijskim sistemom.....	18
3.3.2. Zahteve za orodja OLAP.....	20
3.3.3. Primerjava sistema OLAP z rešitvami podatkovnega rudarjenja.....	21
4. DELOVANJE POLICIJE NA PODROČJU PREPOVEDANIH DROG .....	24
4.1. Obravnava prepovedanih drog s kazensko pravnega vidika .....	24
4.2. Obravnava prepovedanih drog s policijskega vidika .....	25
4.3. Vloga in pomen obveščevalnega in analitičnega dela policije.....	27
4.4. Sistem za sprotno obdelavo transakcij policije in njegovi segmenti.....	30
5. IZDELAVA TESTNEGA SISTEMA OLAP NA PODROČJU PREPOVEDANIH DROG ....	32
5.1. Analiza zahtev in definiranje problematike.....	32
5.1.1. Opis problematike .....	32
5.1.2. Definiranje potreb uporabnikov .....	33
5.1.3. Obstoječe stanje .....	34
5.1.4. Izbira podatkov in njihova priprava .....	34
5.2. Faze večdimenzionalnega modeliranja podatkov.....	37
5.2.1. Izbira procesov .....	37
5.2.2. Določitev osnovne enote.....	38
5.2.3. Postavitev osnovnih dimenzij .....	38
5.2.4. Izbira mer .....	40
5.3. Posebnosti pri večdimenzionalnem modeliranju podatkov.....	41
5.3.1. Uporaba merskih lestvic.....	41
5.3.2. Časovna dimenzija .....	41
5.3.3. Neenakomerna globina dimenzij.....	42

5.3.4. Uporaba pomožnih tabel .....	43
5.3.5. Počasi spreminjajoče se dimenzije.....	44
5.3.6. Mini-dimenzije.....	45
5.3.7. Dimenzije brez atributov.....	46
5.3.8. Uporaba združenih dejstev kot atributov .....	47
5.4. Izgradnja OLAP kocke.....	48
5.4.1. Microsoft SQL Server 2000 Analysis Services .....	48
5.4.2. Definiranje večdimenzionalne strukture OLAP kocke .....	49
5.4.3. Procesiranje kocke in izbira ustreznega načina shranjevanja.....	51
5.4.4. Izbira OLAP odjemalca.....	53
5.4.5. Matrika povezanosti podatkov .....	56
<b>6. MOŽNOSTI IN OMEJITVE UPORABE SODOBNIH INFORMACIJSKIH TEHNOLOGIJ PRI ORGANIH PREGONA.....</b>	<b>58</b>
6.1. Možnosti uporabe sodobnih informacijski tehnologij.....	58
6.2. Pravna podlaga zbiranja, upravljanja, varstva in dostopa do tajnih podatkov .....	60
6.2.1. Zbiranje in upravljanje podatkov policije .....	60
6.2.2. Varstvo osebnih podatkov.....	61
6.2.3. Dostop do tajnih podatkov in njihovo varovanje .....	63
<b>7. SKLEP.....</b>	<b>64</b>
<b>8. POJASNILA KRATIC IN SLOVAR IZRAZOV .....</b>	<b>67</b>
<b>9. LITERATURA .....</b>	<b>70</b>
<b>10. VIRI.....</b>	<b>72</b>

## SEZNAM SLIK

Slika 1: Vloga podatkovnega skladišča v procesu upravljanja poslovnega sistema .....	8
Slika 2: Centralizirana arhitektura podatkovnega skladišča.....	11
Slika 3: Distributivna arhitektura podatkovnega skladišča .....	12
Slika 4: Zvezdasta shema večdimenzionalnega podatkovnega modela .....	14
Slika 5: Prikaz podatkov v tridimenzionalni kocki .....	19
Slika 6: Primer menijskega ekrana aplikacije FIO .....	30
Slika 7: E-R diagram segmentov podatkov za pripravo testnega sistema OLAP .....	36
Slika 8: Večdimenzionalni model podatkov za področje prepovedanih drog.....	38
Slika 9: Uporaba pomožne tabele »Skupina zasega« .....	43
Slika 10: Priprava mini dimenzije »DIMDemografski«.....	46
Slika 11: Prikaz skupne dimenzije »Zakon« .....	49
Slika 12: Prikaz strukture večdimenzionalnega podatkovnega modela OLAP kocke »Zasegi«....	50
Slika 13: Primer kreiranja agregatov pri OLAP kocki »Zasegi«.....	51
Slika 14: Uporaba programa Microsoft Excel kot OLAP odjemalca.....	54
Slika 15: Priprava lokalne OLAP kocke .....	55
Slika 16: Matrika povezanosti dejanj .....	57

## SEZNAM TABEL

Tabela 1: Vrste kriminalističnih analiz in uporabljenih metod .....	28
Tabela 2: Pomeni in primeri atributov, po posameznih dimenzijah.....	39
Tabela 3: Vrednostna lestvica za prepovedane droge .....	41
Tabela 4: Prikaz hierarhij pri dimenziji »DIMZakon«.....	42
Tabela 5: Neposredni in posredni podatki, ki se nanašajo na osebe.....	45
Tabela 6: Uporaba združenih dejstev kot atributov pri dimenziji »DIMZadeva«.....	47
Tabela 7: Opredelitev povezav med dejanji znotraj posamezne osebe .....	56

## 1. UVOD

### 1.1. Problematika

Slovenski strokovnjaki imajo v zvezi s problematiko uporabe prepovedanih drog več različnih pogledov. Od skrajno liberalnih, kjer so mnenja, da je uporaba prepovedanih drog del vsakdanjega življenja in jo je potrebno tako tudi obravnavati, do rigoroznih, kateri menijo, da je uporaba prepovedanih drog škodljiva ne samo posamezniku temveč tudi družbi kot celoti ter jo je zato potrebno prepovedati in ustrezno sankcionirati. Ne oziraje se na mnenja strokovne javnosti je delovanje policije na področju prepovedanih drog usmerjeno predvsem v odkrivanje, preprečevanje in preiskovanje kaznivih dejanj, ki zadevajo tovrstno problematiko.

Specifičen odnos med oškodovancem in storilcem (kjer imata oba korist) otežuje preiskovanje tovrstne kriminalitete, kar narekuje temu, da ostaja večina kaznivih dejanj neodkritih. To dejstvo pa zahteva od policije, da konstantno zbira in vrednoti podatke in informacije z namenom odkrivanja: »Kaj se dogaja in kaj se pripravlja?« Le na ta način pridobljene informacije opredeljujejo tudi vse nadaljnje aktivnosti policije. Ob vsem navedenem pa je potrebno upoštevati še resurse s katerimi razpolaga policija, saj gre za omejena finančna sredstva, materialna sredstva in kadre.

Za uspešno preiskovanje in odkrivanje kaznivih dejanj na področju prepovedanih drog je tako potrebno usklajeno organiziranje aktivnosti in dobro poznavanje storilcev, njihovih motivov in načinu dela, dinamike delovanja trga s prepovedanimi drogami, nezakonitih poteh, obsega organizirane kriminalitete, kakor tudi trende, ki prevladujejo na njem. Toda na tem mestu se kaj kmalu soočimo s problemom, ki narekuje vprašanje na kakšen način iz vseh zbranih podatkov pridobiti vse tiste koristne informacije, ki so potrebne za učinkovito odločanje.

Brez dvoma lahko trdimo, da je v sedanjem obdobju preveč podatkov, poleg tega pa je tudi preveč vnosnih in dostopnih mest, kar povzroča vse večje težave pri obvladovanju podatkov. Zbiranje in priprava podatkov za sprejemanje pravih odločitev je tako postalo zahtevno opravilo, tako glede stroškov kot hitrosti (SRC, 2004). Državne institucije se tako srečujejo z odvečnimi, zastarelimi in včasih nepotrebnimi podatki, njihovi uporabniki pa so ponavadi obremenjeni s številnimi preobširnimi poročili in neuporabnimi informacijami. Kar pa slabo vpliva ne samo na verodostojnost informacijskega sistema, temveč s svojimi negativnimi posledicami tudi na poslovanje (CVI, str. 76, 2001). V takšnih okoljih so uporabniki, ki so dnevno postavljeni v vlogo odločanja soočeni s problemom, kako brez ustreznih podatkov sprejeti odločitev. Prav to pogojuje temu, da so le - ti prisiljeni sprejemati svoje odločitve zgolj na podlagi politike organizacije ali pa so te odvisne od njihove osebnosti (Inmon, 2002, str. 8).

Zato potrebuje organizacija eno samo skladišče podatkov (angl. Data Warehouse), ki omogoča integracijo razpršenih podatkovnih virov, v katerem so podatki celoviti, točni ter na razpolago v tistem hipu, ko so potrebni v procesu odločanja (SRC, 2004).



Podatkovno skladišče v sinergiji z dodatnimi programskimi analitičnimi orodji, kot sta sprotna analitična obdelava podatkov (angl. Online Analytical Processing - OLAP) in podatkovno rudarjenje (angl. Data Mining), predstavlja celovit pogled na poslovanje organizacije skozi različne vidike, ki nudi končnemu uporabniku podporo pri analizi in odločanju (Golob et al., 2001). Zato so sistemi za podporo odločanju, ki jih uporablja vodstvo na različnih nivojih upravljanja, danes večinoma zgrajeni nad skladišči podatkov (Štemberger et al., 2001).

Kljub celovitemu pogledu na podatke posamezne organizacije, pa še vedno ostaja problem kako podatke prikazati jasno in jih narediti uporabne. Za doseganje tega cilja potrebujemo zmogljivo OLAP orodje, ki predstavi podatke na jasen in pregleden način (Roblek, 2003, str. 323). S strani uporabnikov se pomembna prednost OLAP tehnologije kaže pri omogočanju pregledovanja podatkov iz vseh možnih perspektiv, ne da bi ti poznali nepostopkovni strukturirani poizvedovalni jezik (angl. Structured Query Language SQL) in brez nepotrebnega čakanja na pripravo poročil. Proces odločanja je tako znatno hitrejši in učinkovitejši saj uporabniki, ki so dnevno postavljeni v vlogo odločanja lahko hitro in enostavno pridobijo informacije, ki so točne in pomembne za odločanje (Babič 2002, str. 10). Ob tem pa ne pozabimo na pomen dosegljivosti podatkov, pri katerem igra pomembno vlogo spletna tehnologija, ki je doživlja veliko zanimanje v teh letih in bo prav gotovo imela še večji odmev v prihodnosti (Shim, 2002, str. 114).

Z razmahom sistemov za podporo odločanju, podatkovnih skladišč in poplave orodij za sprotno analitično obdelavo podatkov dobiva večdimenzionalno modeliranje podatkov vedno večji pomen. Saj so podatki, ki so predstavljeni v večdimenzionalnem podatkovnem modelu, običajno bolj intuitivni in preprosti za uporabo tudi tistim uporabnikom, ki niso računalniški strokovnjaki. Namen večdimenzionalnega modeliranja je tako približati podatkovni model uporabnikovemu načinu razmišljanja in mu omogočiti hitro in učinkovito delo s podatki. Večdimenzionalno modeliranje podatkov lahko tako opredelimo kot novo ime za staro tehnologijo, pri kateri pripravimo bazo podatkov enostavno in razumljivo (Kimball, 2002, str. 13).

Vendar se v praksi mnogokrat izkaže, da izdelava večdimenzionalnih podatkovnih modelov ni vedno tako preprosto, kakor nas učijo šolski primeri. Veliko strokovnjakov iz področja podatkovnih skladišč in večdimenzionalnega modeliranja podatkov je že pisalo o posameznih problemih pri modeliranju večdimenzionalnih podatkovnih modelov, za katere so tudi predlagali svoje rešitve. Takšne probleme sicer lahko kategoriziramo in si s tem pomagamo iskati podobne rešitve pri problemih, ki jih srečujemo. Vendar pa se moramo zavedati, da na takšen način ne dobimo nujno najboljše rešitve problema. Zato ne gre podcenjevati truda, ki ga je treba vložiti v temeljito pripravo večdimenzionalnega podatkovnega modela za analizo podatkov, da se na koncu izdelata učinkovit, praktično uporaben in nedvoumen model (Ferle, 2001).

## **1.2. Namen specialističnega dela**

Namen specialističnega dela je predstaviti vpliv in možnosti uporabe sodobnih informacijskih tehnologij na področju hranjenja ter analize podatkov v računalniško podprtih podatkovnih sistemih, tako s strani podjetij, ki se srečujejo z vse bolj konkurenčnim poslovnim okoljem, kakor s smeri organov pregona na področju odkrivanja, preprečevanja in preiskovanja kaznivih dejanj povezanih s prepovedanimi drogami.

Z raziskavo želimo predstaviti predvsem funkcionalnost in praktičnost uporabe orodij za sprotno analitično obdelavo podatkov, ki temelji na praktičnih spoznanjih njihovega delovanja na specifičnem področju, t.j. delovanju policije pri odkrivanju, preprečevanju in preiskovanju kaznivih dejanj povezanih s prepovedanimi drogami. Poseben poudarek je na pripravi praktično uporabnega večdimenzionalnega podatkovnega modela, ki je bil neposredna podlaga za analizo podatkov, namenjeno za pripravo strateške analize »Problematika prekomerne uporabe prepovedanih drog v RS« in testiranju uporabe OLAP orodij pri dostopu do podatkov s strani zaposlenih, ki dnevno sodelujejo v procesu odločanja.

Čeprav nekateri izdelovalci orodij in rešitev za sprotno analitično obdelavo podatkov radi prikazujejo svoje izdelke kot silno preproste za namestitev ter pripravo za začetek dela, je treba vedeti, da primeri v praksi nikoli niso brez svojih posebnosti in zapletov. Tudi sami smo se ob modeliranju večdimenzionalnega modela podatkov srečevali s številnimi posebnostmi in problemi. Ob reševanju posameznih problemov smo želeli, da posamezne rešitve ne bi bile preveč kompleksne, oziroma boljše povedano, želeli smo, da so le te čim bolj preproste in razumljive za uporabnika podatkovnega modela, kajti le uporabnik je ta, ki določa uporabnost posameznega sistema. Zato je prav tako namen specialističnega dela predstaviti probleme, s katerimi smo se srečevali pri večdimenzionalnem modeliranju podatkov ter predstavitvi rešitev, ki smo jih uporabili.

## **1.3. Cilj specialističnega dela**

Cilj specialističnega dela je prikazati možnosti uporabe sodobnih informacijskih tehnologij, kot je sprotna analitična obdelava podatkov na konkretnem delovnem področju policije, ki se sooča s problemi zagotovitve podatkov ob pravem času na pravem mestu, njihove preglednosti in kakovosti za potrebe pravih in hitrih odločitev. Z ugotovitvami raziskave želim tako vodstvenemu kadru policije kot zaposlenim v informatiki (nosilcem razvoja informacijske tehnologije v policiji) predstaviti možnosti in priložnosti uporabe sodobnih informacijskih tehnologij, s katerimi bi lahko policija znatno okrepila svoje dejavnosti usmerjene proti organiziranemu kriminalu in izboljšala obveščevalni in analitični del pri odkrivanju, preprečevanju in preiskovanju kaznivih dejanj povezanih s prepovedanimi drogami.

## 1.4. Metode dela

Metode dela, ki smo jih pri izdelavi specialističnega dela uporabili, temeljijo na preučevanju teoretične podlage, ki daje osnovo za celovito razumevanje problematike ter uporabi praktičnih spoznanj pri prilagajanju specifičnosti delovnega področja policije. Za rešitev zastavljenih raziskovalnih problemov smo uporabili nekatere metode, ki spadajo v okvir deskriptivnega raziskovanja: metodo deskripcije, komparativno metodo, zgodovinsko metodo in metodo kompilacije.

Pri preučevanju teoretične podlage smo se naslonili na strokovno literaturo tujih in domačih avtorjev, vire, prispevke in članke z novejšimi teoretičnimi spoznanji s področja sistemov za podporo odločanju, podatkovnih skladišč, večdimenzionalnega modeliranja podatkov in orodij za sprotno analitično obdelavo podatkov. Spoznanja, ki so se kot uporabna pokazala že v praksi, smo prav tako prenesli v raziskavo, hkrati pa je vključeno tudi lastno znanje pridobljeno tekom podiplomskega študija in opravljanja svojega delovnega poklica kriminalističnega analitika.

Celotna faza izgradnje testnega sistema OLAP (od popisa virov podatkov in analize obstoječega stanja do priprave odjemalcev za dostop do podatkov) temelji na samostojnem delu, kakor tudi analiza podatkov, ki je namenjena za pripravo strateške analize »Problematika prekomerne uporabe prepovedanih drog v RS«.

Pri raziskavi smo se omejili na delovno področje, ki je vezano na odkrivanje, preprečevanje in preiskovanje kaznivih dejanj povezanih s prepovedanimi drogami. Razlog za takšno odločitev najdemo v kompleksnosti in obsegu obravnavane problematike, samostojnem delu, omejenem dostopu do podatkov in cilju, da bomo z raziskavo ugotovili konkretna spoznanja, ki temeljijo na praktičnem, nedvoumnem in preizkušenem modelu. Vsekakor pa omejitev raziskave ne pomeni, da smo v specialističnem delu zanemarili druga pomembna vzporedna področja, ki niso neposredno predmet našega preučevanja.

Specialistično delo bomo pričeli z opredelitvijo sistemov za podporo odločanju, kjer bomo spoznali njihov namen, prednosti, ki jih prinaša uporaba teh in osnove gradnike sistemov za podporo odločanju. Pri osnovnih gradnikih bomo spoznali osnovne značilnosti podatkovnih skladišč, večdimenzionalnega modeliranja podatkov in uporabe sistemov za sprotno analitično obdelavo podatkov. V obravnavanem poglavju bomo prav tako predstavili različne predstave o uporabi večdimenzionalnega modeliranja podatkov pri podatkovnih skladiščih in uporabi orodij za sprotno analitično obdelavo podatkov v primerjavi z rešitvami, ki jih omogočajo tehnologije podatkovnega rudarjenja.

V četrtem poglavju vam bomo predstavili delovanje policije na področju prepovedanih drog, kjer bomo spoznali obravnavo prepovedanih drog s pravnega in policijskega vidika. Poseben poudarek bo na opredelitvi vloge in pomena obveščevalnega in analitičnega dela policije pri odkrivanju, preiskovanju in preprečevanju kaznivih dejanj.

V osrednjem poglavju specialističnega dela bomo obravnavali izdelavo testnega sistema OLAP na področju prepovedanih drog, sprva s predstavitvijo problematike, obstoječega stanja, zahtev uporabnikov in izbiro podatkov. Glede na definirano problematiko, vam bomo predstavili faze večdimenzionalnega modeliranja podatkov, ki zajemajo izbiro procesov, določitev osnovne enote, postavitev osnovnih dimenzij in določitev mer. V nadaljevanju si bomo podrobneje ogledali nekaj primerov posebnosti oz. problemov s katerimi smo se srečevali pri postavljanju večdimenzionalnega podatkovnega modela za področje prepovedanih drog. Prikazali bomo možne rešitve teh, ter nakazali smernice za reševanje podobnih problemov. Nadaljevali bomo s predstavitvijo izgradnje OLAP kocke z orodjem Microsoft SQL Server 2000 Analysis Services, kjer bomo spoznali večdimenzionalne strukture OLAP kocke, procesiranje kocke in izbiro ustreznega načina shranjevanja. Ker je OLAP odjemalec zelo pomemben člen sistema OLAP in je z vidika uporabnika celo najpomembnejši, saj mu ta omogoča uporabo sistema in dostop do želenih podatkov, vam bomo prav tako predstavili, katere možnosti imamo pri zbiru OLAP odjemalca in uporabo Microsoft Excel-a, kot OLAP odjemalca.

V šestem poglavju, vam bomo predstavili možnosti in omejitve uporabe sodobnih informacijskih tehnologij pri organih pregona na področju prepovedanih drog. Kar zadeva omejitve uporabe sistema OLAP se bomo podrobneje spoznali s pravno podlago, ki se nanaša na zbiranje, upravljanje, varstvo in dostop do tajnih podatkov s strani varnostnih organov.

V sklepnem delu specialističnega dela, vam bomo predstavili ugotovitve testiranja sistema OLAP, na področju analize podatkov za potrebe strateške analize, kakor tudi uporabe teh orodij za potrebe obveščevalnega dela policije na področju prepovedanih drog.

## 2. SISTEMI ZA PODPORO ODLOČANJU

Vsak poslovni sistem potrebuje za svoje delovanje ustrezen informacijski sistem ali več informacijskih sistemov odvisno od tega kako opredelimo meje tega sistema. Informacijski sistemi lahko namreč pokrivajo večje ali manjše poslovno področje in zadovoljujejo informacijske potrebe enega ali več uporabnikov. Odgovor na vprašanje, ali ima neka organizacija en sam informacijski sistem ali več informacijskih sistemov je v tesni povezavi z vprašanjem ali je bil informacijski sistem grajen z enotno, integralno koncepcijo, ki temelji na informacijskih potrebah celotne organizacije (Vintar, 2003, str. 75-76).

Vendar v praksi obstajajo le redke organizacije, ki imajo en sam integriran informacijski sistem, ki zadovoljuje vse uporabnike v tej organizaciji in jim omogoča, da se vsak podatek zajema samo enkrat in je dostopen poljubnemu številu uporabnikov. Informacijski sistemi se namreč pogosto načrtujejo in gradijo, da najprej zagotovijo informacijsko podporo temeljni dejavnosti organizacije in kasneje v naslednjih fazah podprejo tudi višje upravljaljske ravni (kot sta taktična in strateška raven) v organizaciji. Izkušnje v praksi, so prav tako pokazale, da splošen posloven informacijski sistem, tudi če je bil grajen z integralno koncepcijo, ne more pokrivati, določenih specializiranih informacijskih potreb, ki jih imajo vodilni delavci oz. nosilci odločitev (Vintar, 2003, str. 76).

Pri sprejemanju odločitev se nosilci odločitev velikokrat soočajo s problemi, s katerimi se še v preteklosti niso soočili, oziroma so ti problemi izjemno kompleksni in pomembni za celotno organizacijo. Kar pogojuje k ustvarjalnejšemu pristopu k reševanju tovrstnih problemov in ne k rutini ali reševanju teh po nekem programu. Take odločitve so rezultat procesa notranjega reševanja problemov, presoje, intuicije in ustvarjalnosti (Dimovski, 2002, str. 55). Ker je odločitev odvisna od tega, kako bodo posamezne alternative oziroma rešitve predstavljene, nosilci odločitev običajno od informacijskega sistema pričakujejo kakovostne podatke, katerih način predstavitve jim mora omogočati kreativno uporabo s ciljem tvorjenja novih informacij (CVI, 2001, str. 77).

Vendar so nosilci odločitev v praksi v večini obremenjeni s številnimi preobširnimi poročili in neuporabnimi informacijami, srečujejo se tako z odvečnimi, zastarelimi in včasih nepotrebnimi podatki, kar pa slabo vpliva ne samo na verodostojnost informacijskega sistema, temveč s svojimi negativnimi posledicami vpliva tudi na poslovanje (CVI, 2001, str. 76). Reševanje nastalega položaja v večini poteka (ali je potekalo) preko dolgotrajnih procesov zbiranja in spreminjanja podatkov v zeleno obliko. Kar pa je vse skupaj pripeljalo do nezadovoljstva informatikov, kakor tudi uporabnikov podatkov. Kajti zaradi vse večjega konkurenčnejšega okolja v katerem delujejo organizacije, nihče več ni pripravljen čakati na poročila več dni ali celo tednov. Pridemo do stanja, ko nosilci odločitev prenehajo »prostiti« za poročila in pričnejo zahtevati informacije na svojem računalniku in to pregledne in preproste (Kričej, 1998, str. 8). Na osnovi teh ugotovitev je nastala cela vrsta specializiranih informacijskih sistemov med katerimi prav tako zasledimo sisteme za podporo odločanju.

Sisteme za podporo odločanju (angl. decision support information systems) lahko opredelimo kot množico strojne opreme, programske opreme in ljudi, ki pomagajo v procesu odločanja. Naloga sistemov za podporo odločanju je, da pomagajo v procesu odločanja in »ne« da sprejemajo odločitev namesto uporabnikov, ki so postavljeni v vlogo odločanja. Sistemi za podporo odločanju uporabniku tako priporočajo eno izmed možnih alternativ, poti, variant ali rešitev. Medtem ko je naloga uporabnika, da se odloči za to rešitev ali kakšno drugo (Damij et. al, 1995, 254).

Sistemi za podporo odločanju so tako namenjeni v pomoč vsakemu zaposlenemu, ki sodeluje v procesu odločanja in na vseh organizacijskih ravneh v organizaciji. Poudarek je tako na strukturiranih kot na ne-strukturiranih odločitvah (Bidgoli, 1997, str. 284; po Babič, 2002, str. 2). Pojem strukturiran problem običajno uporabljamo v primerih, ko je narava problema takšna, da omogoča bolj ali manj avtomatizirano sprejemanje ustreznih odločitev, torej takšno ukrepanje, ki ne zahteva neposrednega vpletanja nosilca odločitve in njegove presoje. Vendar sisteme za podporo odločanju karakterizira, pomoč pri reševanju delno strukturiranih problemov. To je problemov, pri katerih nekatere faze v postopku odločanja lahko avtomatiziramo, medtem, ko je pri drugih fazah potrebna sprotne presoja nosilca odločitve in posledično tudi njegovo prevzemanje odgovornosti za izbiro izbrane alternative (Čibej, 1998, str. 8).

Od leta 1990 so se na trgu uveljavila štiri ključna orodja za izgradnjo sistemov za podporo odločanju. Kot prvo orodje lahko štejemo podatkovno skladišče, naslednji dve, ki sta sledili in nadgradili prvo, sta sprotne analitična obdelava podatkov (OLAP) in podatkovno rudarjenje. Medtem, ko je sedaj pri izgradnji sistemov za podporo odločanju zaznati močan trend uporabe orodij povezanih s spletno tehnologijo, ki odpirajo nove zmožnosti sistemov za podporo odločanju (Shim, 2002, str. 114), bo v prihodnosti poudarek predvsem na mobilnih orodjih, mobilnih e-storitvah in brezžičnih protokolih. Na ta način bo omogočen neprestani dostop do informacij in orodij za podporo odločanju, boljše sodelovanje in lažje interaktivno odločanje kljub geografskim omejitvam (Shim, 2002, str. 121).

V splošnem velja, da sistem za podporo odločanju doseže svoj namen, če njegovi uporabniki najdejo korist v njem in si tako težko predstavljajo, kako bi lahko le brez tega učinkovito opravljali svoje delo. A pojavljajo se podjetja, ki ponujajo in tržijo vodstvene informacijske sisteme (angl. executive information systems), ki po svojih lastnostih posegajo na področje sistemov za podporo odločanju (Babič, 2002, str. 4).

Torej v kakšnem odnosu so vodstveni informacijski sistemi in sistemi za podporo odločanju? Poenostavljeno bi sicer lahko rekli, da so si oboji v sorodu, samo, da so vodstveni informacijski sistemi običajno preprostejši in se uporabljajo za bolj vsakodnevne taktične odločitve (Vintar, 2003, str 79). Vendar jih po drugi strani nikakor ne smemo enačiti zaradi njunih razlik v področju delovanja, zlasti glede na strukturiranost problemov, ki jih posamezni sistemi rešujejo (Čibej, 1998, str. 11).

### 3. OSNOVNI GRADNIKI SISTEMOV ZA PODPORO ODLOČANJU

#### 3.1. Podatkovna skladišča

##### 3.1.1. Osnovne značilnosti in opredelitev podatkovnih skladišč

Namen upravljanja poslovnega sistema je doseganje njegovih ciljev, kot so finančna uspešnost, kakovost proizvodov in storitev, izpolnjevanje rokov ter fleksibilnost v smislu prilagajanja poslovnega sistema novo nastalim situacijam. Upravljanje, tako zahteva številne odgovorne odločitve, ki jih lahko sprejema le človek. Pri tem pa si lahko ta bistveno pomaga z ustreznimi informacijami, ki mu jih nudi informacijski sistem, ki zajema podatke tako iz ustreznih poslovnega sistema, kakor tudi iz širšega okolja (Rajkovič, 2004).

Informacijski sistemi na operativnem nivoju ob množici podatkov, ki se v organizacijah zbirajo v operativnih bazah podatkov, pogosto ne zagotavljajo takšnega dostopa do podatkov in njegove uporabe, ki bi bili namenjeni v podporo odločitvenim procesom v organizaciji, kajti njihov cilj je omogočiti učinkovito opravljanje vsakodnevnih opravil v organizaciji. Ker je pridobivanje informacij za potrebe odločitvenih procesov ponavadi vezano na obdelavo velikih količin zgodovinskih podatkov, bi to predstavljalo hude obremenitve operativnih podatkovnih virov, kar bi slabo vplivalo na izpolnjevanje njihovih osnovnih zahtev (kot sta: razpoložljivost in hitrost).

Organizacije se tako soočajo z naraščajočimi potrebami po zbiranju točnih, integriranih in predvsem pravočasnih informacij, potrebnih za podporo odločitvenim procesom. Zato potrebuje organizacija podatkovno skladišče podatkov (angl. Data Warehouse), v katerem so vsi poslovni podatki in to celoviti, točni ter na razpolago v tistem hipu, ko so potrebni v procesu odločanja (SRC, 2004).

Vlogo podatkovnega skladišča v procesu upravljanja poslovnega sistema, vam prikazujemo na sliki 1, na kateri je vloga podatkovnega skladišča prikazana kot vez med informacijskim sistemom in uporabniki, ki so dnevno postavljeni v vlogo odločanja.

Slika 1: Vloga podatkovnega skladišča v procesu upravljanja poslovnega sistema



Vir: Rajkovič, 2004

Podatkovno skladišče je osrednja, od operativnih zbirk ločena podatkovna zbirka podatkov, ki je usmerjena na ključne poslovne vire (kot so trgi, kupci, izdelki...itd.) in namenjena shranjevanju in dostopu do teh podatkov za potrebe odločanja in vodenja. Podatkovno skladišče lahko opredelimo kot zbirko podatkov, ki je (Inmon, 1996, str. 30):

- predmetno usmerjena,
- integrirana,
- nespremenljiva in
- časovno dimenzionirana.

Sistemi za podporo odločanju, ki jih danes uporabljajo organizacije na različnih nivojih upravljanja, so zato večini zgrajeni nad podatkovnimi skladišči (Štemberger et al., 2001). Kajti podatkovna skladišča s svojimi strukturami in procesi, prirejenimi v podporo poslovnemu procesu, omogočajo tistim, ki odločajo in skrbijo za razvoj organizacije, celovit pogled na podatke posamezne organizacije, ne glede na uporabljene strojne in programske rešitve v posameznih operativnih okoljih (Golob et al., 2001).

Glede na to, da podatki v podatkovnih skladiščih izhajajo iz različnih zunanjih virov in operativnih sistemov, ki so včasih težko povezljivi (različne tehnologije na izvoru, različni formati podatkov), pogosto pa tudi nekonsistentni (različne vrednosti pomensko enakih atributov), je pri prenosu podatkov v skladišče potrebno poskrbeti za integracijo in transformacijo teh podatkov (Jaklič, 2001, str. 3). Podatke med prenosom prav tako oplemenitimo, da jih reorganiziramo na način, ki je bolj primeren za uporabnike na višjem nivoju. Podatkovno skladiščenje tako vsebuje procese zbiranja in čiščenja podatkov ter njihove transformacije v obliko, ki je najprimernejša za poslovno uporabo.

Vendar kakovost podatkov sama po sebi ne bo zadošča, kajti iz podatkov moramo poiskati odgovore na manj ali bolj zapletena poslovna vprašanja, kateri odgovori naj pripeljejo do boljših poslovnih odločitev. Z drugimi besedami povedano, iz zbranih podatkov skušamo izluščiti informacije in znanje (Štemberger et al., 2001). Zato si s samim podatkovnim skladiščem ne moremo kaj veliko pomagati, saj so to le na drugačen način predstavljeni podatki. Podatkovno skladišče je zato podlaga za nadaljnje delo, iz katerega lahko z različnimi orodji in tehnikami dobimo kakovostne informacije. Smisel graditve skladišča podatkov je tako, da na njegovi podlagi gradimo sisteme za podporo odločanju, omogočamo sprotno analitično obdelavo podatkov in podatkovno rudarjenje (Povšič, 1998, str. 16).

Orodja za izkoriščanje podatkovnih skladišč uporabnikom tako omogočajo nov, bogatejši pogled na podatke. Namesto papirnih poročil so jim sedaj na voljo moderna orodja, s katerimi lahko raziskujejo podatke, interaktivne tabele omogočajo preglede na globalnem nivoju in drobljenje v podrobnosti podatkov, grafične predstavitve pa lahko velikokrat prikažejo informacijo bolj jasno in učinkovito kot velike papirne tabele. Skratka uporabniki lahko z imenovanimi orodji poiščejo informacije in povezave v trenutku, medtem ko bi s klasičnimi orodji lahko porabili dneve ali mesece (Jaklič, 2001, str. 3).



Medtem, ko je pri klasičnih sistemih za pridobivanje informacij nujno potrebno sodelovanje informatikov, programerjev in drugih strokovnjakov, so informacije iz podatkovnih skladišč hitro in relativno enostavno dostopne. Tako po eni strani uporabnikom skrajšajo čas, ki je potreben zato, da se ustrezni podatki zberejo, po drugi strani pa sprotno osveževanje podatkov v podatkovnem skladišču bistveno skrajša čas, ki poteče od nekega kritičnega dogodka. Če je bilo na primer odločanje prej odvisno od mesečnih poročil, lahko sedaj uporabniki spremljajo dogajanja dnevno. S tem lahko izkoristijo priložnosti, ki bi jim sicer ob pomanjkanju informacij ušle (Jaklič, 2001, str. 3).

Zaradi združevanja potreb po operativnih podatkih in potreb za podporo odločanju se je prav tako razvilo področje t.i. operativnih podatkovnih skladišč, ki pokrivajo potrebe po integriranem podatkovnem viru na operativnem nivoju, ki je predmetno usmerjen, integriran, kjer so podatki detajlni, spremenljivi in izkazujejo trenutne vrednosti. V primerjavi s podatkovnim skladiščem tako ne vsebuje zgodovinskih podatkov, temveč le trenutne, aktualne in detajlne podatke določene organizacije. Operativno podatkovno skladišče je tako hibridna struktura, ki izpolnjuje tako operativne kot tudi analitične zahteve, zaradi svoje dvojne narave je to najbolj kompleksen tip podatkovnega vira (Jaklič, 2004).

Podatkovno skladišče je zgrajeno z namenom zadovoljiti skupne potrebe celotnega podjetja in ne samo določenih delov oz. oddelkov. Njegova struktura v večini primerov tako ni prirejena za delo posameznega oddelka, zato se za posamezna delovna področja oz. oddelke razvijejo t.i. podatkovne tržnice oz. področna podatkovna skladišča (angl. Data Marts). V literaturi se pojavlja več imen za tovrstne vire, tako jih nekateri poimenujejo kot podatkovne baze za oddelčne sisteme za podporo odločanju, OLAP baze, večdimenzionalne podatkovne baze, agregirane baze podatkov, itd.

Ne glede na poimenovanje področnega podatkovnega skladišča predstavlja ta zbirko področnih objektov za podporo odločanju, ki temelji na potrebah danega poslovnega področja v podjetju. Tako ima oddelek za finance svojo področno podatkovno skladišče, enako oddelek za prodajo, trženje itd. Ker ima lahko vsak oddelek v podjetju svojo lastno razlago glede samega izgleda področnega podatkovnega skladišča ter kako bo le-to zadovoljevalo njihove specifične potrebe, je prednost področnih podatkovnih skladišč v izrabi njihove neodvisnosti od ostalih področnih podatkovnih skladišč (Inmon, 1999b).

Odvisno od izbrane arhitekture podatkovnega skladišča so vir področnih podatkovnih skladišč lahko operativne zbirke podatkov ali centralno skladišče podatkov. Glede na vir polnjenja področnih podatkovnih skladišč tako ločimo (Inmon, 1999a):

- odvisna področna podatkovna skladišča, katerih vir izhaja iz centralnega podatkovnega skladišča in
- neodvisna področna podatkovna skladišča, ki se polnijo vsaka zase.

### 3.1.2. Arhitekture podatkovnih skladišč

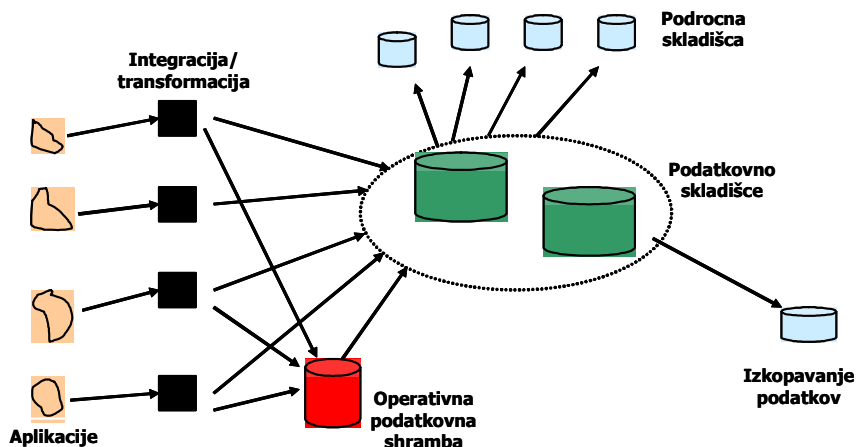
Na uspeh izgradnje podatkovnega skladišča vpliva več faktorjev, vendar ena izmed najpomembnejših odločitev s katero se mora soočiti vsak načrtovalec podatkovnega skladišča že na začetku, je izbira primerne arhitekture. Izbira arhitekture je tako kritična, saj le-ta določa podatkovni model, vlogo področnih podatkovnih skladišč ter sosledje korakov v razvojnem ciklu (Golob, 2001).

Vendar moramo pri tem opomniti, da na področju poznavanja in razumevanja arhitektur podatkovnih skladišč vlada precejšnja zmeda, saj se zaradi vse večjih napredkov v strojni in programski opremitvi nenehno pojavljajo nove rešitve. Dodatno zmedo pa povzročajo zagovorniki posameznih arhitektur s svojimi opozorili o neprimernosti ostalih arhitektur. Ker je cena projektov podatkovnega skladiščenja v primerjavi z ostalimi projekti na področju informatike dokaj visoka, je strah pred nepravilno izbiro arhitekture utemeljen (Golob, 2001).

Zato si v nadaljevanju pogledjmo tri osnovne arhitekture podatkovnih skladišč (Golob, 2001):

- V *centralizirani arhitekturi* podatkovnega skladišča je v središču podatkovno skladišče zaključenega organiziranega sistema, ki se polni iz operativnih podatkovnih baz ter operativnega podatkovnega skladišča (operativna podatkovna shramba). Izhajajoč iz arhitekture centraliziranega podatkovnega skladišča, ki vam jo prikazujemo na sliki 2 je osrednje podatkovno skladišče edini vir podatkov za področna skladišča.

Slika 2: Centralizirana arhitektura podatkovnega skladišča

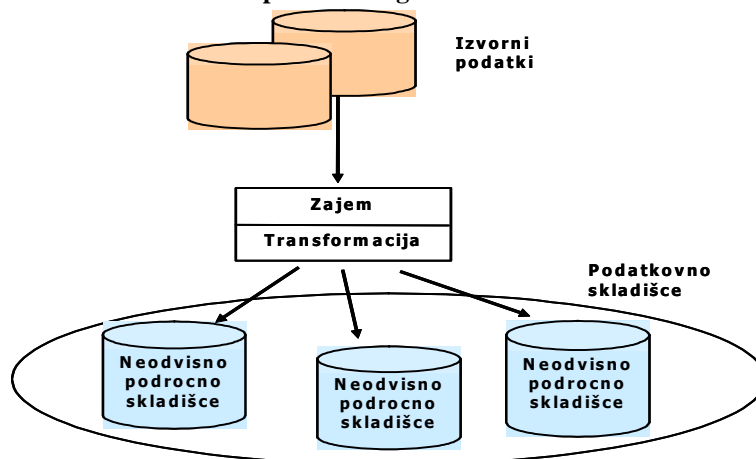


Vir: Jaklič, 2003

Razvojni cikel izgradnje podatkovnega skladišča je podatkovno voden (angl. data driven). Kljub morebitni delni denormaliziranosti strukture oziroma podatkovnega modela podatkovnega skladišča, podatki obdržijo značilnost močne normaliziranosti. Ker je izgradnja slednjega podatkovnega skladišča relativno velik projekt, je priporočljivo slednjega graditi po delih, kajti v nasprotnem primeru lahko za uvajanje porabimo več let (Inmon, 1999c).

- Pri *distributivni arhitekturi* je področno podatkovno skladišče podmnožica podatkovnega skladišča določene organizacije. Kot lahko razberemo iz slike 3, proces izgradnje podatkovnega skladišča zajema postopno izgradnjo posameznih neodvisnih področnih skladišč, ki jih na koncu povežemo v logično podatkovno skladišče celotne organizacije. Tak pristop izgradnje imenujemo tudi pristop »od-spodaj-navzgor« (angl. bottom-up).

**Slika 3: Distributivna arhitektura podatkovnega skladišča**



Vir: Jaklič, 2003

Ker tovrstna arhitektura ne predvideva centralnega podatkovnega skladišča, temveč le unijo neodvisnih področnih skladišč, ta vsebujejo veliko količino zelo podrobnih podatkov iz daljšega obdobja in ne le agregirane podatke omejenega časovnega obdobja, kot jih predvideva centralizirana arhitektura. Osnovni podatkovni model je dimenzijski oz. večdimenzionalni. Področno podatkovno skladišče ima vlogo oddelčnega, krajevnega ali funkcionalnega podatkovnega skladišča in podpira eno ali več specifičnih področij organizacije. Vsako področno skladišče mora biti predstavljeno z dimenzijskim modelom, ki je znotraj enotnega podatkovnega skladišča skladno. Skladna dimenzija je dimenzija, za katero je značilno, da ima enoličen pomen, ne glede na to, s katero tabelo dejstev jo povežemo. Brez upoštevanja koncepta skladnih dimenzij podatkovno skladišče ne more delovati kot integrirana celota.

- *Federativna arhitektura* podatkovnega skladišča je hibridna rešitev zgornjih arhitektur, ki temelji na uporabi skupnega poslovnega modela (angl. common business model) in področjih priprave informacij (angl. information staging areas), ki so v skupni rabi. Osrednje podatkovno skladišče tako vsebuje poljubno število neodvisnih področnih skladišč. Ključni element podatkovne integracije med podatkovnim in področnimi skladišči je skupni poslovni model poslovnih informacij, hranjen in upravljan s strani podatkovnega skladišča. Izdelava skupnega poslovnega modela tako zagotavlja konsistenco pri uporabi imen podatkov in poslovnih definicij v vseh procesih podatkovnega skladišča.

## 3.2. Večdimenzionalno modeliranje podatkov

### 3.2.1. Zakaj večdimenzionalno modeliranje podatkov?

Z razmahom operativnih transakcijskih baz podatkov in v iskanju poti po čim večji učinkovitosti zajemanja transakcij smo se precej oddaljili od enega izmed temeljnih ciljev baz podatkov t.j. uporabe podatkov. Zaradi kompleksnosti transakcijskih podatkovnih modelov, ki praviloma temeljijo na postopkih normalizacije, jih končni uporabniki (z izjemo administratorjev) ne morejo razumeti, kaj šele, da bi ti samostojno izvajali poizvedbe v njih. Ob spoznanju, da si končni uporabniki s tako kompleksnimi modeli ne morejo dosti pomagati, so se snovalci podatkovnih modelov začeli zatekati k enostavnejšim rešitvam, ki so pripeljale do enostavnih podatkovnih modelov, ki jih oblikujemo s tehniko večdimenzionalnega modeliranja (Remškar, 2002, str. 18). Večdimenzionalno modeliranje podatkov lahko tako opredelimo kot novo ime za staro tehnologijo pri kateri pripravimo bazo podatkov enostavno in razumljivo (Kimball, 2002, str. 13).

Čeprav je večdimenzionalna predstavitev podatkov sinonim za večdimenzionalne baze podatkov, moramo opozoriti, da tehnike večdimenzionalnega modeliranja podatkov uporabljamo prav tako pri klasičnih relacijskih bazah podatkov. Torej opravka imamo z istimi tehnikami modeliranja, vendar z različno fizično izvedbo teh. V kolikor je predstavitev večdimenzionalnega podatkovnega modela zasnovana na relacijskih bazah govorimo o t.i. »zvezdni shemi« (angl. Star schema), če pa je ta zasnovana na večdimenzionalnih bazah podatkov ali OLAP tehnologiji govorimo, o t.i. »kockah« (angl. Cube) (Kimball, 2002, str. 13).

Kot pglavitne značilnosti večdimenzionalnega podatkovnega modela lahko opredelimo:

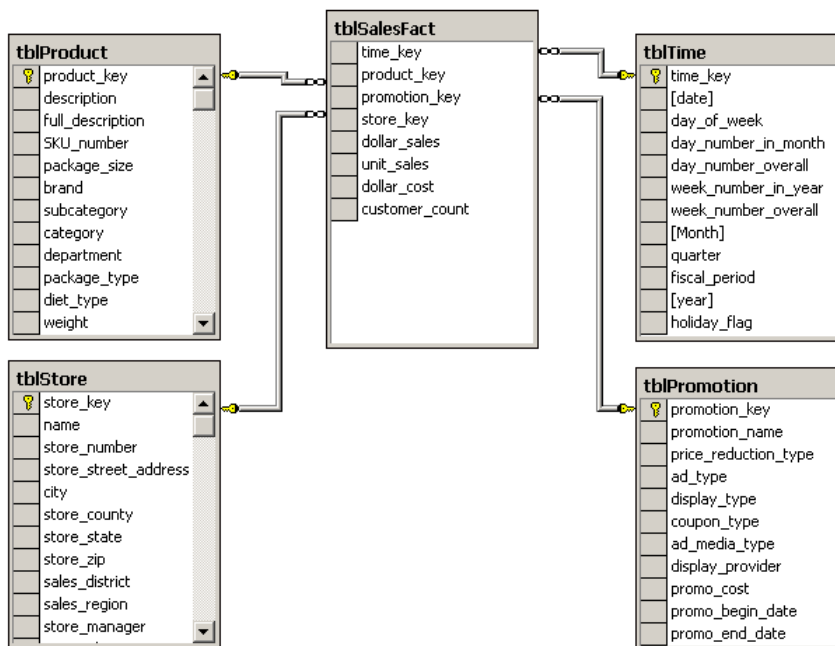
- razumljivost, preglednost, enostavnost uporabe,
- optimiziranost za poizvedovanje (ni težnje po popolni odpravi podvajanja podatkov),
- hierarhičnost (omogoča vrtanje v globino, ...),
- izpeljani podatki (podatki, ki se izračunajo iz podatkov na operativnem nivoju, niso pa vsebovani v transakcijski bazi),
- prilagojenost za analize (blizu poslovnemu pogledu na podatke),
- itd.

Skupno vsem značilnostim pa je, da so vse usmerjene k končnemu uporabniku in njegovemu načinu uporabe podatkov. Prav zaradi slednjega, dobiva večdimenzionalno modeliranje podatkov vedno večji pomen pri podatkovnih skladiščih in orodjih za sprotno analitično obdelavo podatkov. Kajti uporabniki lahko podatkovno skladišče učinkovito uporabljajo le, če natančno razumejo strukturo in relacije med podatki. Namen večdimenzionalnega modeliranja je tako približati podatkovni model uporabnikovemu načinu razmišljanja in mu omogočiti hitro in učinkovito delo s podatki. Po drugi strani pa je mogoče poizvedbe, ki temeljijo večdimenzionalnih modelih, bolje optimizirati in zato hitreje izvesti kot tiste, ki jih izvajamo nad poljubnimi entitetami s poljubnimi odnosi med njimi (Ferle, 2001a).

### 3.2.2. Osnove večdimenzionalnega modeliranja podatkov

Značilni večdimenzionalni podatkovni modeli imajo obliko zvezde, pri čemer je v sredini tabela mer ali dejstev (angl. Facts), obkrožajo pa jo vsaj ena oziroma običajno več dimenzij (angl. Dimensions). Dimenzije predstavljajo kategorije po katerih analiziramo podatke, med tem, ko mere predstavljajo vrednost ali velikost podatka, ki ga analiziramo. Na spodnji sliki 4 si lahko podrobneje pogledamo ogledamo zvezdasto shemo.

Slika 4: Zvezdasta shema večdimenzionalnega podatkovnega modela



Kot lahko razberemo iz zgornje slike 4 je osnova večdimenzionalnega podatkovnega modela tabela dejstev. Zapisi v tabeli dejstev poleg ključev za vse dimenzije vsebujejo posamezne mere (kot so zneski, velikosti, števila ...), ki jih uporabniki izpisujejo ali pregledujejo. Ker uporabniki pregledujejo mere glede na poljubne dimenzije na poljubnih ravneh hierarhije, je zaželeno, da so te vrednosti numerične in aditivne. Kajti le ob takih vrednostih so omogočene poizvedbe, ki iz številnih zapisov pripravijo različne izračune ter omogočajo njihove primerjave. To je še zlasti pomembno, kadar uporabniki obdelujejo zelo veliko število podatkov, ob tem pa ne želijo pregledovati posameznih zapisov, ampak jih zanimajo samo agregirane količine (Ferle, 2001a).

Večdimenzionalni podatkovni modeli pa lahko vsebujejo tudi t.i. neaditivna ali semiaditivna dejstva ali mere. Semiaditivne mere so tiste, pri katerih seštevanje daje smiselne rezultate samo v določenih dimenzijah, v drugih pa ne. Medtem ko nad neaditivnimi merami lahko izvajamo le operacije kot so štetje ali povprečje. Pri posameznih večdimenzionalnih modelih se lahko prav tako srečamo s tekstualnimi merami (napr. deskriptivne vrednosti). Čeprav je teoretično možno uporaba tudi teh, je priporočljiveje te vrednosti uporabiti kot attribute posamezne dimenzije, seveda če te ne odražajo unikatnosti posameznega zapisa (Kimball, 2002, str. 18).

V večdimenzionalnem podatkovnem modelu so dimenzijske tabele praviloma povezane s tabelo dejstev preko primarnega ključa do tujega ključa, v relaciji ena proti mnogo. Tabele dejstev tako običajno nimajo posebej polja za glavni ključ, ampak je le-ta sestavljen iz vseh tujih ključev. Poenostavljeno bi lahko dejali, da vsaka tabela, ki v večdimenzionalnem podatkovnem modelu odraža relacijo mnogo proti mnogo med dimenzijskimi tabelami je tabela dejstev (Kimball, 2002, str. 19).

Dimenzijske tabele vsebujejo opisne tekstualne podatke, ki pojasnjujejo mere ali dejstva. Med tem ko je za tabelo dejstev značilna globina v smislu vrstic, je za dimenzijske tabele značilno, da vsebujejo številne attribute. Ker nam dimenzijski atributi pri izdelavi poročil ali analiz služijo kot kriteriji po katerih pregledujemo podatke, ne smemo skopariti s časom, ki ga namenimo pri postavitvi dimenzijski tabel. Z drugimi besedami povedano, robustni dimenzijski atributi vplivajo na robustnost analize. Dimenzijski atributi imajo tako vlogo, da predstavijo podatke na jasen in pregleden način (Kimball, 2002, str. 20).

Število dimenzij je v splošnem lahko neomejeno, vendar je v praksi težko razumeti modele, ki imajo preveliko število dimenzij, zato je za praktično uporabo dovolj manjše, vendar obvladljivo število dimenzij v podatkovnem modelu (Ferle, 2001a). Tovrstno tabelo dejstev zaradi njenega videza imenujemo tudi »stonoga tabela dejstev«. Zato v kolikor ima večdimenzionalni podatkovni model več kot 25 dimenzij, je to lahko že znak, da posamezne dimenzije niso samostojne oz. da smo zašli v past postavljanja prevelikega števila dimenzij (Kimball 2002, str. 58).

Dimenzije mnogokrat vsebujejo hierarhije, glede na katere je mogoče podatke sešteti ali kako drugače združevati. V običajnih večdimenzionalnih podatkovnih modelih bi naj bila globina hierarhij v dimenziji enakomerna, podatki v tabeli mer ali dejstev pa bi se morali vsi nanašati na najnižjo raven hierarhije (Ferle, 2001a). Postavljanje hierarhije znotraj posamezne dimenzijske tabele vpliva tako na redundanco podatkov oz. nenormaliziranost dimenzijskih tabel. Večina pomislekov v zvezi z zvezdno shemo se tako nanaša na porabo prostora, ki ga zavzemajo ponavljajoči se tekstualni podatki. Zato nekateri uporabljajo »snežinkasto shemo« (angl. snow-flake shema), za katero je značilno, da so posamezne dimenzijske tabele normalizirane (oz. razdeljene v več tabel glede na posamezne nivoje). Posodabljanje tovrstnih dimenzij je res-da olajšano vendar se rezultat teh v večini kaže v razvejani strukturi podatkov, kar se odraža na preglednosti in optimizaciji poizvedb.

Za zvezdasto shemo je tako značilna izredna preprostost in optimiziranost za učinkovito poizvedovanje podatkov, kajti z zmanjševanjem normalizacije dimenzijske tabele se poizvedbe izvajajo na ožjem predelu podatkovne baze, kar povečuje hitrost obdelovanja podatkov. Dodaten prostor, ki ga zasedejo nenormalizirane dimenzijske tabele, pa je dejansko zelo majhen v primerjavi s prostorom, ki ga zasedejo tabele dejstev, ki običajno zavzamejo 90% ali več celote porabljenega prostora (Kimball, 2002, str. 21-22).

### 3.2.3. Različne predstave o večdimenzionalnem modeliranju podatkov

Teorija zagovornikov neodvisnih področnih podatkovnih skladišč oz. distributivne arhitekture podatkovnih skladišč pravi, da lahko uvedeš eno ali več področnih podatkovnih skladišč, jih integriraš, potem pa, ko dosežejo določene mere oz. količino podatkov, pa jih enostavno pretvoriš v podatkovno skladišče.

Torej namesto, da bi se pri projektu podatkovnega skladišča obremenjevali s kompleksnostjo vseh poslovnih funkcij in procesov, ki jih te funkcije zajemajo, ter iskali ključne elemente na ravni celotnega poslovnega sistema, se osredotočimo samo na določen del. Bistvo tovrstnega pristopa je ciljna usmerjenost. Načrtovalci si zastavijo ozek cilj, ki je relativno laže dosegljiv in daje dobre rezultate že v kratkem času (Jug, 1999, str. 33). Dandanes se podjetja nagibajo predvsem k temu pristopu, saj je bolj vabljivo, če je mogoče začeti s povsem skromnimi načrti in potem rasti (seveda če je to potrebno), ob tem pa so rešitve stroškovno ugodnejše in zahtevajo manj časa, rezultati pa so največkrat vidni že v nekaj mesecih (Vagaja, 2003, str. 21).

Čeprav se zdi omenjeni pristop zelo vabljiv, so tudi kritike upravičene. Strokovnjak Bill Inmon je glede navedene teorije mnenja, da je tovrstno razmišljanje predvsem zmotno, saj je bistvo področnih podatkovnih skladišč, da uporabniki te uporabljajo za svoje specifične potrebe in ne da te povezujejo (Inmon, 1999b). Ob tem poudarja, da so večdimenzionalne baze podatkov del tehnologije medtem ko je podatkovno skladišče arhitekturna infrastruktura. Čeprav se slednji dopolnjujeta in imata nekakšno simbiotično povezavo, še to po njegovem ne pomeni, da področna podatkovna skladišča ne potrebujejo centralno podatkovno skladišče (Inmon, 1996, str. 177). Posledice, ki lahko pripeljejo pri takšnem načinu razmišljanja so (Inmon, 1999b): ogromno podvajanje podrobnih podatkov v področnih skladiščih; nekonsistentni in nepravilni rezultati po posameznih področnih skladiščih; in neobvladljivi vmesniki med področnimi podatkovnimi skladišči in aplikacijami.

Osnovne značilnosti, ki ločujejo področno podatkovno skladišče od centralnega podatkovnega skladišča, so (Inmon, 1999b):

- podatkovno skladišče vsebuje razen granularnih podatkov tudi podatke poslovanja celotne organizacije, področno pa le podatke delovanja posameznega segmenta;
- struktura podatkovnega skladišča je normalizirana, med tem, ko so strukture področnih podatkovnih skladišč dimenzijske, ki nudi fleksibilnost pri analizi vendar niso prilagojene za večje količine podatkov;
- strukture podatkovnega skladišča so namenjene neznani uporabi, medtem ko strukture področnega skladišča so načrtovane za specifične, za znane namene;
- podatkovno skladišče vsebuje veliko količino zelo podrobnih podatkov iz daljšega obdobja, področno pa le agregirane podatke omejenega časovnega obdobja;
- različni tipi uporabnikov (pri podatkovnem skladišču so uporabniki bolj raziskovalno usmerjeni).

V nadaljevanju si pogledjmo nekatere lastnosti oz. značilnosti večdimenzionalnih podatkovnih modelov in področnih podatkovnih skladišč, za katere zagovorniki neodvisnih področnih podatkovnih skladišč menijo, da so to napačne predstave oz. miti o večdimenzionalnih podatkovnih modelih in področnih podatkovnih skladiščih (Kimball, 2002, str. 24-26):

- *Večdimenzionalni modeli podatkov in področna podatkovna skladišča so samo za agregirane podatke.* Ker ne moremo predvideti vseh vprašanj, ki si jih bodo zastavljali uporabniki, jim moramo nuditi poizvedovalni dostop do najbolj podrobnih podatkov. Dostop do podrobnih podatkov uporabnikom omogoča, da odvijajo podatke glede na postavljeno vprašanje. Tako naj agregirani podatki le dopolnjujejo podrobne podatke izključno v zagotavljanju izboljšanih učinkov poizvedovanja in naj ti ne služijo kot njihova zamenjava. Upoštevanje tega mita je najpogostejši vzrok za nekatere slabo postavljene večdimenzionalne podatkovne modele.
- *Večdimenzionalni modeli podatkov in področna podatkovna skladišča so rešitve samo za oddelke in ne za celotno organizacijo.* Zagovorniki normaliziranih podatkovnih skladišč pogosto to trditev podkrepijo z izrisom t.i. okolja pajkove mreže (angl. spiderweb environment) in ob tem opozarjajo pred potencialnimi nevarnostmi načrtovanja podatkovnega skladišča brez centralnega normaliziranega skladišča, ki predstavlja vir odvisnih področnih skladišč. Ob svojih trditvah pa ne upoštevajo, da so področna podatkovna skladišča organizirana okrog poslovnih procesov in ne oddelkov. Večina poslovnih funkcij lahko tako analizira iste mere iz enega poslovnega procesa.
- *Večdimenzionalni modeli podatkov in področna podatkovna skladišča niso primerna za velike količine podatkov.* Ideja, da področna podatkovna skladišča niso primerna za velike količine podatkov izvira iz začetkov večdimenzionalnega modeliranja podatkov in ne velja za trenutno resničnost, saj današnje tabele dejstev vsebujejo na milijarde vrstic. Kolikšna naj bo količina shranjenih zgodovinskih podatkov v področnem podatkovnem skladišču je tako odvisno samo od poslovnih potreb.
- *Večdimenzionalni modeli podatkov in področna podatkovna skladišča so primerni samo za znano oz. predvidljivo uporabo podatkov.* Večdimenzionalni podatkovni modeli in področna podatkovna skladišča so zelo fleksibilna in prilagodljiva na spremembe. Skrivnost fleksibilnosti izhaja iz tabele dejstev, katera mora vsebovati atomarne oz. podrobne podatke in ne le agregirane. Problematičnost fleksibilnosti, se tako odraža samo v primerih, ko je načrtovalec zmotno razmišljal o prvem mitu in sedaj želi nastaviti nove dimenzije, attribute, ali dejstva iz prežgodaj agregiranih tabel.
- *Večdimenzionalni modeli podatkov in področna podatkovna skladišča ne moremo integrirati oz. povezati.* Dimenzijski modeli in področna podatkovna skladišča lahko delujejo kot integrirana celota, če upoštevamo koncept skladnih dimenzij, zato jim ne smemo pripisovati vloge samostojnih rešitev, ki temeljijo na nekaterih neuspešnih poizkusih večdimenzionalnega modeliranja podatkov.



### 3.3. Sprotna analitična obdelava podatkov (OLAP)

#### 3.3.1. Lastnosti sistema OLAP v primerjavi s transakcijskim sistemom

Smisel podatkovnega skladišča je, da si z njegovo pomočjo zgradimo in izluščimo iz podatkov, ki jih imamo, uporabne informacije. Vendar si s samim skladiščem ne moremo kaj veliko pomagati, saj so to le na drugačen način predstavljeni podatki, ki smo jih praviloma imeli že prej. Podatkovno skladišče je tako podlaga za nadaljnje delo, iz katerega lahko z različnimi orodji in tehnikami dobimo kakovostne informacije (Povšič, 1998, str. 16).

Ko na splošno govorimo o podatkovnih skladiščih se srečujemo z dvema sistemoma in sicer:

- sistem za sprotno obdelavo transakcij (angl. Online Transactional Processing – OLTP),
- sistem za sprotno obdelavo podatkov (angl. Online Analytic Processing - OLAP).

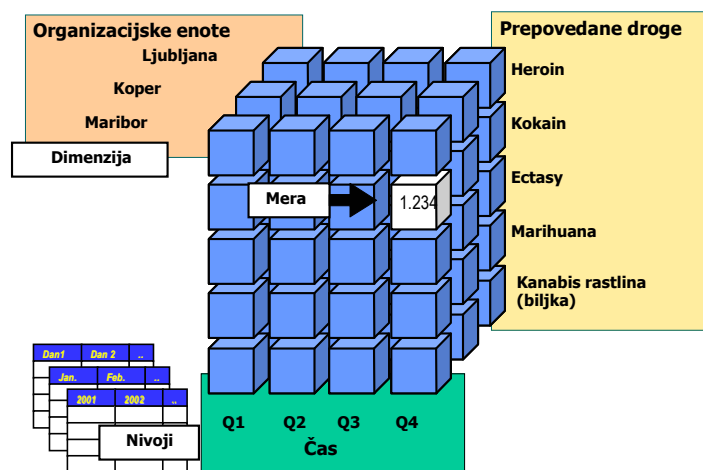
Transakcijski sistemi oz. sistemi za sprotno obdelavo podatkov so za uspešno poslovanje podjetja nujni. Ker je njihov cilj avtomatizacija poslovnih procesov se od njih zahteva: velika odzivnost, učinkovito vnašanje, minimalna možnost vnosa napačnih podatkov, hiter dostop in posodabljanje konkretnega zapisa podatkov (Vagaja, 2004, str. 22). Značilnost transakcijskih sistemov je tako operativna baza podatkov, ki je zbirka medsebojno povezanih operativnih podatkov, ki so shranjeni v računalnikovem pomnilniku brez nepotrebne podvajanja na način, ki omogoča njihovo uporabo številnim uporabnikom z različnimi zahtevami (Grad, 1996, str. 1).

Ker so transakcijskih sistemih zasnovani za obvladovanje velikega števila podatkov, ki nastajajo z vsakodnevnimi opravili v organizaciji se podatki praviloma shranjujejo v relacijski bazi podatkov v normaliziranih strukturah. Zaradi zahtev po normalizaciji je shema relacijske baze podatkov, lahko izredno razvejana in kompleksna. Težava tako nastane, ko hočemo te podatke raziskovati in analizirati, kajti prav ta razdrobljenost in normalizacija operativnih baz podatkov pomenita veliko oviro pri tovrstnih aktivnostih. Poleg tega je transakcijska baza preobsežna, saj vsebuje vse podrobnosti vsake opravljene transakcije in preobremenjena s številnimi transakcijskimi mesti, rezultat tega pa so predvsem počasne poizvedbe. Priprava tovrstnih poizvedb pa prav tako zahteva poznavanje strukture baze in specifično znanje, ki ga običajno poslovni analitiki ali drugi uporabniki nimajo (Vagaja, 2004, str. 22).

Vse te značilnosti preprečujejo, da bi bile transakcijske operativne baze podatkov primerne za iskanje kompleksnih vprašanj, ki si jih vodstvo in poslovni analitiki običajno zastavljajo. Zato se je pojavila potreba po posebni podatkovni bazi, ki je usmerjena v obvladovanje bolj poslovnih, analitičnih in raziskovalnih poizvedb. Takšno strukturo imajo podatkovna skladišča, ki so optimizirana za zajemanje in obdelovanje podatkov s pomočjo orodij OLAP. S podatkovnim skladiščem dobijo tako uporabniki neposreden dostop do podatkov, med tem ko jim uporaba OLAP orodij omogoča, da so podatki predstavljeni na njim jasen in pregleden način.

Predstavitev podatkov v sistemu OLAP je predstavljena na zelo naraven način, podobno, kot podatke pomnimo ljudje. Saj si podatke kar naprej urejamo in grupiramo v določene skupine, ki nam pomagajo pri boljši predstavi določenega pojava (Roblek, 2003, str. 323). Zelo podobno so podatki organizirani v sistemu OLAP, kjer so predstavljeni na večdimenzionalni način. Večdimenzionalno predstavitev podatkov omogočajo tako imenovane podatkovne kocke, ki so sestavljene iz posameznih dimenzij, nivojev in mer (Žorž, 2001, str. 18). Za boljšo predstavo si oglejmo spodnji primer na sliki 5, kjer prikazujemo enostavno kocko s tremi dimenzijami (organizacijska enota, prepovedane droge in čas) in mero količina zasega.

Slika 5: Prikaz podatkov v tridimenzionalni kocki



Vir: Prirejeno po Jaklič, 2004

Iz zgornje slike 5 lahko tako na enostaven način razberemo, da je bilo v zadnjem četrtletju zaseženo 1.234 g prepovedane droge kokain na območju Maribora. Dimenzije predstavljajo kategorije po katerih analiziramo podatke, medtem ko mere predstavljajo vrednost ali velikost podatka, ki ga analiziramo. Vsaka dimenzija predstavlja torej eno dimenzijo kocke in omogoča pregledovanje meritev iz različne perspektive.

Sistemi za sprotno obdelavo podatkov (v nadaljevanju OLAP) tako omogočajo analitikom in ostalim uporabnikom hiter in zanesljiv vpogled na podatke z različnih zornih kotov. Njegovo funkcionalnost opredeljuje predvsem dinamična večdimenzionalna analiza zgoščenih podatkov, pridobljenih iz podatkovnega skladišča, ki nudi končnemu uporabniku podporo pri analizi in odločanju (Babič 2002, str. 10).

Ključne značilnosti rešitve sistema OLAP so: večdimenzionalen pogled na podatke, enostavno raziskovanje in analiziranje velikih količin podatkov, njihovih povezav, zmožnosti zahtevnega izračunavanja ter inteligentnost pri obravnavi časovnega vidika podatkov. Uporabniki lahko tako brez nepotrebnega čakanja na pripravo poročil, hitro in enostavno pridobijo podatke, ki so točni in pomembni za odločanje. Proces odločanja je tako znatno hitrejši in učinkovitejši in prav to je razlog, da je OLAP postal sestavni del sistema za podporo odločanju (Babič 2002, str. 10).

### 3.3.2. Zahteve za orodja OLAP

Izbira OLAP orodja je vedno vprašanje naših potreb in zahtev, ter seveda tudi cene, ki skoraj načeloma pogojuje zmogljivosti le tega. Ker povpraševanje po OLAP orodjih narašča in cene programskih rešitev padajo, vodilni proizvajalci teh orodij na veliko promovirajo svoje izdelke, ki so še pred kratkim predstavljali ozko tržno nišo (Alahjič, 2001). Ker večina proizvajalcev oz. ponudnikov OLAP orodij, predstavlja svoje izdelke v »superlativih«, se nam zastavlja vprašanje: »Kakšna bi morala biti orodja OLAP oz. kaj bi morala le-ta omogočati?«.

Pojem OLAP je bil prvič definiran leta 1993 s strani dr. Teda Codd v raziskovalnem delu z naslovom: »Providing OLAP to User-Analysts: An IT Mandate«. Strokovnjak Codd kot znan izumitelj modela relacijskih baz podatkov in zelo cenjen raziskovalec baz podatkov, je v svojem raziskovalnem delu opredelil dvanajst pravil za opredelitev izdelkov OLAP. Žal je pa njegova opredelitev veljala kot sporna, saj je bila večina kritikov mnenja, da je mnogo funkcij OLAP orodij, definiralo podjetje Arbor Software (danes del podjetja Hyperion Solutions), ki je sponzoriralo njegovo delo (Pendse, 2004).

Ker se delo, ki je prvo definiralo pojem OLAP, smatra za publikacijo podjetja in ne neodvisno raziskovalno delo, sta strokovnjaka Nigel Pendse in Richard Creeth, (avtorja dela OLAP Report) v letu 1995 objavila t.i. FASMI test, ki opredeljuje pravila ali se posamezni izdelek uvršča med sisteme OLAP ali ne. Sistem OLAP mora tako omogočati (Pendse, 2004):

- *Hitrost* (angl. Fast). Hitrost pomeni, da mora sistem OLAP vrniti uporabniku večino rezultatov v času do petih sekund, kajti hitrost manjša od dveh sekund daje uporabniku občutek, da se podatki prikažejo v hipu. Razlagata, da je ta odziv več kot samo cilj, saj je to nujna značilnost sistema OLAP. Tako lahko le redke, kompleksne poizvedbe trajajo več kot dvajset sekund. Ker je hitrost ob veliki količini podatkov težko zagotavljati, se proizvajalci zatekajo k različnim metodam pospešitev, od posebne strojne opreme, posebno strukturiranih agregatov, do različnih načinov shranjevanja podatkov.
- *Analiza* (angl. Analysis). Sistem OLAP mora biti sposoben zajeti kakršnokoli poslovno logiko in statistično analizo, ki jo zahteva nek uporabnik, pri tem pa mora ostati čimbolj enostaven za uporabnika. Čeprav je potrebno programirati določene funkcionalnosti, morajo biti sistemi OLAP narejeni tako, da različne analize lahko nastavi uporabnik sam, brez potrebe po zahtevnem znanju programiranja. Ob tem postavlja zahtevo po intuitivni funkcionalnosti uporabe analitičnih orodij za končnega uporabnika.
- *Skupna raba* (angl. Shared). Sistem OLAP mora zagotoviti varovanje določenih podatkov (zaželeno do posamezne celice). Ker so proizvajalci predvidevali, da bodo podatki v sistemih OLAP namenjeni le za branje, so zato vgrajevali le enostavne načine zaščite. Sistem OLAP mora tako vsebovati vse potrebne varnostne mehanizme za skupinsko analitično delo in omogočati hkratno posodabljanje podatkov na zaščiten način.

- *Večdimenzionalnost* (angl. Multidimensional). Če bi opisovali sistem OLAP z eno besedo bi to bila prav gotovo večdimenzionalnost. Sistem OLAP mora zagotavljati koncept večdimenzionalnega pregledovanja podatkov vključno s podporo hierarhične ureditve podatkov, ob tem se ne predpisuje število dimenzij, ki jih mora sistem podpirati. Prav tako se ne predpisuje tehnologija shranjevanja podatkov (oz. ali naj bo struktura podatkov večdimenzionalna ali relacijska) pravilo je le, da mora sistem OLAP zagotavljati koncept večdimenzionalnega pregleda podatkov za uporabnika.
- *Informacije* (angl. Information). Pri tem pogoju se meri sposobnost posameznih izdelkov v smislu: koliko količine izvornih podatkov so ti sposobni obdelati, in ne kapaciteta podatkov, ki so jo ti sposobni shraniti. Razlike med največjimi OLAP izdelki in majhnimi po sposobnosti shranjevanja podatkov so lahko zelo velike. Pri ugotavljanju te zahteve moramo upoštevati: podvajanje podatkov, potreben spomin, izkoristek trdega diska, storilnost, integracija s skladišči podatkov in podobno.

### **3.3.3. Primerjava sistema OLAP z rešitvami podatkovnega rudarjenja**

Področja uporabe orodij OLAP in orodij za podatkovno rudarjenje na osnovi podatkovnih skladišč so zelo raznolika. Teoretično je njihova uporaba možna povsod, kjer je na voljo veliko podatkov in so analize teh podatkov zelo pomembno poslovno orodje za načrtovanje oziroma sprejemanje odločitev. Vendar se moramo ob tem zavedati, da se ob veliki množici podatkov, zmanjšuje naša sposobnost njihovega razumevanja in odkrivanja informacij, ki se skrivajo v njih.

Uporaba podatkovnih skladišč in OLAP orodij je vsekakor podatke naredila preglednejše in jasnejše, vendar je zaradi velike količine podatkov delo analitikov zelo težavno, počasno in mnogokrat tudi neuspešno. Vloga orodij za podatkovno rudarjenje (angl. Data mining) je tako, da namesto analitika samodejno analizirajo podatke in prikažejo rezultate. Orodja za podatkovno rudarjenje so torej kot nekakšna alternativa pri iskanju informacij, kjer so količine vhodnih podatkov izjemno velike ter informacije skrite in težko dosegljive.

Glede na opisano vlogo podatkovnega rudarjenja, lahko podatkovno rudarjenje opredelimo kot avtomatizirano iskanje veljavnih, prej nepoznanih vzorcev in povezav v velikih količinah podatkov, na podlagi katerih lahko ukrepamo, z namenom boljših odločitev (Jaklič, 2004). Poenostavljeno bi lahko tako dejali, da podatkovno rudarjenje zajema pridobivanje koristnih informacij iz velikih količin podatkov. Medtem ko mora uporabnik pri uporabi OLAP orodjih vedeti in razumeti, kako naj išče podatke in informacije, ki ga zanimajo, orodja podatkovnega rudarjenja omogočajo avtomatsko iskanje skritih trendov in vzorcev v velikih količinah podatkov. Pri tem nas ne sme zavesti beseda »avtomatizirano«, saj je še vedno potrebno poznati probleme oz. področje delovanja, razumeti podatke in poznati analitične metode. Orodja za podatkovno rudarjenje so kot taka lahko le zgolj v pomoč analitiku oz. uporabniku, saj sama ne povedo kakšna je za organizacijo vrednost odkritih informacij. Za takšne ocene so še vedno potrebni ljudje (Berry et al., 2000, str. 8; po Konič, 2003, str. 7).

Čeprav se v orodjih za podatkovno rudarjenje uporabljajo številni algoritmi, ki so znani s področja matematike, statistike, umetne inteligence in drugih je pomembno opozoriti, da se podatkovno rudarjenje razlikuje od statističnih analiz, kajti: »Pri statistični analizi nikoli ne najdeš tistega, česar ne iščeš.« Bistvena razlika je ta, da se pri statistiki za vsak problem predpostavi domneve in se nekako pričakuje rezultate, medtem ko gre pri podatkovnem rudarjenju za avtomatizirano iskanje vzorcev in pravil, ki se jih vnaprej ne da predvideti (Jaklič, 2004). Temeljna razlika med orodji OLAP in orodji podatkovnega rudarjenja je v njihovih rezultatih, kajti med tem, ko so rezultati pri OLAP orodjih usmerjeni v preteklost, je temeljna značilnost orodij podatkovnega rudarjenja njihova sposobnost napovedovanja.

Za boljše razumevanje uporabe orodij podatkovnega rudarjenja, si pogledjmo različno zastavljena vprašanja ali zahteve s katerimi so soočeni analitiki s strani vodstva pri vsakdanjem delu (Berson et al., 2000, str. 34): »Napovej in obrazloži povpraševanje v naslednjem mesecu«, »Obrazloži zakaj nekateri kupci prehajajo k konkurenci«, »Ugotovi nove vzorce obnašanja naših kupcev, katerih še ne poznamo«.

Čeprav uporaba podatkovnega rudarjenja ni specifično povezana z nobenim področjem, se je njegova uporaba razvila predvsem na področjih, kjer obstaja velika količina podatkov in kjer se iz podatkov »splača« naučiti nekaj novega. V splošnem lahko rečemo, da se pridobivanje novih znanj spleta toliko časa, dokler je vrednost teh znanj višja od stroškov njihovega pridobivanja. Podobno velja za uporabo podatkovnega rudarjenja na poslovnem področju, le da je pravilo, kdaj se nekaj spleta, nekoliko strožje (Berry et al., 2000, str. 11; po Končič, 2003, str. 7). S pomočjo podatkovnega rudarjenja lahko tako podjetje odkriva profile strank, ki mu služijo za boljše razumevanje njihovih potreb in donosnosti, in za oblikovanje modelov predvidevanja obnašanja strank z namenom izboljševanja trženjskih in drugih poslovnih aktivnosti podjetja.

Podatkovno rudarjenje uporablja torej drugačen koncept od orodij OLAP, čeprav za doseg enakih ciljev, ki so usmerjeni k višjemu dobičku z zmanjšanjem stroškov ali z zvišanjem prihodkov organizacije. Večina organizacij uporablja podatkovno rudarjenje v namene (Berson et al., 2000, str. 34):

- *Odkrivanje znanja v podatkih.* Cilj odkrivanja znanja je odkrivanje skritih povezav in vzorcev v podatkovnih bazah.
- *Vizualizacija podatkov.* Analitik mora razumeti velike količine podatkov, ki so shranjene v podatkovnih bazah, zato je prioriteten cilj pred vsako analizo prosvetliti maso podatkov s katero imajo opravka in najti učinkovitejše načine predstavitve podatkov.
- *Popravki podatkov.* Ko združujemo podatke iz različnih operativnih baz podatkov, velikokrat ugotovimo, da so vneseni podatki nepopolni, nepravilni in vsebujejo nasprotujoče si informacije. Pri odkrivanju tovrstnih nepravilnosti in njihovem sprotnem popravljanju nam lahko veliko pripomore uporaba podatkovnega rudarjenja.

Sposobnost predvidevanje podatkovnega rudarjenja vsekakor ni čarovnija. Znan pregovor pravi: »Tisti, ki se ne naučijo na napakah preteklosti so obsojeni, da te ponavljajo«. Podatkovno rudarjenje tako uporablja zgodovinske podatke na katerih se uči in predvideva bodoče stanje. Za podatkovno rudarjenje se je prav tako uveljavil izraz »iskanje zlata v gori podatkov«, vendar če želimo odkriti zlato moramo predhodno definirati kaj je zlato je, oz. kateri poslovni problem bi želeli rešiti (Berson et al., 2000, str. 96).

Najpogosteje uporabljene tehnike podatkovnega rudarjenja so (Jaklič, 2004): nevronske mreže, drevesa odločitev, inducirana oz. odločitvena pravila, najbližji sosed, homogenost, genetski algoritmi itn. V splošnem pa lahko tehnike podatkovnega rudarjenja delimo na dve širši kategoriji (Rapaić Boštjančič, 2004, str. 45):

- razkrivajoče tehnike rudarjenja podatkov, so takšne, ki iščejo vzorce v podatkih brez prejšnjega poznavanja, kakšni vzorci v podatkih sploh obstajajo. Pri tem poznamo več tehnik, nekatere izmed njih so: grupiranje; analiziranje povezav; analiziranje pogostosti pojavljanja;
- in napovedne tehnike rudarjenja podatkov, s katerimi iščemo povezave med določenimi spremenljivkami in ostalimi spremenljivkami v podatkih.

Pri podatkovnem rudarjenju lahko prav tako govorimo o dveh različnih pristopih (Berry et al., 2000, str. 8; po: Končič, 2003, str. 4):

- usmerjeno podatkovno rudarjenje, kjer vemo, kaj točno iščemo. Najpogostejša oblika tega pristopa je izdelava napovedovalnih modelov. Ciljna (odvisna) spremenljivka je vnaprej določena in model mora kar se da dobro prepoznati povezave med njo in ostalimi podatki;
- neusmerjeno podatkovno rudarjenje, kjer naš cilj ni določen. V tem primeru podatkovno rudarjenje odkrije vzorce in nam samim prepusti odločitev o njihovi (ne)pomembnosti.

V kolikor želimo natančen model podatkovnega rudarjenja, moramo jasno definirati problem, ki ga želimo rešiti, imeti na razpolago bogato zbirko podatkov, ki se nanaša na obravnavano problematiko, razumeti jedro podatkov in razpolagati z ustreznimi merami, s katerimi lahko ocenjujemo točnost podatkovnega modela (Paul et al., 2002, str. 31). Torej ne-glede na izbran pristop ali tehniko rudarjenja, se pri vpeljavi rešitev podatkovnega rudarjenja srečujemo s šestimi koraki (Paul et al, 2002, str. 9):

- definiranje problema,
- priprava podatkov,
- gradnja podatkovnih modelov,
- preizkušanje podatkovnih modelov,
- vpeljava modela,
- upravljanje z meta podatki povezanimi s transformacijo, čiščenjem podatkov, gradnjo modelov in preizkušanjem modelov.

## 4. DELOVANJE POLICIJE NA PODROČJU PREPOVEDANIH DROG

### 4.1. Obravnava prepovedanih drog s kazensko pravnega vidika

Slovenija sodi med države, ki kot kaznivo dejanje ne inkriminira samega uživanja prepovedanih drog temveč različna dejanja glede pridelave, predelave, prodaje ali omogočanja uživanja prepovedanih drog. Kazenski zakonik Republike Slovenije (v nadaljevanju: KZ) tako v 196. členu določa kot kaznivo dejanje neupravičene proizvodnje in prometa z mamili, ki je opredeljeno v svoji temeljni obliki, v kvalificirani obliki pa kadar je dejanje storjeno v sotorilstvu ali v organizirani mreži prekupčevalcev ali posrednikov oziroma kot pomoč pri tem kaznivem dejanju. Kot samostojno kaznivo dejanje je v KZ določeno tudi omogočanje uživanja mamil v smislu 197. člena.

Poleg teh kaznivih dejanj so z zakonikom določena še nekatera kazniva dejanja, katerih namen je prav tako preprečevanje prometa z mamili. To so: kaznivo dejanje proizvodnje in prometa škodljivih sredstev za zdravljenje 193. čl. KZ, kaznivo dejanje pranja denarja v smislu 252. čl. KZ, in tihotapstvo po 255. čl. KZ.

Presojanje tovrstnih kaznivih dejanj je v sodni praksi dokaj zapleteno. V obrambi obdolženca prevladujejo predvsem trije zagovori, da je (Dežman, 2001, 24):

- droga podtaknjena,
- za lastno uporabo,
- ravnal v zmoti in ni vedel, da gre za mamila,
- in najpogostejši zagovor, da je mamilo za lastno uporabo. Dokazna ocena zagovora je toliko večja, kolikor manjša je količina, ker je toliko bolj sprejemljiv zagovor, da je obdolženec imel v posesti mamila le za lastno uporabo.

Poseben problem predstavlja organizirani kriminal, torej pregon kaznivih dejanj storjenih v okviru organizirane kriminalne združbe z notranjimi pravili ravnanja, ki deluje na podjetniški način in pri tem praviloma uporablja nasilje oz. korupcijo, ter z namenom pridobivanja protipravne premoženjske koristi ali družbene moči (Dežman, 2001, 24). Prav zaradi specifičnosti izvrševanja kaznivih dejanj v organizirani skupini je zakonodajalec v Zakonu o kazenskem postopku (v nadaljevanju ZKP) uvedel vrsto ukrepov, imenovanih posebne metode ali pogosteje imenovane ukrepe v smislu 150. člena ZKP. Za te ukrepe je značilno, da pomenijo globok poseg v človekove pravice in svoboščine. Zato se lahko uporabljajo le pod strogo določenimi pravnimi pogoji in je njihova uporaba opravičena samo v primerih, ko je mogoče utemeljeno sklepati, da se z drugimi ukrepi ne bi dalo zbrati dokazov oziroma bi njihovo zbiranje lahko ogrozilo življenje ali zdravje ljudi.

Da bi bilo lahko pravosodje v boju zoper organizirano kriminaliteto bolj učinkovito, je zakonodajalec prav tako določil, da se sme storilcu kaznivega dejanja, ki prepreči nadaljnje izvrševanje teh dejanj, ali razkrije podatke, ki so pomembni za preiskovanje in dokazovanje že storjenih kaznivih dejanj, kazen za ta dejanja omiliti (tretji odstavek 297. člena KZ).

## 4.2. Obravnava prepovedanih drog s policijskega vidika

Slovenski strokovnjaki imajo v zvezi s problematiko uporabe prepovedanih drog več različnih pogledov. Od skrajno liberalnih, kjer so mnenja, da je uporaba prepovedanih drog del vsakdanjega življenja in jo je tako treba tudi obravnavati, do rigoloznih, kateri menijo, da je uporaba prepovedanih drog škodljiva ne samo posamezniku temveč tudi družbi kot celoti ter jo je zato treba prepovedati in ustrezno sankcionirati. Ne oziraje se na mnenja strokovne javnosti je delovanje policije na področju prepovedanih drog usmerjeno predvsem v odkrivanje, preprečevanje in preiskovanje kaznivih dejanj, ki zadevajo tovrstno problematiko.

Kriminaliteto povezano z mamili, lahko razdelimo na tri vsebinske sklope in sicer na kriminaliteto, povezano s tihotapljenjem (mednarodna trgovina), kriminaliteto, ki zajame druge oblike kriminala (nasilje, podkupovanje, pranje denarja) in kriminaliteto uživalcev mamil (vlomi v lekarne, ropi, tatvine). Skupne značilnosti te kriminalitete so predvsem naslednje (Brejc et. al., 1993, 20):

- ni klasičnega oškodovanca, ki bi sicer lahko prijavil kaznivo dejanje,
- organizirana mednarodna trgovina z mamili prinaša velike dobičke,
- internacionalnost in dobra organiziranost tihotapcev,
- pojavljanje vedno novih načinov proizvodnje, predelave, prenosa in razpečevanja mamil.

Gledano s strani ekonomije imajo prepovedane droge številne ekonomske lastnosti (Novak 2001, 25):

- enostavna in poceni pridelava (gojenje in predelava),
- enostaven in poceni prevoz, saj niso pokvarljive in krhke,
- visoka vrednost na količino,
- ogromno in dobičkonosno povpraševanje,
- so skorajda neomejene,
- proizvodnjo in preprodajo je enostavno skriti,
- delovanje preprodajalcev na tržišču kupcev.

Po mnenju nekaterih poznavalcev znaša letni dobiček z ulično prodajo mamil in psihotropnih snovi več kot 500 milijard ameriških dolarjev ali skoraj desetino celotne svetovne trgovine. Zaradi dobičkonosnosti te dejavnosti so t.i. »mamilarski karteli« sposobni zaposlovati najkvalitetnejše strokovnjake s področij prava, financ, logistike in kemije. Uporabljajo najsodobnejšo opremo in tehnologijo za proizvodnjo, transport in distribucijo mamil ter pomoč pri pranju denarja, pridobljenega z njihovo prodajo. Največji »trgovci« so tako sposobni voditi in oskrbovati svojo dejavnost, ne da bi ob tem prišli v stik z blagom. Trgovske poti po vsem svetu so razvejane in dobro organizirane ter vodene. Prilagajajo se potrebam, preprekam in se prepletajo z mednarodno politiko, terorizmom, vojnimi žarišči. Trgovske niti se tako dotikajo mnogih spoštovanih in uglednih osebnosti (Vrančič, 1998).



Kljub utemeljenim sumom o krivdi in povezanosti z nezakonitimi posli je tako zelo težko zbrati dokaze za kazenski pregon. Za uspešno preiskovanje in preprečevanje kaznivih dejanj povezanih s prepovedanimi drogami je potrebno predvsem dobro poznavanje načina dela samih tihotapcev in preprodajalcev droge. Zato mora policija na tem področju uporabljati standardne, klasične metode preiskovanja kot tudi nove metode, ki so jih razvili v zahodnih državah, kjer je problematika prekomerne uporabe prepovedanih drog prisotna že dalj časa. Preiskovanje tovrstne kriminalitete je specifično in zahteva posebno usposobljene kriminaliste oziroma policiste, ki to problematiko spremljajo dnevno.

Kriminalistična policija je organizirana na način, da je odprta za dogajanja v družbi, kar pomeni, da iz okolja sprejema prijave kaznivih dejanj. Tesna povezanost z okoljem je posebno vidna na lokalni ravni in v procesu zbiranja informacij. Na osnovi zbiranja obvestil in preverjanja informacij, kriminalistična policija načrtuje in izvaja potrebne ukrepe za odkrivanje kaznivih dejanj. Pri tem slovenska policija ne daje svojega prispevka samo v Sloveniji temveč skuša slediti kriminalnim združbam, katerih dejavnosti ne priznavajo regionalnih, nacionalnih in kontinentalnih meja. Tako slovenska policija prispeva k globalnemu zmanjšanju ponudbe prepovedane droge na svetovnem in evropskem tržišču.

Slovenija se v mednarodni trgovini s prepovedanimi drogami pojavlja kot država porabnica prepovedanih drog, kakor tudi tranzitna država preko katere se izvajajo transporti oziroma se sklepajo posli za neupravičeno proizvodnjo in promet s prepovedanimi drogami. Če ob tem pomislimo še na policijsko statistiko, se lahko prepričamo, da je problematika prepovedanih drog z vidika kaznivih dejanj, v stalnem porastu, tako rekoč iz leta v leto od 25 do 35%.

Ob izračunu in razmišljanju, da 5.000 odvisnikov od heroina v Sloveniji potrebuje dnevno najmanj 1 kg heroina in da je to količino potrebno v Slovenijo pretihotapiti, porazdeliti in prodati v manjših primernih količinah, je povsem jasno, da je dnevno storjenih ogromno število kaznivih dejanj. Zato statistika ne predstavlja niti omembe vrednih števil. Letno se namreč ne odkrije niti toliko kaznivih dejanj in prekrškov, kot se jih zgodi v enem samem dnevu (Vrančič, 1998). Za učinkovito policijsko delovanje, se zato slovenska policija vključuje v najrazličnejše preventivne programe.

Največji delež preventivnih programov je namenjen otrokom in mladostnikom na lokalni ravni in sicer v vzgojno-izobraževalnih ustanovah, del pa tudi zunaj njih. Dejavnosti na področju preventive potekajo od začetnega odvracanja od uporabe drog pri vseh starostnih skupinah pa vse do zmanjševanja negativnih zdravstvenih in socialnih posledic uporabe drog, zdravljenja, socialne obravnave in ponovne socialne vključitve oseb, nekdam odvisnih od drog v družbo. Iz preventivnega stališča se policija angažira v obliki raznih predavanj na šolah za slušatelje in pedagoške delavce in starše, se udeležuje raznih okroglih miz in nastopa v raznih javnih medijih. V sodelovanju z drugimi institucijami policija skrbi za čim boljše osveščanje otrok in mladine o nevarnostih prepovedanih drog in nepotrebnosti poseganja po njih.

### 4.3. Vloga in pomen obveščevalnega in analitičnega dela policije<sup>1</sup>

Kriminalistična analitika (angl. crime analysis) je širok pojem, ki je bil izbran za opis analitičnega dela znotraj kriminalistične policije. V slovenskem prostoru je kriminalistična analitika še vedno neprepoznavna in kot taka izpostavljena različnim interpretacijam, ki največkrat nimajo nič skupnega s pojmom o katerem bo govora v nadaljevanju. Največkrat se analitika zamenjuje s statistiko ali pa se govori o njej na način, da se omenjajo le njeni posamezni sestavni deli. Zato trenutni trendi in priporočila Interpola in Europola gredo predvsem v smeri svetovne standardizacije definicij, metodologije in tehnik v vseh kriminalističnih policijah po svetu.

V slovenski kriminalistični policiji se uporabljajo predvsem standardi in pravila, ki sta jih postavila Interpol in Europol. Obe organizaciji namesto izraza kriminalistična analitika uporabljata izraz kriminalistično obveščevalna analitika, ki vključuje bistvo discipline (t.j. izdelava novega vedenja na osnovi obveščevalnih podatkov, ki so namenjeni odločanju) kot specifičnost načina in konteksta, v katerem policija opravlja obveščevalni proces. Izraz je prevod angleškega pojma »intelligence«, ki v bistvu pomeni informacijo, ki smo ji v analitičnem procesu dodali novo vrednost oz. gre za predelano informacijo z novim vedenjem, ki omogoča delovanje. Ker bomo v nadaljevanju uporabljali krajši izraz t.j. kriminalistična analitika, samo opomnimo, da imamo v mislih kriminalistično obveščevalno analitiko.

Obstaja veliko definicij kriminalistične analitike, od leta 1992 pa je Interpol začel uporabljati definicijo, ki govori o kriminalistični analitiki kot o identifikaciji in omogočanju vpogleda v odnose med podatki o kriminaliteti in drugimi potencialno relevantnimi podatki z vidika policijske in sodne prakse. Pri tem pod pojmom podatki o kriminaliteti razumemo podatke o kaznivih dejanjih, njihovih storilcih in žrtvah ter policijskih postopkih preiskovanja in sodnih postopkih. Kriminalistična analitika se tako ukvarja z objekti (kaznivimi dejanji), subjekti (storilci in žrtvami) in postopki (preiskovanjem kriminalitete in sodnimi postopki), ki sestavljajo celovit kompleks kriminalitete kot družbenega pojava, pri tem pa izhaja iz položaja policije in pravosodnih organov.

Kriminalistična analitika je funkcionalno razdeljena na dve področji, ki ju imenujemo: strateška analitika in operativna analitika. Funkcionalna delitev je potrebna zaradi razlik v namenu in predmetu analiz v enem in drugem sklopu analitike. Medtem ko so cilji operativnih analiz kratkoročni, usmerjeni v preiskovanje konkretnega primera in prijetja storilca, odkrivanju kriminalne združbe ali zasegu predmetov kaznivega dejanja, so strateške analize usmerjene k dolgoročnejšim ciljem in so v funkciji vodenja in odločanja. Strateška analiza tako zajema širšo problematiko in se ukvarja z nadzorom in napovedovanjem celotnega razvoja kriminala v srednje in dolgoročnem obdobju oz. ukvarja se z dolgoročnimi cilji (narava kriminala, obseg in projekcija rasti kriminala, določanje prioritet, itd.).

---

<sup>1</sup> Prirejeno po Maver, 1997, str. 97-113 in interni dokumentaciji organizacije

Kriminalistična analitika uporablja enotne tehnike in metode ter je osredotočena na razvijanje hipotez, rekonstruiranje posameznih kaznivih dejanj, identifikacijo serijskih, povezanih kaznivih dejanj, razumevanje kriminalnih mrež in analiziranje, obsega in vzorcev kriminalnih aktivnosti.

Interpol je oba tipa kriminalistične analitike (operativno in strateško analitiko), glede na kaj je analiza osredotočena razdelil v tri skupine. Tako ločijo različne analize glede na kaznivo dejanje, storilca oz. žrtev in metode nadzora kriminala, ki vam jih prikazujemo v tabeli 1.

**Tabela 1: Vrste kriminalističnih analiz in uporabljenih metod**

	STRATEŠKA	OPERATIVNA
KAZNIVO DEJANJE	ANALIZA KRIMINALNIH VZORCEV	ANALIZA PRIMERA
		PRIMERJALNA ANALIZA PRIMEROV (serijska kazniva dejanja)
OSEBA	ANALIZA GENERALNEGA PROFILA	ANALIZA KRIMINALNIH SKUPIN (znani storilci)
		ANALIZA SPECIFIČNEGA PROFILA (neznani storilci)
METODE NADZORA KRIMINALA	ANALIZA METOD KONTROLE KRIMINALITETE	ANALIZA PREISKAV

Vir: Maver, et. al., 1997, str. 113

Do vseh treh objektov analize pristopamo tako s strateškega kot z operativnega vidika. Skupno gre za osem temeljnih metod kriminalistične analitike, ki se konkretizirajo v uporabi posamičnih tehnik. Uporabljene tehnike in metode zagotavljajo standardiziran pristop k delu in hkrati ponujajo fleksibilnost, ki je omejena samo z analitikovo sposobnostjo in domišljijo.

Ko govorimo o obveščevalni dejavnosti, se zelo pogosto pojavlja stereotipno pojmovanje tega termina, kajti večina ljudi ga povezuje z obveščevalnimi službami, ki jim vselej pripisujejo negativno politično vsebino. Čeprav večina kriminalističnih policij v svetu že dolgo uporablja izraz kriminalistično obveščevalno delo, se ta termin v Sloveniji vse prevečkrat enači z obveščevalno službo s predznakom politične policije. Termin kriminalistična obveščevalna dejavnost, je brez sleherne politične vsebine, kajti pri kriminalistični obveščevalni dejavnosti gre za sistematično zbiranje podatkov in informacij zaradi zoperstavljanja kriminaliteti.

Ker ne bomo podrobneje zahajali v definicije same obveščevalne dejavnosti, moramo opomniti, da je zbiranje podatkov le del obveščevalnega procesa oz. obveščevalnega cikla, ki ima več med seboj povezanih in soodvisnih faz (kot so: dodelitev naloge, ocena zahtevka, zbiranje podatkov, vrednotenje podatkov, primerjanje podatkov, integracija podatkov, interpretacija ter razvoj hipotez in zaključkov, diseminacija). Zaključek posamezne faze sproži nadaljevanje naslednje faze, ker pa govorimo o ciklusu, pomeni zaključek zadnje faze ponovno proženje cikla. Osrednji in bistveni del obveščevalnega procesa je analiza obveščevalnih podatkov, saj se brez te obveščevalno delo lahko kar hitro sprevrže v golo zbiranje informacij. Zbiranje podatkov je tako le ena izmed faz obveščevalnega ciklusa, ki je usmerjena v pridobivanje vseh pomembnih informacij iz vseh virov, ki so na voljo. Katere vire bomo uporabili pa je odvisno od narave analize, ki jo bomo izvajali.

Osnovna naloga operativne kriminalistične analitike je zagotavljanje obveščevalnih analiz pri preiskovanju posameznih primerov. Operativne analize so tako usmerjene na identifikacijo posameznika (storilca kaznivega dejanja) ali kriminalne skupine, delovanje oz. metode, ki jih pri tem uporabljajo (modus operandi), njihove zmožnosti, omejitve, ranljivosti in nameni morebitnih storilcev ter njihove vire in podpore pri financiranju.

Zbiranje podatkov, se praviloma prične z zbiranjem podatkov o kriminalni preteklosti oseb, podatkov o njihovem premoženjskem stanju, o njihovih vozilih, zvezah in podobno. Ker se operativna analitika izvaja neposredno ob sami preiskavi, je potrebno zagotoviti neprestano zbiranje in vrednotenje informacij ves čas, dokler preiskava ni zaključena. V procesu zbiranja podatkov je tako pomembno vedenje, kaj se dogaja in kaj se pripravlja v kriminalnih krogih, kajti tako pridobljene informacije opredeljujejo vse druge aktivnosti policije. Vendar, ker organizirane kriminalne združbe delujejo konspirativno in so zavezane k molku, je zbiranje podatkov v relativno zaprtih kriminalnih združbah možno le s prisluškovanjem, sledenjem, opazovanjem, in s prodiranjem v te kriminalne skupine.

Strateška analitika je pomembna tudi zaradi možnosti, da vpliva na odločitve vodilnih delavcev pri oblikovanju dolgoročne politike kriminalistične policije na določenem področju. Viri podatkov za strateške analize so lahko že zbrani podatki znotraj policije vendar je večino informacij, ki so povezane s strateško analizo možno pridobiti le preko »odprtih virov«, zato jih je potrebno predhodno ustrezno ovrednotiti oziroma jih razvrstiti in ločiti na relevantne in nujne podatke.

Ker obstajajo določene povezave v načinu razmišljanja med podjetji in organi pregona, si pogledjmo malce prirejene značilnosti upravljanja s strankami (v našem primeru: osumljenci kaznivih dejanj) s strani organov pregona na področju prepovedanih drog:

- prepoznavna posameznikov, ki predstavljajo vrh kriminalitete in povečevanje truda, da bi bili ti ustrezno kazensko obravnavani in jim bila odvzeta proti premoženjska korist,
- koncentracija virov (finančni, materialni in dela) na kriminalne združbe, ki predstavljajo največjo ogroženost za družbo,
- uporaba različnih ukrepov, ki jih ima policija na razpolago (sledenje, opazovanje, prisluškovanje ...), da se odkrijejo in zavarujejo sledi kaznivega dejanja in predmeti, ki utegnejo biti dokaz pri kazenskem pregonu osumljencev,
- nenehno zbiranje podatkov z namenom lažjega predvidevanja in obnašanja storilcev kaznivih dejanj,
- vzpostavitev odnosa med informatorji in storilci kaznivih dejanj z namenom, da se zberejo vsa obvestila, ki bi utegnila biti koristna za izvedbo kazenskega postopka,

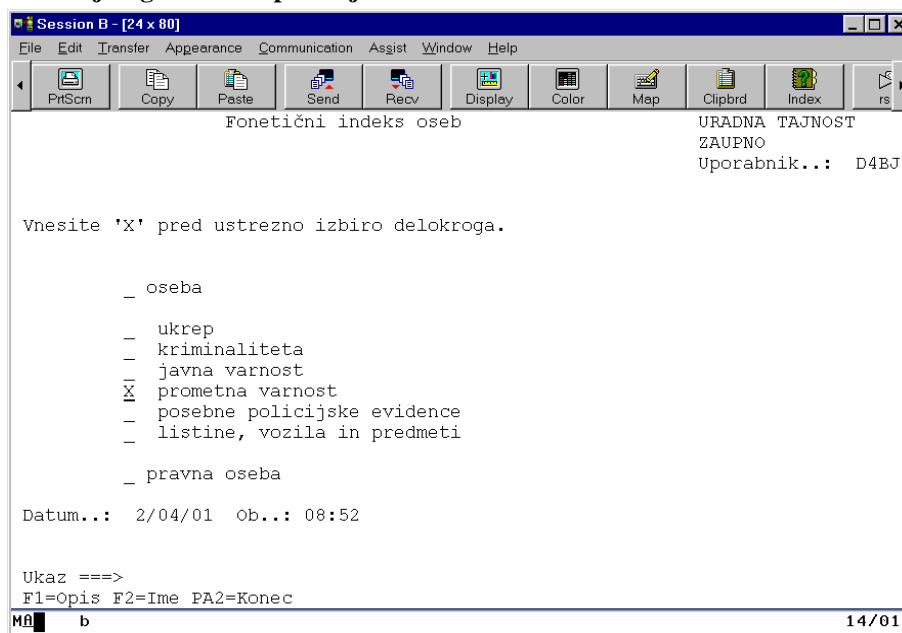
za doseganje skupnega cilja: varovanja življenja, osebne varnosti in premoženja ljudi.

#### 4.4. Sistem za sprotno obdelavo transakcij policije in njegovi segmenti<sup>2</sup>

Na Ministrstvu za notranje zadeve so že zelo zgodaj spoznali možnosti, ki jih ponuja informacijska tehnologija. Začetki mehanske obdelave podatkov namreč segajo v leto 1957, ki ga štejemo za začetek razvoja informacijsko-telekomunikacijske podpore v policiji. Prav posebno mesto v tem razvoju pa pripada Fonetičnemu indeksu oseb (v nadaljevanju besedila: FIO), katerega razvoj se je začel v letu 1992 in je do danes postopno podprl v vse segmente policijskega dela. FIO predstavlja skupek vseh policijskih, z zakoni predpisanih evidenc, ki so zgrajene okoli Indeksa enotne policijske Osebe.

Izdelan je bil po principu kooperativnega procesiranja, ki je kombiniral prednosti centraliziranega zbiranja podatkov s fleksibilnostjo obdelav v lokalnem okolju. S FIO je bila vpeljana relacijska baza podatkov DB2 in programsko okolje CSP (Cross System Product). Na spodnji sliki 7, vam prikazujemo menijski ekran FIO. Ekran lahko razdelimo na tri dele, in sicer: na glavo, kjer je navedena oznaka ekrana, naslov ekrana in stopnja zaupnosti podatkov; podnožje, ki je sestavljeno iz vrstice za sporočila, ukazne vrstice in vrstic, kjer so navedene dovoljene funkcije in funkcijske tipke; in delovno področje ekrana, kjer izvajamo posamezne funkcije.

Slika 6: Primer menijskega ekrana aplikacije FIO



Vir: Lenarčič, 2002, str. 12

Osnovni namen FIO je podpora operativnem delu policije, kar pomeni, da se vse osebe, ki jih je obravnavala policija evidentirajo v enotni bazi. Hkrati pa je postal tudi vir podatkov za med-organizacijsko povezovanje policije znotraj državne uprave in zunaj nje. Z vstopom R Slovenije v Evropsko unijo, pa postaja FIO tudi pomemben vir podatkov za SIS (Schengenski informacijski sistem), Europol in druge informacijske sisteme, ki se uporabljajo v okviru EU.

<sup>2</sup> Povzeto po Lenarčič, 2002

Večina policijskega dela se nanaša na delo z osebami, ki v dogodkih lahko nastopajo kot osumljenci, kršitelji, oškodovanci, udeleženci v prometnih nesrečah, itd. Enoznačnost osebe določa tako imenovani CR osebe (CR je okrajšava za številko iz Centralnega registra. Ime izhaja še iz časov kartotečnega evidentiranja oseb - CEK). Ta enoznačni ključ osebe je tvorjen umetno in iz njega ne moremo sklepati identiteto osebe. Vsaka oseba, ki se na novo pojavi v evidenci, pridobi nov CR. Ker oseba v različnih dogodkih lahko nastopa z različnimi podatki, imamo lahko v okviru istega CR več različnih zapisov.

Podatke, ki se nanašajo na osebe lahko razdelimo v tri osnovne skupine:

- identifikacijski podatki (priimek, ime, spol, podatki o rojstvu, vzdevki, lažna imena);
- naslovi (stalen in začasni naslov);
- značilnosti (državljanstvo, poklic, izobrazba).

Glavni segmenti FIO so osebe, dogodki in predmeti. Vsi trije segmenti so med seboj v relaciji, tako je pristop do podatkov možen preko vseh. Pristopi so načeloma enakovredni (možni so prehodi iz: osebe na njene dogodke in predmete; dogodka na osebe, ki v njem nastopajo in predmete; predmeta na dogodek ali osebo, v zvezi s katerim je evidentiran). Poglejmo si nekoliko podrobneje posamezne segmente:

- *Osebe*, ki so vezane na kazniva dejanja, so lahko fizične ali pravne, nastopajo pa v vlogi osumljenca ali oškodovanca. Med osumljenci in oškodovanci se beležijo njihovi morebitni odnosi (poznanstvo, sorodstvo, ...).
- Podatki o *zaznavah* in *kaznivih dejanjih* so organizirani po nosilnih dosjejih (dosje je zadeva v kateri se obravnava eno ali več kaznivih dejanj).
- Tretji segment so *predmeti*, ki so lahko ukradeni, najdeni, sredstvo, s katerim je bilo storjeno kaznivo dejanje ali na katerem je bila najdena sled. Predmet je lahko udeležen v dogodku (napr.: vozila v prometnih nesrečah) ali pa pričakujemo, da ga bo oseba uporabila pri identifikaciji (napr.: osebna izkaznica). V kaznivih dejanjih lahko nastopajo predmeti, ki so vezani na kaznivo dejanje (napr.: predmet najden na kraju kaznivega dejanja) ali na osebo (napr.: pri oškodovancu ukradeni predmeti; ali pri osumljencu zaseženi predmeti).
- Poseben sklop podatkov so *dokumenti* (zapisniki, zaznamki, odredbe, potrdila, obvestila), ki jih lahko vežemo na dosje, zaznavo, kaznivo dejanje ali osebo. Dokumenti so sestavljeni iz osnovnih formatiranih podatkov in prostega teksta.

FIO je izredno obsežen in kompleksen informacijski sistem, ki pokriva najrazličnejša področja policijskega dela. Seveda pa se s tem zgodba ne konča, saj bo FIO potrebno dopolnjevati tako v vsebinskem kot tehnološkem smislu. Vsebinske prilagoditve so vezane predvsem na prilagajanje spremembam zakonodaje in predpisov. Za prihodnost pa se načrtuje razširitev podatkov FIO, ki so danes zgolj tekstualni, z digitalnimi slikami (fotografija, prstni odtisi, dokumenti). Posebno velik izziv je tudi povezovanje z informacijskimi sistemi znotraj EU. Po tehnološki plati pa se prav tako načrtuje stalno nadgrajevanje systemskega okolja FIO in uvajanje novih načinov, predvsem spletnih tehnologij za dostop do podatkov FIO.

## **5. IZDELAVA TESTNEGA SISTEMA OLAP NA PODROČJU PREPOVEDANIH DROG**

### **5.1. Analiza zahtev in definiranje problematike**

#### **5.1.1. Opis problematike**

Kazniva dejanja v povezavi s prepovedanimi drogami oziroma pojavne oblike tovrstnih kaznivih dejanj, uvrščamo v t.i. tiho kriminaliteto, kjer ne poznamo klasične žrtve kaznivega dejanja. Recimo, če vas nekdo oropa ali pa vam ukrade avtomobil boste to zagotovo prijavili policiji, medtem, ko od odvisnika ali uživalca prepovedanih drog ne moremo pričakovati, da bo ta prijavil svoj vir prepovedane droge.

Specifičen odnos med oškodovancem in storilcem (kjer imata oba korist) otežuje preiskovanje tovrstne kriminalitete, kar narekuje temu, da ostaja večina kaznivih dejanj neodkritih. V takih situacijah lahko policija samo preko znanih indicev, da je storjeno kaznivo dejanje, z vztrajnim analitičnim in operativno kriminalističnim delom utemeljuje sum storitve kaznivega dejanja. To pa zahteva od policije, da konstantno zbira ter vrednoti podatke in informacije z namenom odkrivanja: »Kaj se dogaja in kaj se pripravlja?«

Če so podani razlogi za sum, da je bilo storjeno kaznivo dejanje, za katerega se storilec preganja po uradni dolžnosti, mora policija ukrepati v skladu z zakonom in izvesti vse potrebno, da se storilec kaznivega dejanja izsledi, da se storilec ali udeleženec ne skrije ali ne pobegne, da se odkrijejo in zavarujejo sledi kaznivega dejanja in predmeti, ki utegnejo biti dokaz in, da se zberejo vsa obvestila, ki bi utegnila biti koristna za izvedbo kazenskega postopka. Razloge za sum je potrebno utemeljiti z operativno taktičnimi in tehničnimi ukrepi (npr. zbiranje obvestil, zavarovanje in dokumentiranje sledi, predmetov...) ter drugimi kriminalističnimi aktivnostmi.

Ob vsem navedenem pa je potrebno upoštevati resurse s katerimi policija razpolaga. Gre namreč za omejena finančna sredstva, omejena materialna sredstva in kadre. Zaradi vseh navedenih okoliščin je nujno potrebno maksimalno izkoristiti vse resurse in ustrezno organizirati aktivnosti, da bodo ti najbolj učinkovito uporabljeni. Ob navedenem ne moremo govoriti, da mora policija na vseh svojih področjih delovanja ravnati tako, da bo z najmanjšimi vloženimi sredstvi dosežen določen cilj, saj policija »mora« ob zaznavi vsakega kaznivega dejanja ravnati enako, pa naj si bo to umor ali še tako bagatelno dejanje.

Za uspešno preiskovanje in odkrivanje kaznivih dejanj na področju prepovedanih drog je tako potrebno usklajeno organiziranje aktivnosti in dobro poznavanje storilcev, njihovih motivov in načinov dela, dinamike delovanja trga s prepovedanimi drogami, nezakonitih poti, obsega organizirane kriminalitete, kakor tudi trende, ki prevladujejo na njem. Toda na tem mestu se kaj kmalu soočimo s problemom, ki narekuje vprašanje: »Na kakšen način iz vseh zbranih podatkov pridobiti vse tiste koristne informacije, ki so potrebne za učinkovito odločanje?«

Problematika prekomerne uporabe prepovedanih drog se ne kaže samo v njihovi uporabi oz. uživanju, temveč tudi v negativnih posledicah, ki se kažejo na: socialnoekonomski ogroženosti prebivalstva, ogroženosti zdravja prebivalstva, ogroženosti varnosti prebivalstva. Zato so vse bolj zaskrbljujoči podatki, da v zadnjih desetih letih, konstantno narašča:

- število zasegov in količine prepovedanih drog, kar velja predvsem za tiste droge, ki so najbolj razširjene in katerih uporaba je nevarnejša;
- število kaznivih dejanj, ki so storjena zaradi pridobivanja finančnih sredstev za nakup prepovedanih drog;
- število umrlih zaradi posledic uživanja prepovedanih drog;
- vključenost naših državljanov v mednarodno trgovino s prepovedanimi drogami.

Slovenska policija je zato potrebovala strateško analizo o problematiki prekomerne uporabe prepovedanih drog, ki vsebuje analizo o trenutnem stanju, kar se tiče problematike prepovedanih drog, posledicah prometa s prepovedanimi drogami ter vlogi posameznih skupin organizirane kriminalitete, ki se ukvarjajo s kriminalnimi dejavnostmi v povezavi s prepovedanimi drogami.

Glede na navedene težave in ponujene rešitve smo si v Upravi kriminalistične policije, Sektorju za računalniško kriminaliteto in kriminalistično analitiko, Oddelku za kriminalistično analitiko, zastavili cilj, da spoznamo možnosti in priložnosti uporabe sodobnih informacijskih tehnologij, na področju hranjenja ter analize podatkov, v namene:

- okrepitev dejavnosti policije usmerjene proti organiziranemu kriminalu in
- izboljšanju obveščevalnih in analitični del pri odkrivanju in preiskovanju kaznivih dejanj povezanih s prepovedanimi drogami.

### **5.1.2. Definiranje potreb uporabnikov**

Pri sprejemanju odločitve za testno področje, na podlagi katerega bomo ugotavljali praktičnost uporabe sistema OLAP, smo morali predhodno ugotoviti, kakšne potrebe imajo v resnici uporabniki in ali je izdelava tovrstnega testnega sistema za izbrano področje upravičena. Zato je bilo potrebno predhodno opraviti razgovore z različnimi uporabniki, da se ugotovijo, katere informacije si želijo pridobiti od sistema, na kakšen način sedaj pridobivajo informacije in kaj ob tem pogrešajo.

Ker se je ob izbiri področja za testiranje prav tako izvajala strateška analiza o problematiki prekomerne uporabe prepovedanih drog, sama odločitev za izbiro testnega področja ni bila preveč vprašljiva, saj smo s tem pridobili možnost, da konkretno preizkusimo sistem OLAP.

Definirane potrebe uporabnikov tako izhajajo iz ugotovljenih potreb na podlagi intervjujev z uporabniki, seznanitve z obravnavanim področjem in potreb, ki jih je narekovala izvedba strateške analize na področju prepovedanih drog.



### **5.1.3. Obstoječe stanje**

Vsako delovno področje v policiji ima pred-pripravljene standardne statistične preglede, ki jih lahko dodatno omejujemo preko nekaterih najpomembnejših parametrov. Ko parametre pravilno vnesemo oz. izberemo, se izvede vnaprej pripravljena paketna obdelava, ki nam izdela listo s statističnim pregledom. To listo lahko le pregledujemo, jo izpišemo ali pa prenesemo na osebni računalnik, kjer jo lahko naknadno obdelujemo.

Pri kreiranju poročil so uporabniki tako omejeni z pred-definiranimi polji, katerih postavitev večinoma ustreza pripravi podatkov za standardna letna in polletna statistična poročila, ki jih pripravlja policija. Ker tovrstna poročila za uporabnike ne predstavljajo dodatne vrednosti in zaradi dejstva, da nikoli ne moremo zadovoljiti vseh potreb uporabnikov in s tem izdelati dovolj enostavnih ter hkrati natančnih in dovolj podrobnih predstavitev podatkov, so rezultat številna pred-definirana poročila, ki jih uporabniki ne uporabljajo. Na drugi strani pa je prav tako zaznati pomanjkanje znanja o podatkih s strani uporabnikov, saj imamo opravka s številnimi napačnimi interpretacijami podatkov.

Ker je zaznati vse večjo potrebo po podatkih, za potrebe odločanja in vodenja, so si nekateri oddelki razvili svoje evidence, na podlagi katerih spremljajo in zaznavajo problematiko na njihovem delovnem področju. Rezultat tovrstnih pristopov pa so številni nekonsistentni podatki.

Zaradi kompleksnosti sistemov za pripravo poročil so informatiki prav tako obremenjeni s pripravo podatkov za različne uporabnike, zato je potrebno kljub celovitemu pogledu na podatke posamezne organizacije, še vedno te prikazati na jasn in pregleden način in jih s tem narediti uporabne. Za doseganje tega cilja potrebujemo zmogljivo OLAP orodje, ki predstavi podatke na jasn in pregleden način (Roblek, 2003, str. 323) in zagotavlja celovit pogled na poslovanje organizacije skozi različne vidike, in nudi končnemu uporabniku podporo pri analizi in odločanju (Golob et al., 2001).

### **5.1.4. Izbira podatkov in njihova priprava**

Po Zakonu o policiji, policija upravlja zbirke osebnih podatkov (v nadaljnjem besedilu: evidence), ki jih zaradi opravljanja nalog zbirajo, obdelujejo, shranjujejo, posredujejo in uporabljajo policisti. Pri pripravi testnega sistema OLAP smo razpolagali s podatki, ki izhajajo iz:

- evidence kršiteljev in prekrškov ter
- evidence zaznav kaznivih dejanj.

Z osebni podatki osumljencev oz. obravnavanih oseb kot so: osebno ime, rojstne podatke, EMŠO, vzdevke ali njihovimi lažnimi imeni nismo razpolagali. Prav tako nismo razpolagali s podatki o prijaviteljih in oškodovancih kaznivih dejanj ali prekrškov.

Razpolagali smo samo z osebnimi podatki, ki kažejo na:

- identifikacijsko številko osebe (CR),
- spol,
- državljanstvo,
- prebivališče v času obravnave,
- starost osebe v času obravnave in
- številko zadeve v kateri je bila ta oseba obravnavana.

Zraven kaznivih dejanj in prekrškov zaradi prepovedanih drog smo za obravnavane osebe, prav tako razpolagali s podatki, ki se nanašajo na njihovo obravnavanost zaradi ostalih prekrškov in/ali kaznivih dejanj. Podatki, ki so se nanašali na segment dogodkov, so bili naslednji:

- številka zadeve,
- vrsta dejanja,
- enota obravnave,
- čas in prizorišče dogodka,
- motiv storitve in
- način storitve.

V zvezi s predmeti kaznivih dejanj ali prekrškov smo razpolagali samo s podatki, ki se nanašajo na prepovedane droge. Torej predmet v našem primeru je le prepovedana droga in ne ostali predmeti, ki so lahko udeleženi v posameznem dogodku (napr.: *ukradeno vozilo, ki je vezano na osebo, kateri je bila zasežena prepovedana droga*). Pri prepovedanih drogah smo razpolagali z naslednjimi podatki:

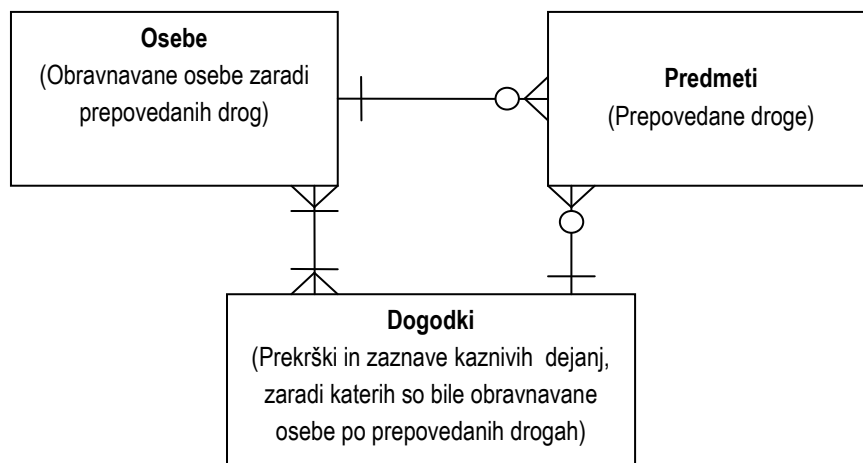
- številka zadeve,
- identifikacijsko številko osebe (CR), kateri je bila zasežena prepovedana droga,
- vrsta prepovedane droge,
- količina zasega in merska enota,
- datum in ura zasega,
- država izvora, tranzita in končna država prepovedane droge,
- organizacijska enota, ki je zasegla predmet,
- kraj in prizorišče zasega.

Osnovni procesi priprave podatkov, so zadevali integracijo različnih evidenc v centralno lokacijo in s tem:

- odpravo neskladnih in napačnih podatkov, ki so se pojavili pri vnosu podatkov (neveljavni datumi);
- odprava manjkajočih vrednosti (manjkajoče vrednosti nekaterih atributov, ker niso bile izmerjene, vnesene ali pa so bile preprosto izgubljene);
- neskladnosti med podatki iz različnih evidenc (različna imena atributov za iste podatke);
- izbiro podatkov, ki so bili pomembni za nas;
- predstavitev podatkov v bolj primerni obliki.

Za boljše razumevanje povezav med posameznimi segmenti podatkov, vam na spodnji sliki 8, prikazujemo E-R diagram segmentov podatkov za pripravo večdimenzionalnega podatkovnega modela.

**Slika 7: E-R diagram segmentov podatkov za pripravo testnega sistema OLAP**



Kot lahko razberemo iz zgornjega diagrama, je lahko posamezna oseba, ki je bila obravnavana zaradi prepovedanih drog obravnavana zaradi ene ali več zadev. V posamezni zadevi je lahko posamezna oseba (ali več) obravnavana zaradi več različnih vrst kaznivih dejanj ali prekrškov (npr.: *Pri hišni preiskavi je bila najdena manjša količina prepovedane droge in znatno število radijskih sprejemnikov, za katere je bilo ugotovljeno, da so odtujeni*).

V okviru posamezne zadeve pri zaznavi enega ali več kaznivih dejanj ali prekrškov, je lahko vključeno več oseb, vendar se zasegi prepovedanih drog evidentirajo oz. vežejo na samo eno osebo (torej na osebo s strani katere je bila prepovedana droga dejansko zasežena). V posamezni zadevi je lahko torej zaseženo več vrst prepovedanih drog (ali nobena). Istovrstne prepovedane droge so lahko v okviru posamezne zadeve, zasežene v različnih količinah in na različnih krajih (npr.: *Z osebno preiskavo osebe se ugotovi, da oseba poseduje prepovedane droge, ker pa obstaja utemeljen sum, da ima oseba, prav tako skrito prepovedano drogo na domu, preiskovalni sodnik odobri hišno preiskavo, s katero se ugotovi, da je bilo najdenih več zavitkov prepovedane droge heroin*).

Posamezno zadevo praviloma vodi samo ena organizacijska enota, vendar se lahko kljub temu v posamezni zadevi evidentira več organizacijskih enot (npr.: *Občan Ljubljane je s strani Policijske uprave Ljubljana obravnavan zaradi preprodaje prepovedanih drog, medtem, ko pa je do dejanskega zasega prepovedanih drog prišlo na območju Kopra, s strani PU Koper. V kolikor imamo opravka z organizirano združbo, ki deluje po celotni Sloveniji je stvar še bolj zapletena*). Ker smo razpolagali s podatki o zaznavi kaznivih dejanj, nismo mogli iz podatkov razbrati katera organizacijska enota dejansko vodi primer (npr.: na podlagi podaje kazenske ovadbe zoper osebe). V kolikor bi te podatke združevali, glede na zadnji ali prvi vnos v okviru posamezne zadeve, bi bili ti neskladni z uradno statistiko policije.

## 5.2. Faze večdimenzionalnega modeliranja podatkov

### 5.2.1. Izbira procesov

Prvi korak pri postavljanju večdimenzionalnega modela podatkov je odločitev, katere procese bomo modelirali. Odločitev o procesih sprejmemo šele po temeljitem razgovoru s končnimi uporabniki, na podlagi katerih spoznamo njihove želje in potrebe. Za sprejetje odločitve pa moramo prav tako poznati in predvsem razumeti razpoložljive podatke, ki jih bomo uporabili (Kimball 2002, str. 33). Priprava večdimenzionalnega modela podatkov je torej proces, v katerem potrebe uporabnikov preslikamo na dejansko razpoložljive podatke (Remškar 2002, str. 22).

V našem primeru smo si za proces izbrali obveščevalno analitično delo policije na področju prepovedanih drog, ki je usmerjeno v analiziranje zbranih podatkov in se nanaša na:

- zaznavo kaznivih dejanj in prekrškov povezanih s prepovedanimi drogami,
- preteklo obravnavo oseb, ki so bile obravnavane zaradi prepovedanih drog,
- zasege prepovedanih drog.

Večdimenzionalni podatkovni model, ki ga bomo pripravili za opis opredeljenega procesa, nam bo tako pomagal odgovoriti na vprašanja, kot so:

- Katere vrste prepovedanih drog so najpogosteje zasežene ob praznikih in ob vikendih?
- Ali se število zasegov, glede na njihovo količino po posameznih tipih prepovedanih drog, razlikuje po času v dnevnu?
- Katere so karakteristike zasegov prepovedanih drog na mejnih prehodih?
- Katere so najpogostejše končne države tranzita prepovedanih drog po njihovih vrstah?
- Ali se motivi storitev osumljencev, ki so bili obravnavani zaradi kaznivih dejanj v povezavi s prepovedanimi drogami, razlikujejo po njihovih starostnih skupinah?
- Katera so najpogostejša prizorišča manjših zasegov prepovedanih drog s strani mladoletnih oseb?
- Ali so storilci kaznivih dejanj v povezavi s prepovedanimi drogami, povratniki in storilci drugih kaznivih dejanj?
- Kako so organizirani storilci kaznivih dejanj v povezavi s prepovedanimi drogami, pri ostalih kaznivih dejanjih?

Osrednji in bistveni del je torej analiza podatkov, ki bo lahko vplivala na odločitve vodilnih delavcev policije in na oblikovanje dolgoročne politike kriminalistične policije na področju prepovedanih drog, saj bo omogočila: usklajeno organiziranje aktivnosti policije, poznavanje storilcev, njihovih motivov in načinov dela, dinamike delovanja trga s prepovedanimi drogami, nezakonitih poti, obsega organizirane kriminalitete, kakor tudi trende, ki prevladujejo na njem.

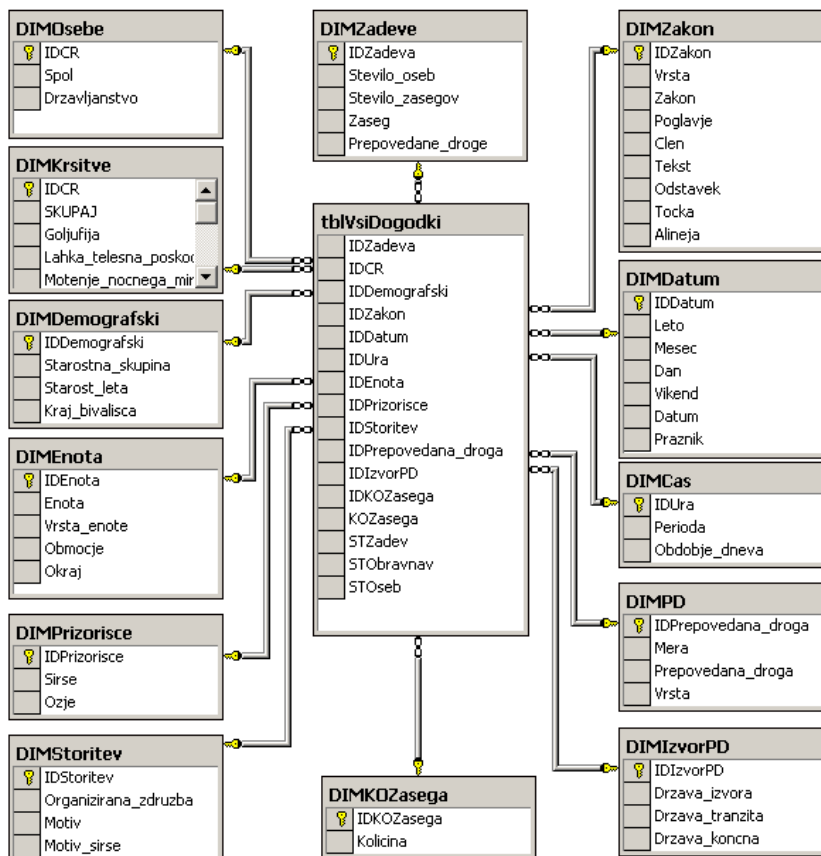
### 5.2.2. Določitev osnovne enote

Ko smo se odločili za proces, ki ga bomo modelirali, se moramo v naslednjem koraku vprašati ali bomo v tabelo dejstev shranjevali podrobne podatke, zbirne podatke (agregirane) ali posnetke stanj. Pri odločitvi o osnovni oz. najmanjši enoti (angl. grain) moramo vedeti, da se skrivnost fleksibilnosti večdimenzionalnega podatkovnega modela skriva v shranjevanju podrobnih podatkov oziroma podatkov, ki jih ne moremo več nadalje deliti. Zato naj bo osnovna enota kar najbolj natančen podatek, ki je na voljo, saj le tako lahko omogočimo poljubne prereze in poglobljene poizvedbe. Agregirani podatki naj tako le dopolnjujejo podrobne podatke, izključno v zagotavljanju izboljšanih učinkov poizvedovanja in naj ti ne služijo kot njihova zamenjava (Kimball 2002, str. 34). Osnovna enota oziroma posamezna vrstica v tabeli dejstev se tako nanaša na posamezen zaseg prepovedane droge in/ali obravnave oseb, zaradi prekrška ali kaznivega dejanja.

### 5.2.3. Postavitev osnovnih dimenzij

Na osnovi navedbe osnovne enote lahko hitro določimo tudi osnovne dimenzije, ki to enoto opisujejo. Glede na razpoložljivost podatkov in zahteve uporabnikov pa se lahko odločimo še za dodatne dimenzije. Vendar moramo ob tem paziti, da s tem ne povzročimo podvajanja vrstic v tabeli dejstev (Kimball 2002, str. 35). Rezultat izbire dimenzij je tipična zvezdasta shema, ki ima skupno 13 dimenzij in vam jo prikazujemo na sliki 9.

Slika 8: Večdimenzionalni model podatkov za področje prepovedanih drog



Ker vam bomo v nadaljevanju podrobneje predstavili posamezne dimenzije, si sedaj pogledjmo kratek pomen podatkov in primere njihove vsebine, ki vam jih predstavljamo v tabeli 2.

**Tabela 2: Pomeni in primeri atributov, po posameznih dimenzijah**

	Atribut	Pomen in primer vsebine
Dimenzija zadeva (DIMZadeva)	IDZadeva	Primarni ključ (številka zadeve): 346786
	Število_oseb	Deskriptivna vrednost števila oseb, ki so vključeni v zadevo: od 01 do 02
	Število_zasegov	Deskriptivna vrednost števila zasegov v posamezni zadevi: od 00 do 01
	Zaseg	Ali je v zadevi prišlo do zasega (Da/Ne): DA
	Prepovedane_droge	Ali je so v zadevi vključene prepovedane droge (Da/Ne): DA
Dimenzija Zakonodaja (DIMZakon)	IDZakon	Primarni ključ (1..N): 69
	Vrsta	Vrsta dejanja (kaznivo dejanje ali prekršek): KD
	Zakon	Ime zakona (KZ RS): KD95
	Poglavje	Poglavje v zakonu: KAZNIVA DEJANJA ZOPER ČLOVEKOVO ZDRAVJE
	Člen	Člen : 196
	Tekst	Tekstualno ime člena: NEUPRAVIČENA PROIZVODNJA IN PROMET Z MAMILI
	Odstavek	Odstavek člena: 1
	Točka	Točka člena:
Alineja	Alineja člena : 1	
Dimenzija datum (DIMDatum)	IDDatum	Primarni ključ (1..N): 254
	Leto	Leto : 2000
	Mesec	Mesec: JANUAR
	Dan	Dan v tednu: TOREK
	Vikend	Vikend (Da/Ne): NE
	Datum	Polna vrednost datuma: 11.1.2000
	Praznik	Praznik (Da/Ne): NE
Dimenzija Čas (DIMCas)	IDUra	Primarni ključ (1..N): 6
	Perioda	Urna perioda: od 06 do 08
	Obdobje_dneva	Tekstualni zapis dela dneva: ZJUTRAJ
Dimenzija prepovedane droge (DIMPDP)	IDPrepovedana_droga	Primarni ključ (1..N): 12
	Mera	Merska enota: GRAM
	Prepovedana_droga	Ime prepovedane droge: HEROIN
	Vrsta	Vrsta oz. skupina prepovedane droge: DEPRESORJI
Dimenzija izvor prepovedane droge (DIMIzvorPD)	IDIzvorPD	Primarni ključ (1..N): 65
	Država_izvora	Država izvora: ČEŠKA
	Država_tranzita	Država tranzita: SLOVENIJA
	Država_končna	Država končna: HRVAŠKA
Dimenzija količina zasega (DIMKOZasega)	IDKOZasega	Primarni ključ (1..N): 3
	Kolicina	Deskriptivna vrednost količine zasega: VELIKA
Dimenzija motiv storitve (DIMStoritev)	IDStoritev	Primarni ključ (1..N): 74
	Organizirana_združba	Oblika organizirane združbe: ZAČASNA ZDRUŽBA
	Motiv	Opredeljen motiv zaradi katerega je bilo storjeno kaznivo dejanje: PRIDOBITEV SREDSTEV ZA NAKUP MAMIL
	Motiv_sirse	Širše opredeljeni motiv: ZADOLŽENOST
Dimenzija prizorišče (DIMPrizorišče)	IDPrizorišče	Primarni ključ (1..N): 29
	Sirše	Širše prizorišče storitve: ŠOLA IN ŠPORT
	Ožje	Ožje opredeljeno prizorišče storitve kaznivega dejanja / prekrška: PROSTOR ZA ŠPORT, REKREACIJO, ZABAVO
Dimenzija enota obravnave (DIMEnota)	IDEnota	Primarni ključ (1..N): 710
	Enota	Organizacijska enota policije v zadevi: PP GORNJA RADGONA
	Vrsta_enote	Vrsta organizacijske enote: POLICIJSKA POSTAJA
	Območje	Območje delovanja enote: MURSKA SOBOTA
	Okraj	Okraj delovanja enote: GORNJA RADGONA
Dimenzija demografski podatki za osebe (DIMDemografski)	IDDemografski	Primarni ključ (1..N): 6099
	Starostna_skupina	Deskriptivna vrednost starosti osebe: od 25 do 30
	Starost_leta	Starost osebe v letih: 28
	Kraj_bivališča	Kraj bivališča osebe: DOMŽALE
Dimenzija najpogostejših kršitev (DIMKršitve)	IDCR	Primarni ključ (CR): 1448694
	Goljufija	Kot attribute smo opredelili najpogostejša kazniva dejanja in prekrške zaradi katerih so bile obravnavane osebe. Najpogostejša dejanja, skupno predstavljajo več kot 90% vseh dejanj.
	Lahka_telesna_poškodba	
	a	
	Motenje_nočnega_miru	Ker je pogostost izvajanja posameznih dejanj različna, smo glede na frekvenčno porazdelitev vrednosti, posameznim atributom (oz. dejanjem) postavili samo dve vrednosti (0 in 1+), ostalim pa več (0; 1; 2-4; 5+).
	Nedostojno_vedenje_do_UO	
	...	
Dimenzija osebe (DIMOsebe)	IDCR	Primarni ključ (CR): 1448694
	Spol	Spol osebe: MOŠKI
	Državljanstvo	Državljanstvo oseb: SLOVENSKO

#### 5.2.4. Izbira mer

Kazniva dejanja s področja prepovedanih drog uvrščamo v t.i. tiho kriminaliteto, kjer ne poznamo klasične žrtve kaznivega dejanja. Recimo, če vas nekdo oropa ali pa vam ukrade avtomobil boste to zagotovo prijavili policiji, medtem ko od odvisnika ali uživalca prepovedanih drog podobnega ravnanja (da bi ta prijavil svoj vir prepovedane droge oz. preprodajalca) ne moremo pričakovati. Zato policija pri statističnem prikazovanju problematike prepovedanih drog uporablja mere, ki kažejo na obseg policijskega delovanja na tem področju:

- količino zasežene prepovedane droge,
- število zaznav kaznivih dejanj in/ali število kazenskih ovadb,
- število obravnavanih in/ali ovadenih oseb.

Mediji in javnost tako neposredno povezujejo podatke o količini zasežene prepovedane droge z dejanskim obsegom oz. razpoložljivostjo te v R Sloveniji. Pri tem moremo poudariti, da se Slovenija nahaja na t.i. Balkanski poti, kar pomeni, da je velika večina odmevnih zasegov namenjena v druge Evropske države. Podatek o tem ali je zasežena droga namenjena za slovenski trg ali drugi je tako poznan samo njihovim organizatorjem. Po nekaterih raziskavah se dnevno v Sloveniji porabi toliko prepovedane droge, kolikor se te zaseže v celotnem letu, količina zasega prepovedane droge je tako le »vrh ledene gore«.

Količina zasega ali število obravnavanih oseb je tako le odraz aktivnosti organov pregona in zato naj ne služi kot indikator razpoložljivosti te na trgu. Kot indikator bi lahko uporabili njeno ceno in čistino (t.j. razmerje med prepovedano drogo in dodatnimi primesmi, ki jih ta vsebuje). Vendar te vrednosti prav tako neposredno ne odražajo obsega trga povpraševanja temveč le njegove ponudbe. Uporabnost cene prepovedane droge kot mere, je prav tako vprašljiva saj se cene spreminjajo glede na čas, kvaliteto in področje prodaje. Prepovedanih drog pa prav tako ne moremo na slepo med seboj primerjati.

Za vrednotenje smo uporabili standardne mere (količina zasega, število posamičnih zadev, oseb in število obravnav oseb) skupaj z vrednostnimi lestvicami oz. deskriptivnimi vrednostmi količine zasega. »Zakaj vrednostne lestvice?« Če uporabimo kot mero samo količino zasega, nam lahko že en sam zaseg prepovedane droge povzdigne njeno statistično vrednost in posledično njeno neprimerljivost (po letih, dnevih, času, kraju, ...). Kadar pa uporabljamo kot mero število obravnavanih oseb ali posameznih zadev, pa se moramo zavedati na neprimerljivost posameznih zadev, saj imamo v skupnem košu zadeve pri katerih je bilo zaseženo 5 kg prepovedane droge heroin in zadeve kjer je bilo zaseženo do 5 tbl. ecstasy-ja.

Za pravilne seštevke in vsote posameznih mer smo uporabljali funkcijo »DistinctCount« in t.i. enakomerne uteži (katerih skupna vsota v okviru posamezne skupine je 1), katerih uporabo poznamo iz t.i. »pomožnih tabel« pri večdimenzionalnem modeliranju podatkov, vendar o uporabi teh nekoliko več v nadaljevanju.

### 5.3. Posebnosti pri večdimenzionalnem modeliranju podatkov

#### 5.3.1. Uporaba merskih lestvic

Vrednostne lestvice smo pripravili za vsako posamezno vrsto prepovedane droge, zaradi različnega pomena teh in uporabe različnih merskih enot pri evidentiranju. Zasege smo tako ločevali, glede na količino zasega na: ni zasega, majhne, srednje in velike zasege. Iz spodnje tabele 3, na kateri vam prikazujemo del vrednostne lestvice za prepovedane droge, lahko razberete, da se je posamezen zaseg, kjer je bilo zaseženo nad 100g heroina obravnaval kot velik zaseg, medtem ko se je ta količina pri prepovedani drogi kanabis rastlina (marihuana) obravnavala kot majhen zaseg.

**Tabela 3: Vrednostna lestvica za prepovedane droge**

Prepovedana droge	Majhna	Velika
AMFETAMIN	do 5 g	nad 100 g
BENZODIAZEPINI	do 5 g	nad 100 g
ECSTASY	do 5 tbl	nad 100 tbl
HEROIN	do 5 g	nad 100 g
KANABIS RASTLINA (BILJKA)	do 10 kom	nad 50 kom
KANABIS RASTLINA (MARIHUANA)	do 100 g	nad 1000 g
SMOLA KANABISA (HAŠIŠ)	do 5 g	nad 1000 g
EKSTRAKTI IN TINKTURE KANABISA	do 5 ml	nad 100 ml
KOKAIN	do 5 g	nad 100 g
METADON	do 100 ml	nad 1000 ml

Uporabnost vrednostnih lestvic se je izkazala kot neizbežna na skoraj vseh področjih analiziranja, saj smo lahko na enostaven način omejevali naše področje opazovanja (napr. Odgovorili smo lahko na vprašanje: *»Ali je prišlo do porasta majhnih zasegov prepovedanih drog na področju Ljubljane v primerjavi z ostalimi leti in na katerih prizoriščih?«* )

#### 5.3.2. Časovna dimenzija

Če pregledujemo statistična poročila policije, lahko opazimo, da se časovna komponenta uporablja predvsem za primerjavo upada ali porasta opazovanih statističnih spremenljivk po letih ali mesecih. Torej pri iskanju odgovorov na vprašanja kot so: *»Ali je več prepovedanih drog zaseženo ob vikendih kakor med tednom?«*, potrebujemo dimenzijo, ki nam bo omogočala t.i. inteligentno obravnavo časa. Pri pripravi dimenzije »DIMDatum«, ki predstavlja časovno dimenzijo smo podatkovno polje datum opredelili kot entiteto, kateri smo dodali naslednje attribute: leto, mesec, dan, vikend, praznik. Po priporočilih za večdimenzionalno modeliranje podatkov, glede izbire ustreznih ključev, smo za ključ izbrali naraščajočo število od 1. Alternativno pa lahko datumsko polje spremenimo v številčno (formata YYYYMMDD - integer) in ga kot takšnega uporabimo za ključ. Za potrebe analiziranja dogajanj po urah, urnih periodah in delih dneva smo prav tako pripravili dimenzijo »DIMCas«.



### 5.3.3. Neenakomerna globina dimenzij

Čeprav pri opredelitvi hierarhij oz. nivojev v naših dimenzijah nismo imeli težav, so v praksi te več ali manj vedno problematične, saj le redkokdaj naletimo na značilen šolski primer. Kot primer šolske dimenzije, kjer so hierarhije enake globine in smiselne v vseh ravneh, vam prikazujemo na spodnji tabeli 4, pri prikazu hierarhij dimenzije »DIMZakon«.

**Tabela 4: Prikaz hierarhij pri dimenziji »DIMZakon«**

IDZaki	Vrsta	Zakon	Poglavje	Člen	Tekst	Odstavek
250	KD	KD95	K.D. ZOPER PRAVNI PROMET	258	OVERITEV LAŽNE VSEBINE	1
251	KD	KD95	K.D. ZOPER PRAVNI PROMET	258	OVERITEV LAŽNE VSEBINE	2
252	KD	KD95	K.D. ZOPER URADNO DOLŽNOST IN JAVN/	261	ZLORABA URADNEGA POLOŽAJA ALI URADNIH PRAVIC	1
253	KD	KD95	K.D. ZOPER URADNO DOLŽNOST IN JAVN/	261	ZLORABA URADNEGA POLOŽAJA ALI URADNIH PRAVIC	4
254	KD	KD95	K.D. ZOPER URADNO DOLŽNOST IN JAVN/	268	DAJANJE PODKUPNINE	1
255	KD	KD95	K.D. ZOPER URADNO DOLŽNOST IN JAVN/	270	KRŠITEV ČLOVEŠKEGA DOSTOJANSTVA Z ZLORABO URADNEG	1
256	KD	KD95	K.D. ZOPER PRAVOSODJE	286	OPUSTITEV OVADBE KAZNIVEGA DEJANJA ALI STORILCA	1
257	KD	KD95	K.D. ZOPER PRAVOSODJE	287	POMOČ STORILCU PO STORITVI KAZNIVEGA DEJANJA	1
258	KD	KD95	K.D. ZOPER PRAVOSODJE	287	POMOČ STORILCU PO STORITVI KAZNIVEGA DEJANJA	2
259	KD	KD95	K.D. ZOPER PRAVOSODJE	288	KRIVA OVADBA	1
260	KD	KD95	K.D. ZOPER PRAVOSODJE	288	KRIVA OVADBA	2
261	KD	KD95	K.D. ZOPER PRAVOSODJE	288	KRIVA OVADBA	4

Ker orodja OLAP za končne uporabnike pričakujejo enakomerno globino ravni zaradi preprostega vrtanja iz višjih v nižje ravni, je za praktično uporabo potrebno tovrstno dimenzijo pravilno modelirati. Rešitev, ki omogoča preprosto vrtanje v globino je v predelavi dimenzije tako, da se vse hierarhije izpolnijo do najnižje ravni s podvajanjem zapisov. Na ta način je sicer uvedeno nekaj podvajanja, vendar pa ima sedaj dimenzija enakomerno globino v vseh vejah, uporabnik pa lahko pregleduje podatke na poljubno izbrani ravni, vrta v globino in nazaj. Podvajanje zapisov lahko prav tako opazite iz zgornjega prikaza v tabeli 4.

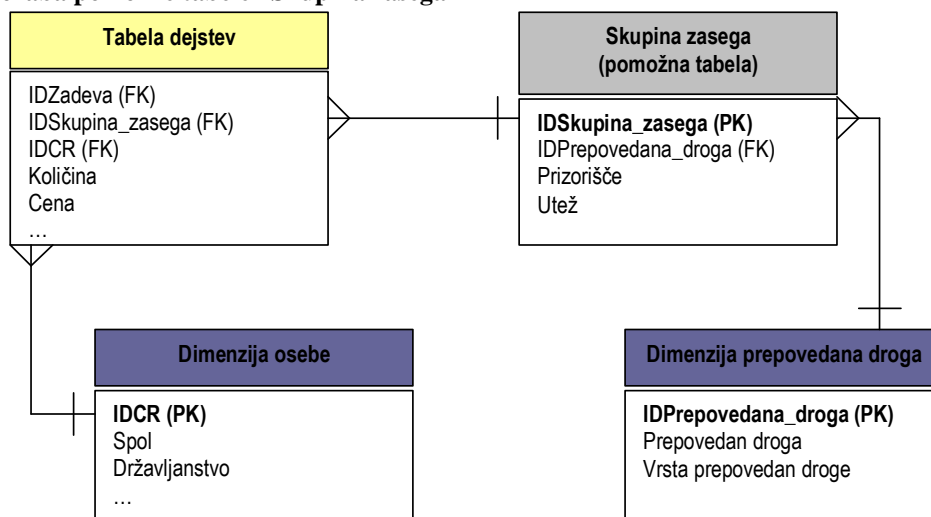
Kadar je hierarhija izrazito neuravnotežena, na primer če ena veja sega mnogo globlje od ostalih, lahko prav tako zmanjšamo število hierarhij. Torej ali je bolj smiselno ožiti ali širiti dimenzijo? Menimo, da je to odvisno predvsem od potreb in želja uporabnikov, vendar moramo ob tem opozoriti, da nam pri analiziranju preveliko število hierarhij nekako zamegli rezultate poizvedovanja, medtem, ko nam manjše število poda preveč grobo sliko.

Pri modeliranju dimenzij z izrazito neenakomerno globino lahko uporabimo rešitev *rekurzivnega kazalca*, tako da vsakemu elementu dimenzije določimo njemu nadrejen element v hierarhiji (v praksi se na tak način rešujejo predvsem problemi organiziranosti zaposlenih »nadrejen/podrejen«). Z uporabo takšnega kazalca bi lahko za vsak element v hierarhiji ugotovili, kateri je njegov nadrejeni element in preko njega tudi ugotovili kateri element je najvišji v hierarhiji (Kimball, 2001). Vendar pa ta kazalec ne predstavlja najboljše rešitve, saj ga je nemogoče neposredno uporabljati v SQL poizvedbah. Kot alternativo rešitev neenakomernih globin hierarhij in nepravilne drevesne strukture hierarhije (to je kadar ima nek element lahko več kot en sam nadrejen element) lahko uporabimo rešitve, ki jih poznamo iz uporabe t.i. pomožnih tabel, ki vam jih prikazujemo v nadaljevanju.

### 5.3.4. Uporaba pomožnih tabel

Pri izbiri dimenzij velja pravilo, da naj bi dimenzije določili tako, da je za opis zapisa v tabeli dejstev dovolj ena vrednost dimenzije. Torej, da je razmerje med dimenzijsko tabelo in tabelo dejstev v razmerju ena proti mnogo. Zgodi pa se nam lahko, da to ni mogoče in da ne gre drugače kot, da dimenzija vsebuje večvrednostne podatke. Da bi se izognili povezavi mnogo proti mnogo, lahko uporabimo t.i. pomožno tabelo (angl Help table) (Kimball, 2001). Primer uporabe pomožne tabele pri zasegih prepovedane droge vam prikazujemo na spodnji sliki 9.

Slika 9: Uporaba pomožne tabele »Skupina zasega«



Kot smo že dejali je lahko v posamezni zadevi obravnavano več oseb in je zaseženo več različnih vrst prepovedanih drog, medtem, ko se posamezni zasegi prepovedane droge vežejo oz. evidentirajo samo na osebo, kateri je bila ta dejansko zasežena. V kolikor bi želeli analizirati zasege prepovedanih drog po vseh osebah, ki so bile vključene v posamezno zadevo in ne samo po osebah katerim je bila ta dejansko zasežena, bi morali vzpostaviti pomožno tabelo med dimenzijo prepovedanih drog in tabelo dejstev, kot jo prikazujemo na sliki 9.

S tem bi pridobili detajlnejši pregled oseb, ki so bile obravnavane zaradi posameznih vrst prepovedanih drog. Za pravilne meritve, pa bi morali vpeljati parametre uteži in poenotiti merske enote, saj se pri meritvah količine zasega prepovedane droge uporabljajo različne merske enote (kot so: grami, mililitri, kosi, ...). Pri poenotenju enot bi prav tako lahko uporabili zadnjo ceno prepovedane droge, ki je veljala na trgu.

Vendar je uporaba pomožne tabele ostala le kot ideja. Zakaj? Vloge oseb znotraj posamezne zadeve (kot so: organizator, tihotapec, preprodajalec) se ne evidentirajo v evidenci zaznav kaznivih dejanj, zato tovrsten prikaz podatkov ne bi bil kaj dosti analitično zanimiv. Sprememba bi vplivala na izbrano najmanjšo enoto in na primerljivost podatkov z uradno statistiko policije.

Primer uporabe pomožne tabele bi prav tako lahko uporabili v primeru nepravilnih drevesnih struktur hierarhije. To je kadar ima nek element več kot en sam nadrejen element. V pomožno tabelo dodamo polja kot so: oznaka nadrejenega elementa; oznaka elementa; oddaljenost od nadrejenega elementa; zastavica najnižje ravni; in zastavica najvišje ravni. Kadar element nima nadrejenih elementov, govorimo, da je zastavica najvišje ravni dvignjena, medtem pa kadar je dvignjena zastavica najnižje ravni, govorimo, da je element v dnu hierarhije. Kadar pa imamo opravka s *spreminjajočo se hierarhijo*, v pomožno tabelo dodamo še parametra, ki označujeta časovni začetek in konec veljavnosti relacije med elementoma v hierarhiji. Prednost, ki jo prinese takšna vmesna tabela je v tem, da lahko enostavno poizvedujemo po posameznih ravneh hierarhije in podatke pravilno združujemo (Kimball, 1998).

Uporaba pomožne tabele ima v praksi kar nekaj pomanjkljivosti (Ferle, 2001a):

- prenos podatkov v pomožno tabelo je razmeroma zapleten;
- pomožna tabela vsebuje zelo veliko število zapisov, saj so v njej vsi pari vseh ravni v hierarhiji, kar pri velikem številu zapisov v dimenziji predstavlja mnogo večje število zapisov v pomožni tabeli;
- uporaba pomožne tabele je prezapletena za običajne uporabnike, ker jih je treba naučiti, kako jo pravilno uporabljati;
- pomožna tabela ni primerna za uporabo v značilnih orodjih za analizo podatkov, ker ne dovoljuje značilnega vrtnja v globino.

### 5.3.5. Počasi spreminjajoče se dimenzije

Praviloma je posamezna dimenzija neodvisna od drugih dimenzij in časa. Torej pričakujemo, da so vrednosti posameznih atributov fiksne oziroma, da se te ne bodo spreminjale. Sedaj pa se vprašajmo, ali lahko podobno trdimo za dimenzijo zakonodaja. Dimenzije, katerih vrednosti se s časom spreminjajo, opredeljujemo kot t.i. počasi spreminjajoče se dimenzije (ang. Slowly Changing Dimension - SCD). Pri zagotovitvi neodvisne strukture tovrstnih dimenzij, imamo na razpolago več načinov posodabljanja teh (Kimball 2002, str. 95):

- prepíšemo stare vrednosti z novimi,
- dodamo nov zapis v dimenzijsko tabelo,
- v dimenzijski tabeli lahko uvedemo polja za trenutno in prejšnje vrednosti atributov.

Z izbiro prve možnosti t.j., da prepíšemo stare vrednosti z novimi, poenostavimo vzdrževanje dimenzijske tabele, a s tem izgubimo vse zgodovinske podatke. Zato, v kolikor želimo ohraniti pregled nad spremembami dimenzijskih vrednosti izberemo drugo možnost. Vendar pa pri izbiri slednje, zaradi dodajanja novih zapisov v dimenzijsko tabelo, izgubimo želeno relacijo t.j. ena proti mnogo med dimenzijsko tabelo in tabelo dejstev, zato lahko alternativno uvedemo polja s trenutno in prejšnjo vrednostjo, a po drugi strani uporaba teh doprinese le probleme pri izvajanju poizvedb. Odločitev za katero od navedenih možnosti se bomo odločili pri obravnavi počasi spreminjajočih se dimenzij, je tako odvisna predvsem poznavanja zahtev in pomena podatkov za končne uporabnike.

### 5.3.6. Mini-dimenzije

Izvorne podatke, ki so se nanašajo na osebe lahko v grobem razdelimo na: neposredne in posredne podatke. Pri posrednih podatkih imamo v mislih podatke, ki posredno opredeljujejo lastnosti posamezne osebe zaradi obravnave posameznega kaznivega dejanja zaradi katerega je bila ta oseba obravnavana (napr.: *Oseba deluje v organizirani kriminalni združbi*). Za lažje razumevanje si pogledjmo spodnjo tabelo 5, na kateri vam prikazujemo vrste podatkov, in uporabljene dimenzije, ki se nanašajo na osebe.

**Tabela 5: Neposredni in posredni podatki, ki se nanašajo na osebe**

Podatki, ki se nanašajo na osebe	Povezava	Nastale dimenzije
Identifikacijska številka osebe	neposredna	DIMOsebe
Državljanstvo		DIMDemografski
Spol		
Starost	posredna	DIMMotiv
Prebivališče		DIMPrizorišče
Motiv		
Organizirana združba		
Prizorišče		

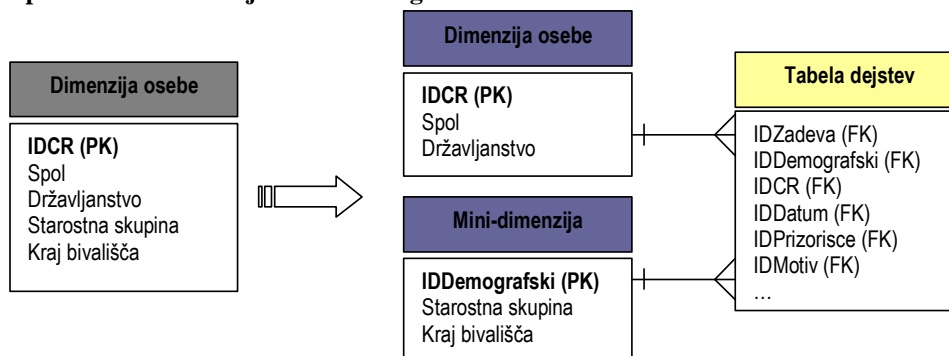
Kot lahko opazite iz tabele 5, smo zaradi pričakovanih problemov posodabljanja podatkov, ki se nanašajo na osebe, postavili več dimenzij. Pri določanju dimenzij smo se odločili, da posredne podatke, ki se nanašajo na osebe, opredelimo kot samostojne dimenzije (motiv in prizorišče), medtem, ko je pri opredelitvi neposrednih podatkov, ki se nanašajo na osebe, obstajalo več alternativ:

- podatke lahko združimo in jih opredelimo v eni dimenziji (pri starosti to ne bi predstavljalo problema, saj lahko uporabimo izračunana polja, medtem, ko bi pri atributu prebivališče izgubili zgodovinske podatke);
- pripravimo dve neodvisni dimenziji t.j. starost in prebivališče (kar bi po drugi strani pomenilo, da se že preveč zatekamo k de-normalizaciji tabele dejstev);
- pripravimo pomožno tabelo, ki predstavlja vez med tabelo dejstev in dimenzijo oseb (v pomožni tabeli bi bili podatki o starosti, prebivališču ter parametra, ki označujeta časovni začetek in konec veljavnosti te relacije);
- pogosto spreminjajoče attribute ločimo iz dimenzije in jih dodamo v skupno novo nastalo dimenzijo.

Izbrali smo zadnjo možnost. Zakaj? Podatki kot so: identifikacijska številka osebe, spol in državljanstvo osebe, se le izjemoma spreminjajo, medtem, ko za starost in prebivališče osebe ne moremo trditi podobno. Ker smo želeli obdržati razmerje 1:N nasproti tabele dejstev in obdržati zgodovinske podatke (tako, da jih dodajamo kot nove zapise), se je rešitev ponujala v vpeljavi t.i. mini dimenzije (angl. Mini dimensions).

Kot lahko razberete iz slike 10, na kateri vam prikazujemo pripravo mini-dimenzije (DIMDemografski), tovrstno dimenzijo pripravimo tako, da pogosto spreminjajoče podatke iz obravnavane dimenzije izločimo in jih skupno dodamo v novo nastalo mini-dimenzijo (Kimball 2002, str. 156). Tej novi dimenziji dodamo nov ključ in ne vključujemo ključev iz stare dimenzije (v našem primeru bi to bila identifikacijska številka osebe).

Slika 10: Priprava mini dimenzije »DIMDemografski«



Iz zgornjega prikaza lahko prav tako opazite, da je povezava med mini-dimenzijo in dimenzijo oseb le posredna, t.j. preko tabele dejstev. Občasne spremembe demografskih podatkov (kot sta: starost in kraj bivališča) oseb, tako neposredno več ne vplivajo na dimenzijo oseb.

### 5.3.7. Dimenzije brez atributov

Ob postavljanju dimenzijskih tabel včasih naletimo na dimenzije, ki poleg tujega ključa v tabeli dejstev nimajo dodatnih atributov in dejansko sploh ne potrebujejo svoje dimenzijske tabele. Tovrstno dimenzijo imenujemo brez atributna dimenzija (angl. degenerate dimensions - DD).

Večkrat se tak primer pojavi, ko imamo opravka z naročili ali računi, kajti številka računa se v klasičnem relacijskem podatkovnem modelu uporablja kot ključ oz. nosilec posamezne transakcije naročila (ki vsebuje attribute: datum, kupec, izdelek, količina, znesek, itd.). Vendar, ker ob modeliranju večdimenzionalnega podatkovnega modela izločimo tovrstne attribute v samostojne dimenzije, sama številka računa nima več svoje prvenstvene vloge. Služi nam lahko le pri grupiranju dejstev, ali pa kot primarni ključ v tabeli dejstev (seveda, če le ta ni sestavljen iz tujih ključev dimenzijskih tabel) (Kimball 2002, str. 50).

V našem primeru je bil predstavnik tovrstne dimenzije številka zadeve, katero smo uporabljali kot grupirani ključ, na podlagi katerega smo lahko pregledovali osebe, ki so bile obravnavane v isti zadevi. Vendar, ker nas z analitičnega vidika seveda niso zanimale posamezne zadeve oz. osebe, smo tovrstno dimenzijo razširili z dodatnimi izpeljanimi atributi, ki so se nanašali na posamezno zadevo. Uporabili smo torej pristop združenih dejstev kot atributov, ki vam ga prikazujemo v nadaljevanju.

### 5.3.8. Uporaba združenih dejstev kot atributov

Kot smo že dejali, je lahko v posamezni zadevi obravnavano več oseb, zaradi različnih vrst kaznivih dejanj ali prekrškov, zaseženo pa je lahko tudi več vrst prepovedanih drog, da bi lahko uporabniki ob tako postavljenem modelu na enostaven način dostopali do podatkov, ki se nanašajo samo na zadeve, kjer so bili zaseženi pravi repertoarji prepovedanih drog ali obravnavano več oseb, smo uporabili pristop združenih dejstev kot atributov (ang. Aggregated Facts as Attributes) (Kimball, 2002, str. 152).

Priprava združenih dejstev kot atributov, se tako nanaša na pripravo deskriptivnih vrednosti iz posameznih uporabljenih mer. Kot lahko opazimo iz spodnje tabele 6, smo na podlagi številke zadeve združevali uporabljene mere (kot sta: število oseb in število zasegov) in jih opredelili kot deskriptivne vrednosti posamezne zadeve.

Tabela 6: Uporaba združenih dejstev kot atributov pri dimenziji »DIMZadeva«

IDZadeva	Število_oseb	Število_zasego v	Zaseg	Prepovedane_d roge
5	do 01	NI	0	1
50	od 02 do 03	NI	0	0
55	od 04 do 06	NI	0	0
60	od 02 do 03	od 03 do 05	1	1
81	od 04 do 06	NI	0	1
108	do 01	NI	0	0
....				

Iz tabele 6 lahko prav tako opazimo uporabo atributov zaseg in prepovedane droge, s katerima smo omejevali poglede na podatke. Omejevanje pogledov na podatke je bilo potrebno, saj je tabela dejstev vsebovala podatke za vsa pretekla dejanja, zaradi katerih so bile posamezne osebe obravnavane. Z uporabo slednjih atributov so lahko uporabniki na enostaven način izločili vsa ostala kazniva dejanja, ki niso imela neposredne zveze z zasegom prepovedane droge ali obravnavo zaradi nje.

V klasični relacijski bazi se vsekakor ne bi zatekali k shranjevanju izpeljanih polj, zaradi skladnosti z izvirnimi podatki. Torej, v kolikor pride do spremembe vrednosti v tabeli dejstev moramo prav tako dosledno posodobiti shranjene izpeljane vrednosti v vezani dimenziji, saj bi drugače prišli do nepravilnega združevanja. Zaradi lažjega posodabljanja smo kot ključ povezave ohranili številko zadeve, ki nam omogoča enostavno spremembo (oz. ponoven izračun) atributov v vezani dimenziji (napr.: *Ob zaznavi kaznivega dejanja je bila obravnavana le ena oseba, tekom preiskave pa se ugotovi, da je v zadevo dejansko vključeno več oseb. V vezani dimenziji, bi se tako morale deskriptivne vrednosti popraviti.*)

Zaradi problemov posodabljanja, da uporabo tovrstnih atributov omejimo samo na tiste attribute, ki so najpogosteje uporabljeni in te ne uporabljamo za numerične kalkulacije. Torej vrednosti teh naj služijo zgolj v deskriptivne namene (Kimball 2002, str. 152).

## **5.4. Izgradnja OLAP kocke**

### **5.4.1. Microsoft SQL Server 2000 Analysis Services**

Ob vstopu Microsofta na trg OLAP z svojim izdelkom Microsoft SQL Server 2000 Analysis Services (v nadaljevanju besedila: Analysis Services), lahko rečemo, da se je v zgodovini sprotne analitične obdelave podatkov pričelo novo poglavje, kajti z Analysis Services je postala zmožnost sprotne analitične obdelave podatkov na osnovi večdimenzionalnih baz podatkov dosegljiva vsakomur in ne le več ozkemu krogu dovolj bogatih in pogumnih izbrancev (Ferle, 2001b).

Ker je Analysis Services del relacijskega strežnika Microsoft SQL Server 2000 je štetje uporabnikov Analysis Services vprašljivo, saj je težko ugotoviti, ali so uporabniki kupili strežnik Microsoft SQL Server zaradi relacijske ali večdimenzionalne baze podatkov. Vendar strokovnjaki hitro rast tržnega deleža Analysis Services na področju OLAP storitev pripisujejo čedalje bolj pogosti uporabi strežnika Microsoft SQL Server 2000 za v namene poslovne inteligence. Ne glede na število dejanskih uporabnikov je Microsoft s svojim izdelkom vplival na znižanje cen izdelkov konkurenčnih izdelovalcev (Ferle, 2001b).

Analysis Services vsebuje srednje nivojski (angl. middle tier) strežnik, ki je zgrajen za namene upravljanja večdimenzionalnih podatkovnih struktur in zagotavljanju hitrega dostopa do podatkov. Analysis Services organizira podatke iz podatkovnih virov v t.i. OLAP kocke oz. večdimenzionalne baz podatkov. Analysis Services prav tako omogoča kreiranje modelov za podatkovno rudarjenje na podlagi relacijskih kakor tudi večdimenzionalnih baz podatkov (Microsoft, 2005).

Delo s strežnikom Analysis Services poteka preko orodja Analysis Manager. Čeprav je Analysis Services del podatkovne zbirke Microsoft SQL Server 2000, lahko kot vir podatkov uporabimo tudi druge baze podatkov (npr. Oracle, DB2, ali katerokoli drugo, ki je dostopna preko standardnega vmesnika ODBC). Osnove izdelovanja OLAP kock ne zahtevajo veliko zahtevnega znanja, kajti orodje Analysis Manager ima vgrajene številne čarovnike, ki bistveno poenostavijo njihovo izdelavo. Pri upravljanju s strežnikom obstajajo prav tako različna orodja za nadzor nad delovanjem strežnika, do različnih orodij za shranjevanje podatkov, kreiranje uporabnikov, zaščite in optimizacije njegovega delovanja. Z vgrajenim pregledovalnikom podatkov v Analysis Services lahko podatke v podatkovnih kockah poljubno pregledujete, režete, raziskujete njihovo globino in jih omejujete glede na različne dimenzije.

Na strani uporabniških zahtev pa obstaja prav tako mnogo orodij, ki so sposobna črpati podatke iz Analysis Serverja in jih prikazati uporabniku na prijazen način. Tovrstna orodja uporabljajo Microsoftov OLE DB gonilnik za OLAP 8.0 (Microsoft, 2005). Analysis Services odlikujejo predvsem možnost nadgradnje, preprostost upravljanja in hitra odzivnost na zapletene analitične poizvedbe.

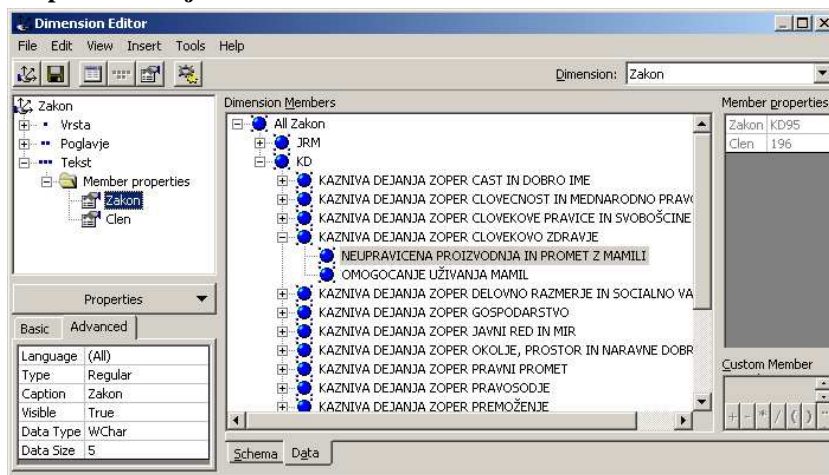
## 5.4.2. Definiranje večdimenzionalne strukture OLAP kocke

Vhodni podatki za sistem OLAP se lahko nahajajo v podatkovnem skladišču, področnih skladiščih ali v transakcijskih bazah podatkov. Ob tem ni pomembno ali so njihove podatkovne strukture normalizirane ali de-normalizirane, pomembno je le, da vsebujejo podatke o vseh merah in dimenzijah, ki jih potrebuje sistem OLAP. Predhodno izdelani večdimenzionalni podatkovni modeli (zvezdne sheme) v relacijski zbirki podatkov torej niso predpogoj za pravilno izdelavo OLAP kock.

Pred samo izgradnjo OLAP kocke je potrebno (v našem primeru: ponovno) definirati strukturo večdimenzionalnega podatkovnega modela. Pri tem moramo opozoriti, da pri tehnikah večdimenzionalnega modeliranja podatkov v klasičnih relacijskih bazah podatkov, praviloma uporabljamo več dimenzij, nivojev in mer, saj želimo s postavitvijo modela zagotoviti prilagodljivost modela. Ker pa so podatki v posamezni kocki namenjeni za specifične potrebe uporabnikov, je potrebno točno vedeti, katere odgovore se želi dobiti z izgradnjo določene kocke. Iz posamezne zvezdne sheme, lahko tako pripravimo več različnih OLAP kock, ki se razlikujejo po uporabljenih dimenzijah, nivojev podrobnosti in mer, ki pomagajo uporabnikom pri odločanju ali analiziranju.

Pri postavljanju dimenzij v Analysis Services ločimo: privatne dimenzije (angl. Private Dimensions), ki jih uporabljamo le v posameznih kockah in skupne dimenzije (angl. Shared Dimensions), ki so skupne v vsem kockam v sklopu določene zbirke večdimenzionalnega podatkovnega modela. Skupne dimenzije v Analysis Services pripravljamo z orodjem »Dimension Editor«, ki vam ga prikazujemo na sliki 10.

Slika 11: Prikaz skupne dimenzije »Zakon«



Vir: Ekranska slika, Microsoft Analysis Manager, Dimension Editor

Iz zgornjega prikaza lahko razberemo, da je skupna dimenzija »Zakon«, sestavljena iz treh nivojev in sicer: vrsta dejanja (kaznivo dejanje ali prekršek), poglavje v zakonu in tekstualni zapis posamezna člena v zakonu (z opredeljenima lastnostima: ime zakona in številka člena). Za potrebe analiziranja podatkov smo tako pripravili več skupnih dimenzij in znotraj posameznih kock določili dodatne privatne dimenzije.

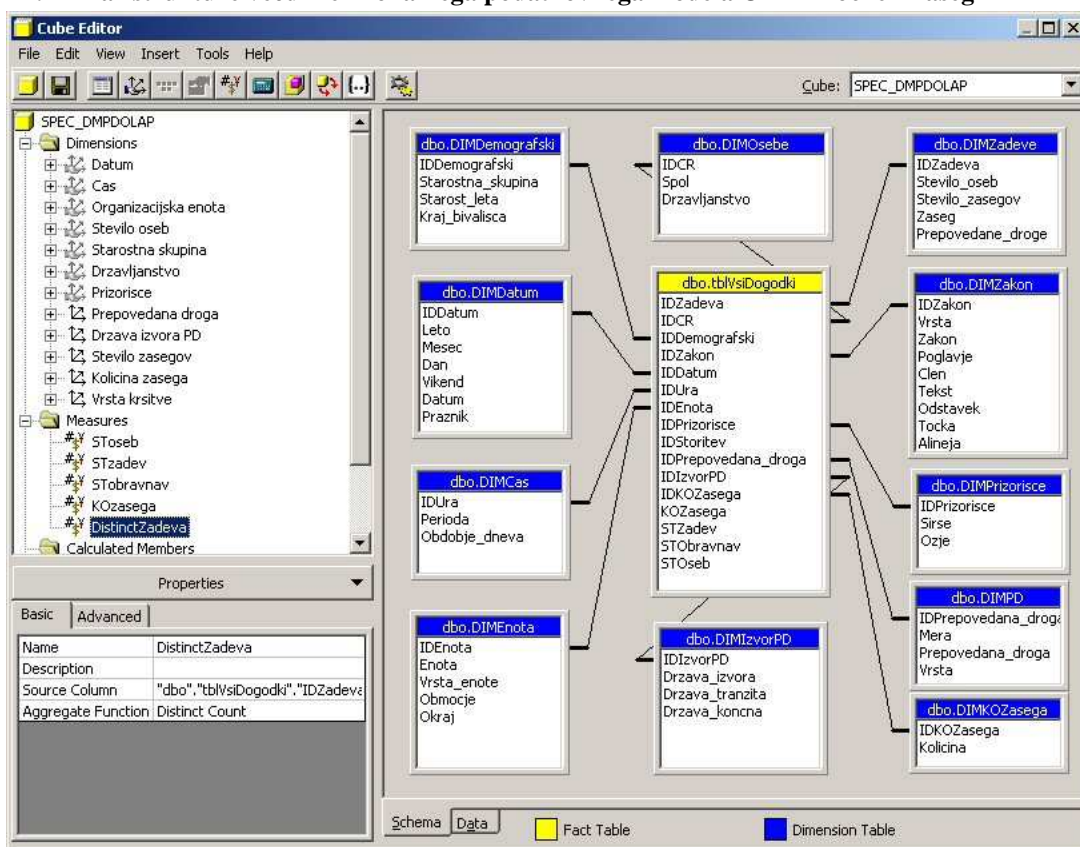


Na podlagi večdimenzionalnega podatkovnega modela, pripravljenega v relacijski zbirki podatkov, smo izdelali več OLAP kock, katerih priprava je bila usmerjena v analiziranje: zasegov prepovedanih drog; zaznav kaznivih dejanj in prekrškov povezanih s prepovedanimi drogami; zaznav ostalih kaznivih dejanj in prekrškov, s strani oseb, ki so bile obravnavane zaradi prepovedanih drog; in povezanost izvajanja posameznih kaznivih dejanj in prekrškov.

Za dostop do podatkov s stani uporabnikov, ki niso neposredno sodelovali pri analiziranju podatkov za potrebe strateške analize »Problematika prekomerne uporabe prepovedanih drog v RS«, smo prav tako pripravili enotno podatkovno kocko, v kateri so bile opredeljene vse dimenzije, vendar z omejenimi nivoji podrobnosti.

V Analysis Services pripravljamo večdimenzionalno strukturo podatkov OLAP kocke z orodjem »Cube Editor«. Ko iz vira podatkov izberemo ustrezno tabelo dejstev, jo orodje obarva z rumeno barvo, medtem, ko izbrane tabele dimenzij obarva z modro barvo. To lahko opazimo iz spodnje slike 11, kjer je na desni polovici slike prikazana struktura večdimenzionalnega podatkovnega modela OLAP kocke »Zasegi«.

Slika 12: Prikaz strukture večdimenzionalnega podatkovnega modela OLAP kocke »Zasegi«



Vir: Ekranska slika, Microsoft Analysis Manager, Cube Editor

Na levi polovici slike 11, pa so prikazane: skupne dimenzije (datum, čas, organizacijska enota, število oseb, starostna skupina, državljanstvo oseb, prizorišče); privatne dimenzije (prepovedana droga, država izvora prepovedane droge, število zasegov, količina zasega, vrsta kršitve); in uporabljene mere (število oseb, zadev, obravnav in količina zasega).

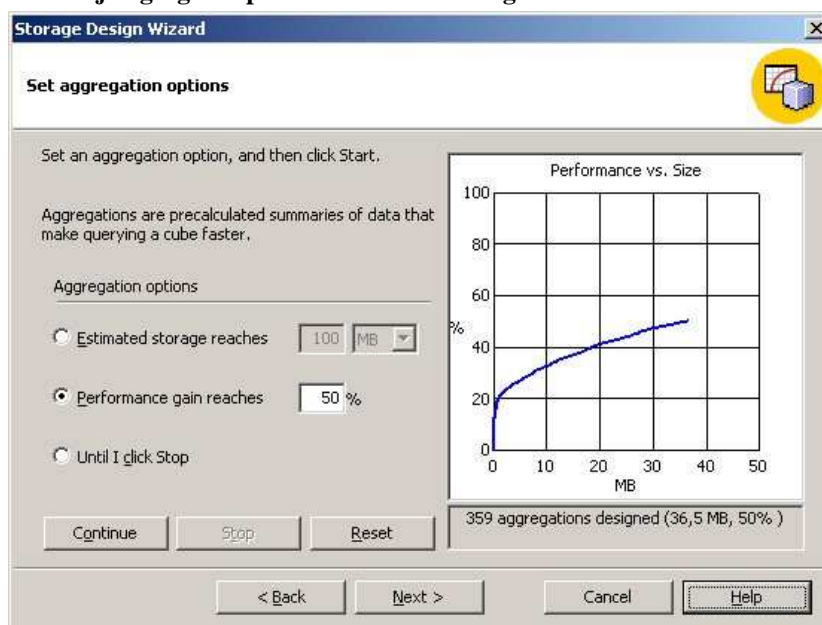
### 5.4.3. Procesiranje kocke in izbira ustreznega načina shranjevanja

Sistem izgradnje OLAP kocke deluje tako, da zajame podatke iz izbranega vira in na podlagi definirane strukture vnaprej izračuna njihove agregate. Temu postopku pravimo tudi procesiranje kocke. Namen agregatov je predvsem v skrajšanju odzivnega časa pri pregledovanju podatkov. Strežniku tako ni potrebno uporabiti vseh osnovnih podatkov za preračunavanje, ampak uporabi že predhodno izračunane podatke.

Načeloma več agregatov pomeni večjo hitrost sistema, vendar če bi hoteli vnaprej izdelati vse agregate, ki so možni, bi prišlo do ogromnega povečanja zahtev po prostoru na disku in procesorske moči. Če upoštevamo, da se ob tem izvede še množica izračunov, je jasno, da večdimenzionalna baza postane zelo obsežna, ob tem pa vsebuje relativno malo podatkov. Govorimo o t.i. eksploziji podatkov. Vendar je danes pri uporabi naprednih algoritmov za kreiranje agregatov, stiskanje podatkov in shranjevanju le celic s podatki, bojazen glede eksplozije podatkov, odvečna (Šmid, 2002, str. 36).

V Analysis Services, lahko uporabniki sami vplivajo na algoritem kreiranja agregatov, s parametri, ki so navedeni v čarovniku »Storage Design Wizard«. Iz slike 12 je lepo razvidno, da je algoritem za kreiranje optimalnega števila agregatov izračunal, da je za konkreten primer OLAP kocke »Zasegi« potrebno izdelati 359 agregatov, ki bodo zasedli 36,5 MB prostora na disku, oziroma, da se doseže želeno 50 % pospešitev poizvedb.

Slika 13: Primer kreiranja agregatov pri OLAP kocki »Zasegi«



Vir: Ekranska slika čarovnika »Storage Design Wizard«, Analysis Services

Pri procesiranju kocke imamo tako opravka z dvema tipoma podatkov. To so vhodni podatki in agregacije. Glede na kakšen način bomo te shranili, razlikujemo MOLAP, ROLAP in HOLAP.

Kako izbira posameznega načina shranjevanja vpliva na čas procesiranja, prostor shranjevanja in hitrost pregledovanja, si pogledajmo posamezne značilnosti teh:

- *MOLAP* (angl. Multidimensional OLAP) shranjuje oba tipa podatkov v večdimenzionalnih strukturah. Kot smo spoznali je to optimizirana struktura za izvajanje poizvedb in večdimenzionalne analize podatkov. Kljub temu pa lahko štejemo kot slabost, da morajo biti isti izvorni podatki zapisani dvakrat, enkrat v relacijski, drugič v večdimenzionalni obliki. Ob veliki količini podatkov, ga tako odraža večja poraba časa za procesiranje kock in velika poraba prostora za shranjevanje podrobnih in agregatnih podatkov v večdimenzionalni strukturi (Šmid, 2002, str. 40). Ker ob pregledovanju, vir podatkov ni več potreben, nam lahko omogoča delo na drugih lokacijah (torej na računalnikih, ki niso neposredno povezani z omrežjem organizacije).
- *ROLAP* (angl. Relational OLAP) shranjuje oba tipa podatkov v relacijski podatkovni bazi. Ker se v ozadju pregledovanj izvaja poizvedovalni jezik SQL je hitrost poizvedb odvisna od postavljenih indeksov, nivoja predhodnega združevanja. V povprečju je odziven čas slabši kot pri orodjih MOLAP. Torej dolg odzivni čas in neučinkovita izraba zmogljivosti sistemov OLAP so njegove karakteristike (Aličehajič et. al., 2001). Vendar je po drug strani čas procesiranja občutno krajši kot pri MOLAP, kajti procesiranje vhodnih podatkov ni potrebno izvrševati. Praktično to pomeni, da ko se podatek zapiše v podatkovno skladišče, je že predstavljen v OLAP odjemalcu. Temu načinu pravimo tudi kocka v realnem času (angl. real-time cube). Vendar mora sistem kljub temu zagotoviti, da se agregacije osvežujejo pri vsaki spremembi vhodnih podatkov (Roblek, 2003, str. 324).
- *HOLAP* (angl. Hybrid OLAP) shranjuje vhodne podatke v relacijski bazi, agregacije pa v večdimenzionalni strukturi. Podatke torej upravlja kot ROLAP, medtem, ko agregate kot MOLAP. Torej če pregledujemo samo agregate, nam omogoča enako hitrost pregledovanja kot MOLAP, hkrati pa prihrani čas osveževanja kock in prostor (Peterson, Pinkelman, 2000; po Babič, str. 17).

Pri izgradnji OLAP kock z orodjem Analysis Services prav tako ločimo tri načine procesiranja kocke: polno procesiranje, osveževanje kocke in inkrementalno procesiranje kocke. Pri *polnem* procesiranju ponovno izgradimo celotno kocko. Tak način običajno izberemo, ko se je spremenila struktura kocke (nova dimenzija, nivo ali mera). *Osveževanje* podatkov poteka tako, da se ponovno opravi izračun kocke brez kreiranja strukture. Izbor te opcije je primeren za primer, ko se podatki v skladišču podatkov spremenijo, ne spremeni pa se struktura kocke. Pri *inkrementalnem* procesiranju pa se podatki enostavno samo dodajo k obstoječim, zato se procesirajo novo zapisani podatki oziroma tisti podatki, ki izpolnjujejo določen pogoj (Roblek, 2003, str. 324).

Pri izbiri načina shranjevanja smo se odločili za MOLAP način shranjevanja.

#### 5.4.4. Izbira OLAP odjemalca

Odjemalec OLAP je zelo pomemben člen sistema OLAP in je z vidika uporabnika celo najpomembnejši, saj mu ta omogoča uporabo sistema in dostop do zelenih podatkov. Tako se nam lahko pripeti, da imamo še tako zmogljiv sistem OLAP, ki je s strani uporabnikov neuporabljen, saj so njegovi odjemalci za uporabnika neprijazni (Babič 2002, str. 26). Uporabnik mora tako z enim do treh klikov priti do področja, ki ga v dani situaciji zanima. V nasprotnem primeru je velika verjetnost, da uporabnik enostavno ne bo uporabljal sistema, bodisi zaradi tega ker nima časa ali pa ker se mu enostavno ne ljubi, saj ima že dovolj drugih zahtevnih opravil.

Zato so za OLAP odjemalce značilni relativno enostavni in intuitivni uporabniški vmesniki, ki zajemajo zanimive načine interaktivnih predstavitev. Uporabniki sistema OLAP se lahko njihove uporabe naučijo hitro in ne izgubljajo časa s podrobnostmi, če jih le-te ne zanimajo. Nekatera odjemalska orodja pa prav tako omogočajo zahtevnim uporabnikom veliko svobode in odlične zmogljivosti za podrobnejše in raznolike analize podatkov.

Najbolj razširjeni odjemalci so: ProClarity, Cognos PowerPlay, HuggyDog, IntelliBrowser, Microsoft Excel, Crystal Reports, Office Web Components, itd. Uporabimo lahko tudi razna razvojna orodja, ki omogočajo razvoj lastnih aplikacij za analizo podatkov (Roblek, 2003, str. 327).

Pri izbiri odjemalca tako obstaja več možnosti, ki imajo svoje prednosti in slabosti, vsem pa je skupno, da mora OLAP odjemalec zadostiti naslednjim kriterijem:

- biti mora preprost za uporabo;
- zagotoviti mora intuitivno metodo izdelovanja novih poročil;
- podpirati mora grafično prikazovanje podatkov;
- zagotoviti izvajanje naslednjih operacij: združevanje, vrtenje v globino, filtriranje, listanje, omejevanje, vrtenje dimenzij itd.

Microsoft Excel (v nadaljevanju besedila Excel) z uporabo vrtilnih tabel (angl. pivot table), grafičnih prikazov in ostalih funkcij, omogoča analizo podatkov, shranjenih v centralni podatkovni zbirki, hkrati pa lahko nastopa tudi v vlogi OLAP odjemalca za Analysis Services. Za razliko od navadnih vrtilnih tabel, ki bi jih oblikovali neposredno iz podatkov na delovnem listu, omogočajo OLAP vrtilne tabele obdelavo večjih količin podatkov. Kajti OLAP strežnik je zadolžen za obdelave podatkov, medtem ko sam Excel skrbi za prikazovanje teh. Tak pristop omogoča delo z veliko večjimi količinami izvornih podatkov, kot bi lahko, če bi bili podatki organizirani v tradicionalni zbirki podatkov, kjer Excel mora pridobiti vsak posamezni zapis.

Excel pri dostopu do podatkov OLAP uporablja tehnologijo OLE DB, ki je skupek vmesnikov za poseganje v večdimenzionalno strukturo in omogoča izvajanje MDX (angl. Multidimensional Expressions) poizvedb (Roblek, 2003, str. 327).

Podatke OLAP lahko v Excelu prikažete le kot poročilo vrtilne tabele ali vrtilnega grafikona. Poleg osnovnih funkcij, ki se zahtevajo s strani OLAP odjemalcev, pa lahko podatke prav tako poljubno oblikujete ali pa jih v izračunih uporabite kot katere koli druge podatke, ki bi jih uporabljali v Excelu. Uporabo Excela kot odjemalca OLAP podatkov, vam prikazujemo na sliki 13.

Slika 14: Uporaba programa Microsoft Excel kot OLAP odjemalca

	A	B	C	D	E	F	G	H
1	Starostna skupina	(Multiple Items)						
2	Cas	ZJUTRAJ						
3	Kolicina zasega	(Multiple Items)						
4	Prepovedana Droga	KANABIS RASTLINA (M...	HUANA)					
5								
6	STzadev		Leto					
7	Prizorisce	Prizorisce ozje	2000	2001	2002	2003	2004	Grand Total
8	CESTA							
9	GOSTINSKI OBJEKT							
10	JAVNA PRIREDITEV							
11	NASTANITVENI PROSTOR							
12	OSTALO							
13	POSLOVNI PROSTOR							
14	PREVOZNA SREDSTVA							
15	ŠOLA IN ŠPORT	PROST. ZA VZGOJNOVAR.						
16		PROSTOR ZA ŠPORT, REK...						
17	ŠOLA IN ŠPORT Total							
18	Grand Total							
19								
20								
21								
22								
23								
24								
25								

Vir: Ekranska slika, Microsoft Excel 2003, Pivot table

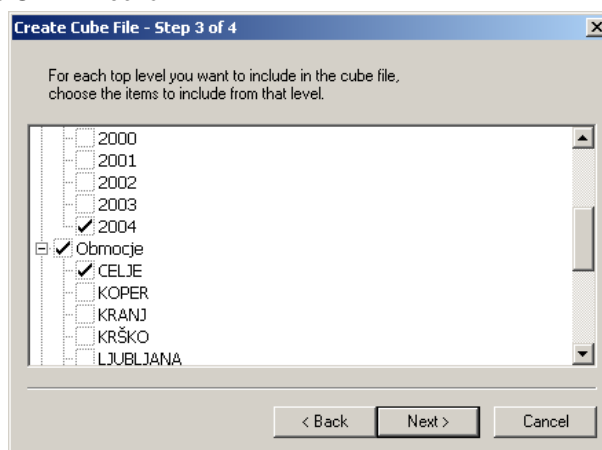
Na zgornji sliki 13, vam prikazujemo uporabo vrtilne tabele pri dostopu OLAP kocke »Zasegi«. Ker na prikazu niso razvidni vsi postavljeni pogoji, povejmo, da smo s postavitvijo vrtilne tabele želeli odgovoriti na vprašanje: »Ali je zaznati povečano število manjših zasegov prepovedane droge »marihuana« s strani mladoletnih oseb, v jutranjih urah na območju šole?»

Prednost, ki jo prinaša uporaba Excela je možnost priprave lokalne kocke oz. kockovne datoteke brez povezave (angl. Offline Cube File). Kockovna datoteka brez povezave (v nadaljevanju besedila: lokalna kocka) je datoteka, ki je namenjena za shranjevanje izvornih podatkov iz OLAP strežnika. Uporaba tovrstne datoteke je primerna za distribucijo OLAP podatkov in njihovo uporabo s strani uporabnikov, ki pregledujejo podatke brez povezave z Analysis Services strežnikom.

Uporaba lokalne kocke lahko prav tako pospeši pregledovanje podatkov v primerih, ko so omrežne povezave s strežnikom OLAP počasne. Da pa ne bi bilo potrebno skrbeti za obilico prostora v primeru velikih OLAP zbirk, lahko pri izgradnji lokalne kocke poljubno izbiramo dimenzije in nivoje podrobnosti, ki jih želimo vključiti v datoteko.

Pri pripravi lokalne kocke imamo prav tako možnost izbire posameznih vrednosti vrhnjih nivojev dimenzij, ki jih želimo shraniti. Tovrstna možnost je zelo uporabna pri distribuciji podatkov OLAP. Kot lahko razberemo iz slike 14, na kateri vam prikazujemo eno izmed ekranskih oken »Čarovnika za pripravo lokalne kocke«, lahko obseg shranjenih vrednosti na enostaven način omejimo, tako da odkljukamo posamezne vrednosti, ki jih ne želimo shranjevati. Iz prikazanega primera na sliki 14, bi tako pripravljena lokalna OLAP kocka, vsebovala samo podatke, ki se nanašajo na obdobje v letu 2004 in na območje Celja.

**Slika 15: Priprava lokalne OLAP kocke**



Vir: Ekranska slika, Microsoft Excel 2003, Čarovnik za pripravo lokalne kocke

Uporaba Excela nam prav tako omogoča objavo interaktivnih spletnih strani, katerih glavna komponenta je vrtilna tabela, ki omogoča prikaz podatkov v spletnem odjemalcu z uporabo gradnikov Office Web Components. V kolikor pa bi želeli, da uporabniki samo vidijo podatke, ki jih objavimo, lahko podatke objavite kot ne-interaktivne podatke. Ob tem naj še opozorimo, da interaktivnih spletnih strani ne moremo odpreti in spreminjati z Excelom, zato je dobro, da shranimo kopijo izvirnega delovnega zvezka, iz katerega je bila pripravljena interaktivna stran, za primer, če bi želeli kaj spremeniti in ponovno objaviti.

Ker je bila izdelava sistema OLAP zgolj v testne namene, nismo razmišljali o izdelavi samostojne aplikacije in/ali nakupu že izdelane aplikacije od specializiranega podjetja, temveč smo za OLAP odjemalec uporabili rešitev, ki jo ponuja Excel in Office Web Components pri dostopu do vira OLAP podatkov. Odločitev je bila podprta ne samo iz cenovnega vidika temveč tudi iz razloga, ker uporabniki relativno dobro poznajo programe paketa Office in je ta dostopen na vseh osebnih računalnikih v organizaciji. Ker uporabniki, ki so testirali dostop do podatkov niso imeli povezave do OLAP strežnika, smo za testiranje dostopa do podatkov, pripravili lokalne kocke, na podlagi katerih smo pripravili interaktivne spletne strani, katere smo objavili na skupnem strežniku. Uporabnikom smo tako preko brskalnika Microsoft Internet Explorer omogočili dostop do podatkov OLAP in njihovo analiziranje. Določene funkcionalnosti, ki jih sicer omogoča neposreden dostop do OLAP strežnika, smo s tem izgubili, vendar je bil po drugi strani tak spletni pregled za marsikoga povsem zadovoljiv.

#### 5.4.5. Matrika povezanosti podatkov

Uživalci prepovedanih drog se vsakodnevno srečujejo s problemom kako priti do droge, ki je njihova vsakdanja potreba. Za nabavo prepovedanih drog pa so potrebne večje vsote denarja in eden od razlogov, da nekdo začne s kriminalom, so finančni problemi oziroma pomanjkanje denarja za nabavo droge. Uživalci prepovedanih drog se tako zatekajo raznim oblikam tako imenovane sekundarne ali indirektno kriminalitete kot so vlomi, ropi, tatvine, izsiljevanja, prostitucija, pornografija, itd.

Sekundarna kriminaliteta je statistično težko merljiva, saj pri evidentiranju kaznivih dejanj v policiji ne razlikujemo, oz. policija ne deli kaznivih dejanj na tista, ki jih storijo uživalci prepovedanih drog in tista, ki jih storijo drugi osumljenci. Osumljenci - odvisniki pa prav tako redko priznajo oz. označijo storjeno dejanje, kot oskrbovalni delikt (zaradi nakupa prepovedanih drog), kar je razumljivo, saj jih je strah, da bi morali spregovoriti tudi o svojih oskrbovalcih z drogo.

Da bi lahko na enostaven način prišli do odgovorov, ki kažejo na povezanost podatkov oz. izvajanj posameznih kaznivih dejanj ali prekrškov, smo pripravili tudi t.i. *povezovalno matriko*. Res, da takšen postopek priprave kot ga bomo opisali ni primeren za velike količine podatkov ali v primerih, ko razpolagamo s programskimi orodji namenjenimi za statistično obdelavo podatkov (kot je SPSS), vendar po drugi strani na dokaj enostaven način pridobimo podatke o povezanosti podatkov. Za postavitev matrike smo:

- omejili število dejanj na najpogostejša, ki skupno predstavljajo več kot 90% vseh obravnavanih dejanj,
- odstranili vsa nepotrebna podvajanja, saj je lahko posamezna oseba za isto dejanje obravnavana večkrat,
- da bi vzpostavili neposredno vez med dejanji, zaradi katerih je bila posamezna oseba obravnavana, smo pripravili tabelo povezav med dejanji. Primer tovrstne tabele vam prikazujemo v tabeli 7. V tabeli so tako opredeljene vse možne kombinacije med posameznimi dejanji, zaradi katerih je bila obravnavana posamezna oseba. Torej, če je bila ta obravnavana zaradi treh dejanj, imamo opredeljenih devet povezav.

**Tabela 7: Opredelitev povezav med dejanji znotraj posamezne osebe**

IDCR	Krsitev_A	Krsitev_B
2060234	MOTENJE NOČNEGA MIRU	MOTENJE NOČNEGA MIRU
2060234	MOTENJE NOČNEGA MIRU	POSEDUJE PREPOVEDANE DROGE
2060234	MOTENJE NOČNEGA MIRU	PREPIRANJE, VPITJE
2060234	POSEDUJE PREPOVEDANE DROGE	MOTENJE NOČNEGA MIRU
2060234	POSEDUJE PREPOVEDANE DROGE	POSEDUJE PREPOVEDANE DROGE
2060234	POSEDUJE PREPOVEDANE DROGE	PREPIRANJE, VPITJE
2060234	PREPIRANJE, VPITJE	MOTENJE NOČNEGA MIRU
2060234	PREPIRANJE, VPITJE	POSEDUJE PREPOVEDANE DROGE
2060234	PREPIRANJE, VPITJE	PREPIRANJE, VPITJE

- iz tako pripravljene tabele, smo pripravili OLAP kocko, pri kateri smo za dimenzije uporabili atributa: Kršitev\_A in Kršitev\_B, za mero pa smo izbrali DistinctCount identifikacijskih števil oseb.



Rezultat je spodnja matrika, katero vam prikazujemo na sliki 16. Iz matrike lahko razberemo, koliko posamičnih oseb je sodelovalo pri preseku posameznih dejanj. Sedaj si lahko zastavimo vprašanje kot je: »Koliko oseb se je nedostojno vedlo do uradne osebe in je bilo prav tako obravnavano zaradi motnja nočnega miru ali prepiranja in vpitja?«

Slika 16: Matrika povezanosti dejanj

	GOLJUFIJA	LAHKA TELESNA POŠKO	MOTENJE NOČNEGA MIR	NEDOST.VEDENJE DO UR	NEUPRAVIČENA PROIZV	OMOGOČANJE UŽIVANJA	POSEDUJE MANJŠO KOL	POSEDUJE PREPOVEDAN	POŠKODOVANJE TUJE S	PREPIRANJE, VPITJE	PRETEP, ŽALJIVO VEDEN	PRIKRIVANJE	TATVINA	VELIKA TATVINA	VZNEMR. OKOLJA V PIJAJ	Grand Total
GOLJUFIJA	676	138	104	142	258	94	478	235	121	365	157	138	418	305	54	676
LAHKA TELESNA POŠKODE	100	775	244	245	288	88	457	277	255	815	485	152	388	259	73	775
MOTENJE NOČNEGA MIRU	181	243	1558	418	488	212	545	588	343	543	248	228	628	485	138	1558
NEDOST.VEDENJE DO UR.O	142	245	418	1374	411	138	884	528	385	814	411	227	528	388	214	1374
NEUPRAVIČENA PROIZVOD	258	288	430	411	2983	408	888	718	382	1157	488	429	784	528	130	2983
OMOGOČANJE UŽIVANJA	94	88	212	158	488	1487	582	488	188	512	134	188	414	284	98	1487
POSEDUJE MANJŠO KOLIČ	478	457	545	888	388	8716	1245	138	2418	814	617	1725	1154	504	8716	
POSEDUJE PREPOVEDANE I	235	277	388	528	718	400	1245	4588	448	1428	588	388	882	652	267	4588
POŠKODOVANJE TUJE STV	121	255	343	388	302	188	738	448	1188	832	388	182	528	407	143	1188
PREPIRANJE, VPITJE	365	815	543	814	1157	512	3418	1428	822	2988	588	540	1411	882	410	2988
PRETEP, ŽALJIVO VEDENJE	157	485	348	411	448	134	814	528	380	888	1073	248	517	483	158	1073
PRIKRIVANJE	138	152	228	237	428	188	817	388	182	548	248	888	511	425	82	888
TATVINA	418	388	628	528	784	414	1725	882	528	1411	517	511	2677	1182	298	2677
VELIKA TATVINA	385	259	485	388	528	304	1154	852	427	882	483	425	1182	8748	130	1154
VZNEMR. OKOLJA V PIJAJ	84	78	138	214	138	98	584	287	143	418	153	82	288	138	188	788
Grand Total	676	775	1558	1374	2983	1487	8716	4588	1188	2988	1073	888	2677	1182	788	14285

Ker iz tako pripravljene matrike ne moremo razbrati osebe, ki so bile obravnavane zaradi prepovedanih drog več kot 5 krat, smo pripravili dimenzijo »DIMKršitev«, v kateri so bila prav tako opredeljena najpogostejša dejanja kot atributi posamezne osebe. Vendar so bile numerične vrednosti atributov (kolikokrat je bila posamezna oseba obravnavana zaradi določenega dejanja) spremenjene v deskriptivne vrednosti (intevalne vrednosti). Ker je pogostost izvajanja posameznega dejanja različna, smo glede na frekvenčno porazdelitev vrednosti, posameznim atributom postavili samo dve vrednosti (0 in 1+), ostalim pa več (0; 1; 2-4; 5+).

Torej, če k tako pripravljene matriki dodamo še dimenzijo kršitev, ki prikazuje deskriptivne vrednosti najpogostejših kaznivih dejanj posameznih oseb, lahko uporabniki sedaj na enostaven način zastavljajo bolj kompleksnejša vprašanja. Zgornji primer glede nedostojnega vedenja do uradne osebe lahko tako dopolnimo s vprašanjem: »Koliko je bilo oseb, ki so se vedle do uradne osebe nedostojno in so motile nočni red in mir, če so bile te obravnavane več kot tri krat zaradi vznemirjanja okolja v pijanem stanju?« V kolikor pa dodamo dimenzijo kršitev v OLAP kocko »Zasegi«, pa si lahko zastavimo vprašanja kot so: »Koliko je bilo zasegov prepovedanih droge s strani oseb, ki so bile obravnavane zaradi tatvin več kot pet krat? Ali preverili hipotezo, da so osebe katerim je bila zasežena prepovedana droga heroin pogosteje obravnavane zaradi tatvin kot ostale osebe«.



## 6. MOŽNOSTI IN OMEJITVE UPORABE SODOBNIH INFORMACIJSKIH TEHNOLOGIJ PRI ORGANIH PREGONA

### 6.1. Možnosti uporabe sodobnih informacijski tehnologij

V nadaljevanju si pogledjmo samo na nekatera delovna področja policije na področju prepovedanih drog, pri katerih menim, da bi ta z uporabo sistema OLAP v namene: okrepite dejavnosti policije usmerjene proti organiziranemu kriminalu in izboljšanju obveščevalnih in analitični del pri odkrivanju in preiskovanju kaznivih dejanj, lahko povečala učinkovitost in uspešnost njenega delovanja:

- dinamično spremljanje dogajanj na trgu ponudbe in povpraševanja po prepovedanih drogah. Povečanje aktivnosti policije na posameznih prizoriščih (napr. okolice šol in igrišč), vplivajo na preselitev deliktov na druga prizorišča. Policija lahko tako predvidi in načrtuje svoje aktivnosti na več prizoriščih hkrati;
- merjenje sekundarne kriminalitete na področju prepovedanih drog je izredno težavno. Zato lahko le ob dnevnem spremljanju dogodkov, ki so povezani s premoženjskimi delikti ugotovljamo njihov izvorni vzrok;
- s spremljanjem izvora prepovedanih drog in njihovega načina tihotapljenja in prikrievanja, lahko zagotovimo učinkovitejši nadzor na mejnih prehodih;
- celovito spremljanje dogodkov (koncerti) in aktivnosti, kjer prihaja do občasne uporabe »mehkih drog«, predvsem iz vidika preventivnega delovanja;
- sprejemanje in zgodnja prepoznavo novih sintetičnih drog. Številni smrtni primeri v svetu, samo potrjujejo, da so sintetične droge (ectasy) izjemno nevarne, saj so te narejene iz različnih kemikalij v skrivnih laboratorijih, katerih končni učinek ni vedno pričakovan;
- dinamični pregled nad razpoložljivimi finančnimi sredstvi, materialnimi sredstvi in kadri, bi znatno pripomogel k učinkovitejšemu načrtovanju, organiziranju in koriščenju resursov, ki jih ima policija na razpolago;
- na področju izvajanja strateških analiz, ki so usmerjene v opazovanje (preučevanje) kriminalnih skupin, njihovega načina dela, ocenjevanja njihovih namenov, zmožnosti, omejitev in ranljivosti. Kakor tudi v namene ocenjevanja lastne organizacije;
- in nenazadnje pri iskanju odgovorov na kompleksnejša vprašanja, ki prihajajo s strani ministrstva, državnega zbora in medijev, ki so v večini primerov zastavljena tako, da se kar kmalu soočimo s problemom kako iz vseh zbranih podatkov pridobiti koristne informacije.

Pri policijskem delovanju v boju z organiziranim kriminalom je osrednjega pomena obveščenost o dejavnostih kriminalnih skupin, kar pomeni konstantno zbiranje in vrednotenje podatkov in informacij. Operativno delo policije pri preiskovanju kaznivih dejanj povezanih s prepovedanimi drogami je tako temeljno in ga nikakor ni možno nadomestiti z drugim pristopi.

Na področju uporabe tehnologij podatkovnega rudarjenja nekatera podjetja zavzeto promovirajo uporabo podatkovnega rudarjenja v namene odkrivanja: bančnih goljufij, zavarovalniških, davčnih prevar in podobnih gospodarskih kaznivih dejanj. Če nekoliko podrobneje pogledamo način storitve teh dejanj, lahko opazimo, da storilci izkažejo lažne podatke na nek obrazec oz. vlogo, verodostojnost podatkov pa podkrepijo s ponarejenimi listinami. Uporaba podatkovnega rudarjenja pri odkrivanju tovrstnih goljufij in prevar je skorajda neizbežna, saj je zaradi količine vlog in omejitev časa, ki ga imajo organizacije na razpolago, izločitev sumljivih vlog izjemnega pomena. Na področju prepovedanih drog, bi lahko uporabili podobno razmišljanje pri carini, saj na podlagi podatkov, ki jih ti dobivajo od špediterjev (glede izvora, tranzita, končne destinacije, vrste, teže, vrednosti .... tovara) lahko učinkovito izločijo sumljive transporte, katere nadalje podrobneje pregledajo. Medtem, ko policija za vsako preiskavo (hišno ali osebno) mora pridobiti odredbo preiskovalnega sodnika. Za pridobitev odredbe pa je potrebno utemeljiti sum, da bo s preiskavo najden predmet kaznivega dejanja. Kako bi se končalo utemeljevanje razlogov za sum s pojasnjevanjem algoritmov podatkovnega rudarjenja si lahko samo predstavljamo.

Problemi in delo policije se ne sme končati pri identificiranju »malih rib« temveč oseb, ki predstavljajo vrh kriminalne združbe. Pred zaključkom akcije je tako potrebno z operativno taktičnimi in tehničnimi ukrepi zbrati ustrezna dokazna sredstva, ki bodo koristna za izvedbo kazenskega postopka. Menim, da lahko s pomočjo sodobnih informacijskih tehnologij, kot sta OLAP in podatkovno rudarjenje na področju pojasnjevanja vpletenosti posameznikov v organizirano kriminalno združbo in njihovega načina organiziranosti. Na področju ugotavljanja višine pridobljene protipravne premoženjske koristi ali prikrivanja izvora denarja nam je lahko v veliko pomoč OLAP tehnologija, saj lahko ob dobro postavljenem modelu v relativno kratkem času preverimo številne hipoteze, ki smo si jih zastavili.

Komunikacija med člani organizirane kriminalne združbe je še vedno njihova šibka točka. Pri analizi telefonskih pogovorov so poleg vsebine pogovora, prav tako pomembni atributi kot so: kraj klica, čas in sosledje pogovorov in dogodkov, ki so se pripetili. Ker je večina organiziranih kriminalnih združb organizirana hierarhično, na začetku preiskave policiji ni znana umeščenost posameznika v hierarhijo in s tem njegova vloga v kriminalni združbi. Uporaba podatkovnega rudarjenja pri ugotavljanju vzorcev klicev med člani združbe bi lahko veliko pripomogla pri preiskavi in pojasnjevanju njihove umeščenosti v združbo. Iskanje vzorcev sedaj poteka s pomočjo vizualizacije podatkov (program Analyst Notebook podjetja i2), vendar moramo pri tem opomniti, da je lahko takšno delo predvsem zamudno in nepregledno, če razpolagate z veliko količino podatkov.

Informacije, ki jih policija pridobi z operativni delom na posameznem primeru, postanejo kaj kmalu nepregledne zaradi velike količine ter njihove ne-strukturirane oblike. Ideja uporabe »text maninga« z uporabo »avtomatiziranega strukturiranja dokumentov« je zelo zanimiva, saj bi se s tem enormno zmanjšalo delo in stroški v kriminalistični analitiki. Vendar pri delovanju teh vidim oviro predvsem v jezikovnem področju.

## 6.2. Pravna podlaga zbiranja, upravljanja, varstva in dostopa do tajnih podatkov

### 6.2.1. Zbiranje in upravljanje podatkov policije

Pooblastilo za zbiranje podatkov in informacij o kaznivem dejanju in storilcu najdemo v Zakonu o kazenskem postopku, ki namesto pojmov »podatki in informacije« uporablja izraz »obvestila«, in sicer v naslednji povezavi: »Če so podani razlogi za sum, da je bilo storjeno kaznivo dejanje, za katero se storilec preganja po uradni dolžnosti, mora policija ukreniti potrebno, da se izsledi storilec kaznivega dejanja, ..., in da se zberejo vsa *obvestila*, ki bi utegnila biti koristna za uspešno izvedbo kazenskega postopka.« (prvi odstavek 148. člena ZKP). Zakon torej neposredno ne govori o analiziranju podatkov, vendar je jasno, da je analiza vsebovana že v namenu zbiranja podatkov, ta pa se izraža v uspešni izvedbi kazenskega postopka. Zbiranje podatkov samo po sebi torej ne more biti cilj (Maver, 1997, str. 105).

V Zakonu o policiji je določeno, da policisti zbirajo osebne in druge podatke neposredno od osebe, na katero se ti podatki nanašajo in od drugih, ki o tem kaj vedo, ali iz že obstoječih zbirk podatkov. Pri zbiranju osebnih in drugih podatkov od drugih oseb ali iz obstoječih zbirk, policisti niso dolžni o tem obvestiti osebe, na katero se podatki nanašajo, če bi to onemogočilo ali otežilo izvršitev določene naloge. Za policijo, ki preiskuje kaznivo dejanje je slednje določilo izrednega pomena, kajti v fazi ko storilec še ni znan ali ko obstaja le sum storitve kaznivega dejanja, je potrebno večji del podatkov zbrati od drugih oseb (od oškodovancev, prič, večjega kroga potencialno sumljivih oseb, itd.) (Maver, 1997, str. 104).

Kadar policija z zbiranjem obvestil ugotovi, da za določeno osebo obstajajo razlogi za sum, da je storila ali sodelovala pri storitvi kaznivega dejanja ji mora policija, preden začne od nje zbirati obvestila, povedati, katerega kaznivega dejanja je osumljena in kaj je podlaga za sum zoper njo, ter jo poučiti, da ni dolžna ničesar izjaviti in odgovarjati na vprašanja, da se ni dolžna izpovedati zoper sebe ali svoje bližnje ali priznati krivdo in, da ima pravico do zagovornika, ki si ga svobodno izbere in, ki je lahko navzoč pri njenem zaslišanju, ter, da se bo lahko vse, kar bo izpovedala, na sojenju uporabilo zoper njo (četrti odstavek 148. člena ZKP)

Policija po Zakonu o policiji upravlja z evidencami, ki jih zaradi opravljanja nalog zbirajo, obdelujejo, shranjujejo, posredujejo in uporabljajo policisti. Policija vodi in vzdržuje skupno 19 evidenc: (evidenco ovadenih oseb in kaznivih dejanj; evidenco kršiteljev in prekrškov; evidenco iskanih oseb; evidenco identifikacij; evidenco zaznav kaznivih dejanj; evidenco operativnih informacij; ...; in evidenco izdanih odredb za prepoved približevanja). Zakon o policiji taksativno našteva osebne podatke, ki jih je dovoljeno zbirati v vsaki evidenci, za vsako izmed njih pa našteva tudi dodatne podatke, včasih zelo podrobno (na primer vrsta, kraj, čas itd.), drugič spet bolj ohlapno (na primer motiv storitve, način storitve itd.). Zakon o policiji prav tako ureja tudi pravice posameznikov, na katere se nanašajo zbrani osebni podatki.

Podatki se v policijskih evidencah hranijo od enega do treh let, v večini evidenc pa do izbrisa obsodbe ali kazni, če te ni pa do zastaranja pregona. Po preteku rokov se podatki arhivirajo v zrcalno kopijo baze FIO (Fonetični indeks oseb), do katerih imajo dostop le ustrezno pooblaščen delavci. Podatki, za katere pa se ocenjuje, da po prepisu v arhiv niso pomembni za vsakdanje delo, pa se naknadno prepisujejo še na arhivski (off-line) medij (Lenarčič, 2002, str. 16).

Po Zakonu o policiji je policist dolžan varovati državno, uradno ali drugo tajnost, s katero se seznanil pri opravljanju nalog. Dolžnost varovanja tajnosti traja tudi po prenehanju delovnega razmerja.

Osnovna zaščita podatkov se izvaja preko registracije prijave, kjer ima vsak uporabnik svoje uporabniško ime in geslo, v okviru katerega mu skrbnik zaščite določa pooblastila. V večini primerov delimo zaščito podatkov na funkcije: vpogleda, spreminjanja podatkov, brisanja podatkov, dostopa do arhivskih podatkov, posebne funkcije verifikacij in dostopa do statističnih podatkov. Poleg tega pa se vsi dostopi do podatkov zapisujejo v poseben dnevnik dela. Rekonstrukcija pristopov je možna za 10 let nazaj, iz dnevnikov pa je možno razbrati kdo je, kdaj in kje je gledal ali spremenil kateri podatek (Lenarčič, 2002, str. 16). Za zbiranje, obdelovanje, shranjevanje, posredovanje in uporabo podatkov policijskih evidenc se uporabljajo določbe zakona o varstvu osebnih podatkov.

### **6.2.2. Varstvo osebnih podatkov**

Varstvo osebnih podatkov je pravica, ki je zagotovljena z Ustavo RS. Ustava določa, da pravica varstva osebnih podatkov vsebuje izrecno prepoved uporabe osebnih podatkov v nasprotju z namenom njihovega zbiranja in vsakomur daje pravico seznaniti se z zbranimi osebnimi podatki, ki se nanašajo nanj ter pravico do sodnega varstva ob njihovi zlorabi. Cilj varstva osebnih podatkov je varstvo posameznika, na katerega se ti podatki nanašajo in njegove informacijske zasebnosti.

Z veljavnim Zakonom o varstvu osebnih podatkov (v nadaljevanju besedila: ZVOP-1), se določajo pravice, obveznosti, načela in ukrepi, s katerimi se preprečujejo neustavni, nezakoniti in neupravičeni posegi v zasebnost in dostojanstvo posameznika oziroma posameznice pri obdelavi osebnih podatkov.

Osebni podatek je v ZVOP-1 opredeljen kot katerikoli podatek, ki se nanaša na posameznika, ne glede na obliko, v kateri je izražen. Medtem, ko je bil osebni podatek po starem zakonu (ZVOP) definiran kot podatek, ki kaže na lastnosti, stanja ali razmerja posameznika, ne glede na obliko, v kateri je izražen. Ob teh definicijah osebnih podatkov lahko opazimo, da je bila do sedaj veljavna definicija zelo široka in da je bilo včasih zelo težko določiti, kaj je in kaj ni osebni podatek. Novi zakon pa je pri opredelitvi osebnega podatka še bolj ohlapen, saj je sedaj osebni podatek katerikoli podatek, ki se nanaša na posameznika (Žurej, 2004, str. 30).

V ZVOP-1 je določeno, da se osebni podatki lahko obdelujejo le, če obdelavo osebnih podatkov in osebne podatke, ki se obdelujejo, določa zakon ali če je za obdelavo določenih osebnih podatkov podana osebna privolitve posameznika. Ob tem mora biti namen obdelave osebnih podatkov določen v zakonu, v primeru obdelave na podlagi osebne privolitve posameznika pa mora biti posameznik predhodno pisno ali na drug ustrezen način seznanjen z namenom obdelave osebnih podatkov.

Pri pridobivanju osebnih podatkov iz zbirk osebnih podatkov s področja zdravstva, policije, obveščevalno-varnostne dejavnosti države, obrambe države, sodstva in državnega tožilstva ter kazenske evidence in prekrškovnih evidenc ni dovoljena uporaba istega povezovalnega znaka (ko je na primer: EMŠO ali davčna številka) na način, da bi se za pridobitev osebnega podatka uporabil samo ta znak.

Možnost uporabe istega povezovalnega znaka namreč omogoča pridobitev osebnih podatkov iz več zbirk, ki skupaj dajejo »sliko posameznika« in pomenijo veliko nevarnost za zlorabe pri uporabi tako pridobljenih podatkov. Zato je v nekaterih državah uporaba takšnih povezovalnih identifikacijskih znakov prepovedana. ZVOP-1 sicer prepoveduje uporabo istega povezovalnega znaka na način, da bi se za pridobitev osebnih podatkov uporabil samo ta znak, vendar pa izjemoma omogoča njegovo uporabo, če je to edini podatek v konkretni zadevi, ki lahko omogoči, da se odkrije ali preganja kaznivo dejanje, zavaruje življenje ali telo posameznika ali zagotovi izvajanje nalog obveščevalnih in varnostnih organov (Rovšek, 2005, str. 76).

Nedopustno je torej povezovanje policijskih evidenc preko istega povezovalnega znaka z evidencami, ki so določene z Zakonom o zbirkah podatkov s področja zdravstvenega varstva (v nadaljevanju besedila: ZZPPZ). V ZZPPZ je na področju uporabe prepovedanih drog, določeno vodenje Evidence obravnave uživalcev drog, v okviru katere se zbirajo pomembni epidemiološki podatki o obravnavanih uporabnikih prepovedanih drog. Upravitelj evidence je Inštitut za varovanje zdravja RS, ki jo vodi v namene spremljanja pojava uživanja drog in obravnave uživalcev drog, njihovih značilnosti in ovrednotenje izvedenih ukrepov, načrtovanje zdravstvene in socialne mreže; razvoja preventivne in zdravstveno-vzgojne strategije; in načrtovanje bodočih usmeritev.

Zloraba osebnih podatkov je v 154. členu Kazenskega zakonika RS opredeljena tudi kot kaznivo dejanje, ki se preganja po uradni dolžnosti. Določeno je, da se z denarno kaznijo ali zaporom do enega leta kaznuje, kdor v nasprotju z zakonom uporabi osebne podatke, ki se smejo voditi samo na podlagi zakona ali na podlagi osebne privolitve posameznika, na katerega se podatki nanašajo ali kdor vdre v računalniško vodeno zbirko podatkov z namenom, da bi sebi ali komu drugemu pridobil kakšen osebni podatek. Če dejanje stori uradna oseba z zlorabo uradnega položaja ali uradnih pravic, se kaznuje z zaporom do dveh let.

### **6.2.3. Dostop do tajnih podatkov in njihovo varovanje**

Zakon o tajnih podatkih, določa skupne osnove enotnega sistema določanja, varovanja in dostopa do tajnih podatkov z delovnega področja državnih organov Republike Slovenije, ki se nanašajo na javno varnost, obrambo, zunanje zadeve ali obveščevalno in varnostno dejavnost države, ter prenehanja veljavnosti takšnih podatkov. Po njem se morajo ravnati državni organi, organi lokalnih skupnosti, nosilci javnih pooblastil ter drugi organi, ki pri izvajanju zakonsko določenih nalog pridobijo ali razpolagajo s podatki iz prej navedenih področij.

Za tajnega se lahko določi podatek, ki je tako pomemben, da bi z njegovim razkritjem nepoklicani osebi nastale, ali bi očitno lahko nastale, škodljive posledice za varnost države ali za njene politične ali gospodarske koristi. Podatek, ki mu je bila tajnost določena zato, da bi se na ta način prikrilo storjeno kaznivo dejanje, prekoračitev ali zloraba pooblastil, ali prikrilo kakšno drugo nezakonito dejanje ali ravnanje, ni taje. Pooblaščen oseba mora ob določitvi stopnje tajnosti podatka, oceniti možne škodljive posledice, če bi bil podatek odkrit nepoklicani osebi.

Tajni podatki imajo glede na možne škodljive posledice, eno od naslednjih stopenj tajnosti: »Strogo tajno« - njihovo razkritje bi ogrozilo vitalne interese RS ali bi nepopravljivo škodovalo; »Tajno« - njihovo razkritje bi lahko hudo škodovalo varnosti ali interesom RS; »Zaupno« - njihovo razkritje bi lahko škodovalo varnosti ali interesom RS; »Interno« - njihovo razkritje bi lahko škodovalo delovanju ali izvajanju nalog organa.

Pravico dostopa do tajnih podatkov imajo samo tiste osebe, ki imajo dovoljenje za dostop do tajnih podatkov in osebe, ki se morajo s temi podatki seznaniti zaradi opravljanja funkcije ali delovnih nalog (potreba po vedenju). Pristojni organ izda dovoljenje, če se z varnostnim preverjanjem ne ugotovi varnostnega zadržka. Varnostno se lahko preveri osebo, ki je s tem pisno soglašala in izpolni vprašalnik za varnostno preverjanje. V okviru varnostnega preverjanja pristojni organ varnostno oceni navedbe v izpolnjenem vprašalniku ter v zvezi s temi navedbami od osebe, na katero se ti podatki nanašajo, od drugih organov, organizacij in oseb ali zbirk podatkov, ki se vodijo na podlagi zakona zbira osebne in druge podatke, pomembne za ugotavljanje varnostnih zadržkov.

Osebe, ki opravljajo funkcijo ali delajo v organih, so dolžne varovati tajne podatke, ne glede na to, kako so zanje izvedele. Oseba, ki se je v okviru svojega dela seznanila s tajnimi podatki, teh ne sme uporabljati za druge namene kot za izvajanje delovnih nalog in funkcij. V vsakem organu se mora v skladu z Zakonom o tajnih podatkih in predpisih, vzpostaviti sistem postopkov in ukrepov varovanja tajnih podatkov, ki ustreza določeni stopnji tajnosti in onemogoča njihovo razkritje nepoklicanim osebam.

## 7. SKLEP

Z postavitvijo testnega sistema OLAP na področju prepovedanih drog smo v Generalni policijski upravi, Upravi kriminalistične policije, Sektorju za računalniško kriminaliteto in kriminalistično analitiko, Oddelku za kriminalistično analitiko, spoznali funkcionalnost in praktičnost uporabe sistema OLAP, v namene: okrepiteve dejavnosti policije usmerjene proti organiziranemu kriminalu in izboljšanju obveščevalnih in analitičnih del pri odkrivanju in preiskovanju kaznivih dejanj povezanih s prepovedanimi drogami.

Čeprav je zastavljen cilj zadeval celotni oddelek kriminalistične analitike, so vse opisane faze izgradnje testnega sistema OLAP na področju prepovedanih drog, temeljile na samostojnem delu avtorja specialističnega dela.

Priprave na vpeljavo testnega sistema OLAP na področju prepovedanih drog sem pričel z analizo zahtev in definiranjem problematike. Na podlagi tako ugotovljenih potreb uporabnikov je bila sprejeta odločitev o izbranih podatkih, osnovni enoti, dimenzijah in merah večdimenzionalnega podatkovnega modela. Model sem postopno dopolnjeval glede na dodatne potrebe in posebnosti, ki so se pojavljale pri modeliranju. Pri reševanju problemov in posebnosti sem si pomagal s strokovnimi napotki, ki jih lahko zasledimo v strokovni literaturi tujih in domačih avtorjev s področja večdimenzionalnega modeliranja podatkov. Vendar pa moram ob tem omeniti, da na takšen način ne dobimo tudi nujno najboljše rešitve problema, zato sem v pripravo večdimenzionalnega podatkovnega modela prav tako vključil lastno znanje, pridobljeno tekom podiplomskega študija in znanje pridobljeno, na podlagi praktičnih izkušenj pri opravljanju svojega delovnega poklica kriminalističnega analitika.

Pripravljen večdimenzionalni podatkovni model je bil osnova za izgradnjo OLAP kock. Za potrebe analiziranja in testiranje dostopa do podatkov sem pripravil več OLAP kock. Za potrebe analize podatkov iz sistema OLAP sem uporabil OLAP odjemalec Microsoft Excel. Ker pa vsi uporabniki, niso imeli povezave do OLAP strežnika, sem pripravil lokalno kocko in interaktivne spletne strani, katere sem objavil na skupnem strežniku. Uporabnikom je bil tako omogočen dostop do podatkov preko spletnega brskalnika Microsoft Internet Explorer. Določene funkcionalnosti, ki jih sicer omogoča neposreden dostop do OLAP strežnika, smo s tem izgubili, vendar je bil po drugi strani tak spletni pregled za marsikoga povsem zadovoljiv.

Ugotovljene možnosti in priložnosti uporabe sistema OLAP na področju prepovedanih drog sem predstavil vodstvenemu kadru kriminalistične policije, strokovnim sodelavcem v kriminalistični policiji, kakor tudi zaposlenim v službi za informatiko. Končna ocena, ki smo jo postavili glede na rezultat testiranja sistema OLAP je, da bi vpeljavo trajnega sistema OLAP, lahko vplivala na odločitve vodilnih delavcev policije in na oblikovanje dolgoročne politike kriminalistične policije na področju prepovedanih drog, saj bi omogočila: usklajeno organiziranje aktivnosti policije, poznavanje storilcev, njihovih motivov in načinov dela, dinamike delovanja trga s prepovedanimi drogami, nezakonitih poteh, obsega organizirane kriminalitete, kakor tudi trende, ki prevladujejo na njem.

Z vzpostavitvijo sistema OLAP bi uporabniki pridobili hiter in zanesljiv vpogled na podatke, ki zadevajo: zaznavo kaznivih dejanj in prekrškov povezanih s prepovedanimi drogami; preteklo obravnavo oseb, ki so bile obravnavane zaradi prepovedanih drog; in zasege prepovedanih drog. Uporabniki več ne bi bili obremenjeni s številnimi pred-pripravljenimi standardnimi statističnimi pregledi in hkrati tudi omejeni s pred-definiranimi polji, katerih postavitev večinoma ustreza pripravi podatkov za standardna letna in polletna statistična poročila. Funkcionalnost sistema bi se odražala v večdimenzionalnem pogledu na podatke, enostavnem raziskovanju in analiziranju velikih količin podatkov, njihovih povezav in omogočanju zahtevnega izračunavanja ter inteligentne obravnave časovnega vidika.

Pri policijskem delovanju v boju zoper organizirani kriminal je osrednjega pomena obveščenost o dejavnostih kriminalnih skupin, kar pomeni konstantno zbiranje in vrednotenje podatkov in informacij. Predvsem pa je pomembno vedenje, kaj se dogaja in kaj se pripravlja, saj tako pridobljene informacije opredeljujejo vse druge aktivnosti policije. Vendar moramo ob tem poudariti, da je osrednji in bistveni del obveščevalnega procesa analiza obveščevalnih podatkov, saj se brez te obveščevalno delo lahko kar hitro sprevrže v golo zbiranje informacij. Zato je osrednji del testiranja predstavljala analiza podatkov iz sistema OLAP, za potrebe strateške analize »Problematika prekomerne uporabe prepovedanih drog v RS«.

Ker so podatki, ki so predstavljeni v večdimenzionalnem podatkovnem modelu običajno bolj intuitivni in preprosti za uporabo tudi tistim uporabnikom, ki niso računalniški strokovnjaki, je bil naš namen približati podatkovni model uporabnikovemu načinu razmišljanja in jim omogočiti hitro in učinkovito delo s podatki. Vendar ob tem moramo opomniti, da so uporabniške zahteve mnogo bolj zapletene od enostavnih šolskih primerov, zato je bilo potrebno večdimenzionalni podatkovni model prilagajati številnim posebnostim in problemom. Ob reševanju posameznih problemov smo želeli, da so rešitve čim bolj preproste in razumljive za uporabnika, kajti le on je ta, ki določa uporabnost posameznega sistema.

Ob tem moramo poudariti, da je bila analiza podatkov iz sistema OLAP, le del strateške analize, kajti pri izdelavi strateške analize, je bilo med viri podatkov veliko t.i. »mehkih« podatkov, katerih problem se kaže pri njihovem ocenjevanju in vrednotenju. Pri izdelavi strateške analize je bil zato uporabljen širok spekter analitskih tehnik, ki v največji meri omogočajo realen, oziroma nepristranski prikaz problematike prepovedanih drog v slovenskem prostoru.

S tem je bila ustvarjena analitična osnova za predvidevanje bodočih stanj in dogodkov na tem področju, kajti če poznamo vzrok, lahko predvidimo posledice, z njimi pa tudi metode za njihovo učinkovito reševanje. Rezultati strateške analize bodo podlaga za sprejemanje ukrepov in izdelavo strategij v boju proti prepovedanim drogam, reševanju posledic uporabe in zlorabe prepovedanih drog ter izdelavo pristopov in programov za preprečevanje oziroma zmanjševanje posledic njihove uporabe.



Velika večina uporabnikov je bila pozitivno presenečena nad uporabo sistema OLAP, saj so imeli namesto papirnih poročil, na voljo orodje, s katerimi lahko raziskujejo podatke, tako na globalnem nivoju, kakor jim je tudi omogočeno drobljenje v podrobnosti teh. Vendar moramo pri tem omeniti, da je bilo testiranje uporabe OLAP odjemalca (Microsoft Excel in Office Web Components), omejeno samo s strani uporabnikov, ki poznajo delo z vrtilnimi tabelami. Zato je za uporabnike, ki dnevno sodelujejo v procesu odločanja, potrebno izbrati specializirani OLAP odjemalec, ki bo omogočal relativno enostavne uporabniške vmesnike in prav tako omogočal zahtevnim uporabnikom veliko svobode in zmogljivost za podrobnejše in raznolike analize podatkov. Uporaba specializiranega odjemalca je prav tako potrebna zaradi zagotovitve varnosti, ki se odraža pri dostopu do podatkov in zagotovitvi sledljivosti teh dostopov. Problemi se tako kažejo v primerih, ko želimo uporabiti enotno OLAP kocko, do katere dostopajo vsi uporabniki, vendar ob tem želimo s strani posameznega uporabnika, omejevati njihove vpogled v podrobnost podatkov, glede na dodeljen nivo dostopa do podatkov.

Pri vzpostavitvi testnega sistema OLAP smo želeli, da bi bili pridobljeni rezultati skladni s podatki, ki so prikazani v uradnih letnih in polletnih statističnih poročilih policije. Zaradi postavljene zahteve po primerljivosti podatkov, nismo spreminjali sistemov merjenja (kot so: število obravnavanih oseb, število zadev, količine zasega), temveč smo le opozorili na določena odprta vprašanja, ki zadevajo uporabo navedenih mer in zanesljivosti posameznih podatkov. Problemi zanesljivosti podatkov izvirajo predvsem iz postopkov in načinov evidentiranja teh. V informacijskem sistemu policije je namreč omogočeno veliko število vnosnih mest, kar po svojih prednostih prinaša tudi nevarnosti pri vnos napačnih vrednosti podatkov. Večina evidentiranih podatkov prav tako temelji na izjavi obravnavane osebe oz. osumljenca, zato je vprašljiva tudi verodostojnost zbranih podatkov (napr.: *Pri obisku zdravnika boste na vprašanja, ki vam jih ta zastavlja odgovarjali resnično, saj ste izkazali svoj interes glede pomoči. Pričakovati zanesljive odgovore s strani osumljenca, kateremu grozi odobritev pripora, je zato vprašljivo.*)

Pri analiziranju podatkov se je prav tako izrazila potreba po pridobitvi dodatnih podatkov. Na področju osebnih podatkov, bi tako lahko uporabili podatke, ki so evidentirani v ostalih policijskih evidencah in kažejo na: osebni opis, družinske in premoženjske razmere, šolsko izobrazbo, poklic in zaposlitev osumljencev.

Rešitev je tako samo v postavitvi primerne strukture podatkovnega skladišča, ki bo s svojimi strukturami in procesi, omogočal celovit pogled na podatke organizacije. Zato si je policija v okviru svojih delovnih nalog zastavila kot cilj, izgraditev podatkovnega skladišča policije. V kolikor pa govorimo o možnostih graditve integriranega podatkovnega skladišča, s povezovanjem podatkovnih zbirk s strani državnih, vladnih in ne-vladnih ustanov, ki imajo skupen cilj zmanjševanje uporabe prepovedanih drog in posledice prekomerne uporabe le-teh, pa se že zatekamo k teorijam »Velikega brata«, zato je povezovanje teh zakonsko prepovedano.

## 8. POJASNILA KRATIC IN SLOVAR IZRAZOV

**Aggregated facts as attributes** - Združena dejstva kot atributi. Priprava združenih dejstev kot atributov se nanaša na pripravo deskriptivnih vrednosti iz posameznih uporabljenih mer oz. dejstev.

**Bottom-up** – »od-spodaj-navzgor«. Pojem večinoma uporabljamo pri postopni izgradnji posameznih neodvisnih področnih skladišč, ki jih na koncu povežemo v logično podatkovno skladišče celotne organizacije.

**Common business model** – Skupen poslovni model.

**CR** - CR je okrajšava za številko iz Centralnega registra in predstavlja enoznačni ključ za osebe. Ključ je tvorjen umetno in iz njega ne moremo sklepati na identiteto osebe.

**Crime analysis** – Kriminalistična analitika. Kriminalistična analitika je širok pojem, ki je bil izbran za opis analitičnega dela znotraj kriminalistične policije. Obstaja zelo veliko definicij kriminalistične analitike. Interpolova definicija pa govori o kriminalistični analitiki kot o identifikaciji in omogočanju vpogleda v odnose med podatki, o kriminaliteti in drugimi potencialno relevantnimi podatki z vidika policijske in sodne prakse.

**Cross System Product (CSP)** – programsko okolje.

**Cube** - Kocka (OLAP kocka). Kadar je predstavitev večdimenzionalnega podatkovnega modela zasnovana na večdimenzionalnih bazah podatkov ali OLAP tehnologiji govorimo, o t.i. OLAP kockah.

**Data driven** - podatkovno voden. Izraz se nanaša na proces izgradnje podatkovnega skladišča, kjer izhajamo iz razpoložljivih podatkov.

**Data Marts** - Podatkovne tržnice oz. področna podatkovna skladišča. Področnega podatkovnega skladišča predstavljajo zbirko področnih objektov za podporo odločanju, ki temeljijo na potrebah danega poslovnega področja v podjetju.

**Data mining** - Podatkovno rudarjenje. Podatkovno rudarjenje lahko opredelimo kot avtomatizirano iskanje veljavnih, prej nepoznanih vzorcev in povezav v velikih količinah podatkov, na podlagi katerih lahko ukrepamo, z namenom boljših odločitev.

**Data warehouse (DW)** - Podatkovno skladišče. Podatkovno skladišče je osrednja, od operativnih zbirk ločena podatkovna zbirka podatkov, ki je usmerjena na ključne poslovne vire (kot so trgi, kupce, izdelki...itd.) in namenjena shranjevanju in dostopu do podatkov za potrebe odločanja in vodenja.

**DB2** – relacijska baza podjetja IBM

**Decision support information systems (DIS)** - Sistemi za podporo odločanju (SPO). Sisteme za podporo odločanju lahko opredelimo kot množico strojne opreme, programske opreme in ljudi, ki pomagajo v procesu odločanja.

**Degenerate dimensions (DD)** – Brez atributna dimenzija. Ob postavljanju dimenzijskih tabel včasih naletimo na dimenzije, ki poleg tujega ključa v tabeli dejstev nimajo dodatnih atributov in dejansko sploh ne potrebujejo svoje dimenzijske tabele. Tovrstno dimenzijo imenujemo brez atributna dimenzija.

**Dimensions** - Dimenzija. Del zgradbe OLAP kock ali zvezdne sheme v kateri so podatki organizirani po nivojih (npr. v dimenziji naslov bi uporabili nivoje: država/mesto/ulica). Dimenzije predstavljajo kategorije po katerih analiziramo podatke.

**Entity Relationship diagram (E-R Diagram)** – Entitetno relacijski diagram

**Executive information systems** - Vodstveni informacijski sistemi. Poenostavljeno bi lahko rekli, da so vodstveni informacijski sistemi v primerjavi s sistemi za podporo odločanju, običajno preprostejši in se uporabljajo za bolj vsakodnevne taktične odločitve.

**Facts** – Dejstva ali mere. Mere predstavljajo vrednost ali velikost podatka, ki ga analiziramo. Ker uporabniki pregledujejo mere glede na poljubne dimenzije na poljubnih ravneh hierarhije, je zaželeno, da so te vrednosti numerične in aditivne.

**Fonetični indeks oseb (FIO)**. FIO je operativni informacijski sistem policije, ki podpira operativno delo policije, kar pomeni, da se vse osebe, ki jih je obravnavala policija evidentirajo v enotni bazi. Hkrati pa je postal tudi vir podatkov za med-organizacijsko povezovanje policije znotraj državne uprave in zunaj nje.

**Grain** - Osnovna enota.

**Help table** - Pomožna tabela. Pri izbiri dimenzij velja pravilo, da naj bi dimenzije določili tako, da je za opis zapisa v tabeli dejstev dovolj ena vrednost dimenzije. Zgodi pa se nam lahko, da to ni mogoče, in da ne gre drugače kot, da dimenzija vsebuje večvrednostne podatke, da bi se izognili povezavi mnogo proti mnogo, lahko uporabimo t.i. pomožno tabelo.

**Hybrid (HOLAP)**. Je način shranjevanja OLAP podatkov, kjer shranjujejo vhodni podatki v relacijski bazi, agregacije pa v večdimenzionalni strukturi.

**Information staging areas** – Področje priprave informacij.

**Kazenski zakonik (KZ)**

**Kazniva dejanja (KD)**

**Middle tier** – Srednje nivojski strežnik.

**Mini dimensions** – Mini dimenzije. Tovrstno dimenzijo pripravimo tako, da pogosto spreminjajoče podatke iz dimenzije, izločimo in jih skupno dodamo v novo nastalo dimenzijo. Občasne spremembe podatkov tako ne vplivajo več neposredno na osnovno dimenzijo.

**Multidimensional Expressions (MDX)** - Večdimenzijske izjave. Sintaksa, ki se uporablja za poizvedbo po večdimenzionalnih bazah podatkov (OLAP kockah). Prav tako pa se uporablja pri poizvedbah z uporabo vrtilnih tabel z OLE DB.

**Multidimensional (MOLAP)** Je način shranjevanja OLAP podatkov, kjer shranjujejo vhodni podatki in agregacije v večdimenzionalni strukturi (OLAP kockah).

**Offline cube file** - Kockovna datoteka brez povezave ali lokalna kocka je datoteka, ki je namenjena za shranjevanje izvornih podatkov iz OLAP strežnika.

**Online Analytical Processing (OLAP)** – Sistemi za sprotno analitično obdelavo podatkov. Tehnologija zbirk podatkov, optimizirana za poizvedovanje in poročanje. Sistemi za sprotno obdelavo podatkov omogočajo analitikom in ostalim uporabnikom hiter in zanesljiv vpogled na podatke z različnih zornih kotov.

**Online Transaction Processing (OLTP)** - Sistem za sprotno obdelavo transakcij. Sistem procesiranja podatkov, ki je namenjen zajemanju in shranjevanju vseh poslovnih transakcij neke organizacije v trenutku, ko so te nastale. Značilnost transakcijskih sistemov je operativna baza

podatkov, kjer so podatki shranjeni brez nepotrebnega podvajanja na način, ki omogoča njihovo uporabo številnim uporabnikom z različnimi zahtevami.

**Pivot table** – Vrtilne tabele. Vrtilna tabela. Interaktivno, navzkrižno poročilo v programu Microsoft Excel.

**Private dimensions** - Privatne dimenzije. V sklopu OLAP orodja Analysis Services, lahko pripravimo privatne dimenzije, ki se lahko uporabljajo samo v okviru posamezne OLAP kocke.

**Relational OLAP (ROLAP)**. Je način shranjevanja OLAP podatkov, kjer shranjujejo vhodni podatki in agregacije v relacijski bazi.

**Real time cube**. Kocka v realnem času. Kocka v realnem času omogoča, da ko se podatek zapiše v podatkovno skladišče, je ta že predstavljen v OLAP odjemalcu

**Schengenski informacijski sistem (SIS)**.

**Shared Dimensions** - Skupne dimenzije. V sklopu OLAP orodja Analysis Services lahko pripravimo skupne dimenzije, ki jih lahko uporabljajo v okviru določene zbirke večdimenzionalnega podatkovnega modela.

**Slowly Changing Dimension (SCD)** - Počasi spreminjajoče se dimenzije. Dimenzijam, katerih vrednosti se s časom spreminjajo opredeljujemo kot t.i. počasi spreminjajoče se dimenzije.

**Snow-flake shema** - Snežinkasta shema. Večina pomislekov v zvezi z zvezdno shemo se nanaša na porabo prostora, ki ga zavzemajo ponavljajoči se tekstualni podatki. Zato nekateri uporabljajo snežinkasto shemo za katero je značilno, da so posamezne dimenzijske tabele normalizirane.

**Spider-web environment** - Okolja pajkove mreže. Prepletenost podatkovnih virov.

**Star shema** - Zvezdna shema. V kolikor je predstavitev večdimenzionalnega podatkovnega modela zasnovana na relacijskih bazah govorimo o t.i. zvezdni shemi, saj imajo večdimenzionalni podatkovni modeli obliko zvezde, pri čemer je v sredini tabela mer ali dejstev, obkrožajo pa jo vsaj ena oziroma običajno več dimenzij.

**Structured Query Language (SQL)** – Nepostopkovni strukturirani poizvedovalni jezik za kreiranje poizvedb iz relacijskih zbirk podatkov.

**Zakon o kazenskem postopku (ZKP)**

**Zakon o policiji (Zpol)**

**Zakon o tajnih podatkih (ZTP)**

**Zakon o varstvu osebnih podatkov (ZVOP-1)**

**Zakon o zbirkah podatkov s področja zdravstvenega varstva (ZZPPZ)**

**Zakonu o kazenskem postopku (ZKP)**

## 9. LITERATURA

1. Aličehajič Sabrina, Vidmar Zoran: Orodja OLAP. Arhiv referatov na posvetovanju dnevi slovenske informatike 2001. Slovensko društvo informatika. 2001, [URL: <http://www.drustvo-informatika.si/dogodki/arhiv/dsi2001/>], 28.12.2004.
2. Babič Matjaž: Analize in poročila OLAP kot del sistema za podporo odločanju. Diplomsko delo. Ljubljana : Ekonomska fakulteta, 2002. 45. str.
3. Berry Michael J. A., Linoff Gordon: Mastering Data Mining The Art and Science of Customer Relationship Management. New York : John Wiley & Sons, Inc., 2002. 494 str.
4. Berson Alex, Smith Stephen, Thearling Kurt. Building Data Mining Applications for CRM. United States of America : McGraw-Hill, 2000. 449 str.
5. Bidgoli Hossein: Modern Information Systems for Managers. Kalifornija : Academic press, 1997. 438 str.
6. Brejc Nataša: Raziskovalna delavnica: kriminološki problemi zlorabe mamil. Pravna praksa, Ljubljana, 1993, 286, str. 20.
7. Čibej Jože Andrej: Sistemi za podporo poslovnemu odločanju. Zapiski predavanj. Ljubljana : Ekonomska fakulteta, 1998. 141 str.
8. Damij Talib, Grad Janez, Jaklič Jurij: Izbrane teme iz informacijske tehnologije. Ljubljana : Ekonomka fakulteta, 1995. 316 str.
9. Dežman Zlatko: Preprečevanje razširjanja uporabe nedovoljenih drog. Program strokovnega srečanja »Z znanjem proti odvisnostim od nedovoljenih DROG«. 2001, [URL: [http://www.maribor.si/MOM\\_INT/AKTUALNO/DROGE.pdf/](http://www.maribor.si/MOM_INT/AKTUALNO/DROGE.pdf/)], 22.03.2003.
10. Dimovski Vlado: Temelji organiziranja in odločanja. Ljubljana : Ekonomska fakulteta, 2002. 336 str.
11. Ferle Maja: Večrazsežnostno modeliranje podatkov v praksi. Arhiv referatov na posvetovanju dnevi slovenske informatike 2001. Slovensko društvo informatika. 2001a, [URL: <http://www.drustvo-informatika.si/dogodki/arhiv/dsi2001/>], 28.12.2004.
12. Ferle Maja: Sprotna analiza podatkov za vsakogar. Monitor, Ljubljana, februar 2001, 2001b, str. 120 – 122.
13. Golob Izidor, Welzer Tatjana: Arhitekture podatkovnih skladišč. Arhiv referatov na posvetovanju dnevi slovenske informatike 2001. Slovensko društvo informatika. 2001, [URL: <http://www.drustvo-informatika.si/dogodki/arhiv/dsi2001/>], 28.12.2004.
14. Grad Janez, Jaklič Jurij: Baze podatkov. Ljubljana : Ekonomska fakulteta, 1996. 254 str.
15. Inmon, W. H.: Building the Data Warehouse, Second Edition. Canada : John Wiley & Sons, 1996. 393 str.
16. Inmon, W.H.: Choosing the Right Approach to Data Warehousing Big Bang vs. Iterative. Inmon Associates, Inc. 1999c, [URL:<http://www.inmoncif.com/library/articles/bigbang.asp>], 18. 2. 2005
17. Inmon, W.H.: Independent data marts vs dependent data marts. Inmon Associates, Inc. 1999a, [URL: <http://www.inmoncif.com/library/whiteprs/datamart.htm>], 18. 2. 2005
18. Inmon, W.H.: Why You Can't Turn A Data Mart Into A Data Warehouse. Inmon Associates, Inc. 1999b, [URL: <http://www.inmoncif.com/library/articles/artdmtdw.asp>], 18. 2. 2005

19. Jaklič Jurij: Upravljanje podatkov. Gradiva pri predmetu Baze podatkov in njihova uporaba. Ljubljana : Ekonomska fakulteta. 2004. [URL: [ftp://ftp.ef.uni-lj.si/\\_dokumenti/predmeti/bp\\_spec.ppt](ftp://ftp.ef.uni-lj.si/_dokumenti/predmeti/bp_spec.ppt) ], 15.12.2004.
20. Jaklič Tatjana: Podatkovna skladišča v finančnih institucijah. INFO SRC.SI, Ljubljana, 2001, 30, str. 3 - 5
21. Jug Marjan: Mit, stvarnost in perspektiva. Gospodarski vestnik, Ljubljana, 1999, 42, str. 33
22. Kimball Ralph, Ross Margy: The Data Warehouse Toolkit. The Complete Guide to Dimensional Modeling (Secund Edition). Canada : John Wilwy & Sons, 2002. 436 str.
23. Kimball Ralph: Managing Helper Tables. Intelligententerprise, 2001, [URL: [http://www.intelligententerprise.com/010810/412warehouse1\\_2.jhtml](http://www.intelligententerprise.com/010810/412warehouse1_2.jhtml)], 11. 1. 2005
24. Konič Blaž: Uporaba podatkovnega rudarjenja pri odkrivanju nezaželene elektronske pošte. Diplomsko delo. Ljubljana : Ekonomska fakulteta, 2003. 45 str.
25. Kričej, Dušan: Hudič se skriva v podatkih. Finance, Ljubljana, 1998, 42, str. 8
26. Lenarčič Inge: Fonetični indeks oseb (FIO) 1992-2002. Ljubljana : Ministrstvo za notranje zadeve, 2002. 59 str.
27. Maver Darko, Rupnik Andrej, Golja Janez. Kriminalistika. Ljubljana : Uradni list Republike Slovenije, 1997. 280 str.
28. Microsoft: Microsoft. Analysis Services Features. Microsoft Corporation, 2005, [URL: [http://msdn.microsoft.com/library/default.asp?url=/library/en-us/olapmdad/agintroducing\\_8pfb.asp](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/olapmdad/agintroducing_8pfb.asp)], 05.02.2005
29. Novak Jani: Odnos države do obrobni skupin prebivalstva. Diplomsko delo. Ljubljana : Ekonomska fakulteta, 2001. 42 str.
30. Paul Seth, Gautam Nitin, Balint Raymond: Preparing and Mining Data with Microsoft® SQL Server™ 2000 and Analysis Services. USA : Microsoft Corporation, 2002. 154 str.
31. Pendse Nigel, Creeth Richard: What is OLAP? The OLAP Report. Optima Publishing Ltd. 2004, [URL: <http://www.olapreport.com/fasmi.htm>], 12.02.2005
32. Povšič Sabrina: Obetavno brskanje za informacijami. Gospodarski vestnik, Ljubljana, 1998, 19, str. 16
33. Rajkovič Vladislav: Je podatkovno skladiščenje del splošne informacijske kulture. Kranj : Fakulteta za organizacijske vede, 2004 [URL: <http://www.fov.uni-mb.si/programiranje/uros/files/BP/DW98.pdf>], 28.12.2004
34. Rapaić Boštjančič Urška: Analitično upravljanje odnosov s strankami za ponudnike telekomunikacijskih storitev. Magistrsko delo. Ljubljana : Ekonomska fakulteta. 2004. 49 str.
35. Remškar Andrej: Analiza obiska v spletni trgovini s pomočjo podatkovnega skladišča. Diplomsko delo. Univerza v Ljubljani. Fakulteta za računalništvo in informatiko. 2002, 55. str
36. Roblek Peter: Analiza podatkov z OLAP tehnologijo. Raziskovalno delo podiplomskih študentov v Sloveniji. Ljubljana : Društvo mladih raziskovalcev Slovenije – združenje podiplomskih študentov, 2003, str. 322-331.
37. Rovšek Jernej: Zasebno in javno v medijih. Pravna ureditev in praksa v Sloveniji. Ljubljana : Mirovni inštitut, 2005. 101 str.

38. Shim J.P. et al: Past, present, and future of decision support technology. Decision support systems and Electronic commerce, North-Holland, 33, 2002, str. 111-126.
39. Šmid Tomaž: Uporaba računalniškega razpoznavanja govora in transformacije naravnega jezika v MDX sintakso pri kreiranju večdimenzijskih poizvedb iz velikih skladišč podatkov. Magistrsko delo. Ljubljana : Ekonomska fakulteta, 2002. 108 str.
40. Štemberger Indihar Mojca, Jurij Jaklič, Aleš Groznik, Andrej Kovačič: Se slovenski managerji zavedajo pomena kakovostnih informacij za poslovno odločanje?. Arhiv referatov na posvetovanju dnevi slovenske informatike 2001. Slovensko društvo informatika. 2001, [URL: <http://www.drustvo-informatika.si/dogodki/arhiv/dsi2001/>], 28.12.2004.
41. Vagaja Aleksandra: Pri gradnji podatkovnih skladišč ne zaostajamo. Finance, Ljubljana, 2003, 127, str. 21
42. Vagaja Aleksandra: Učinkovito podporo bo ponujalo odprto skladišče. Finance, Ljubljana, 2004, 129, str. 22
43. Vintar Mirko: Informatika. Ljubljana : Bons, 2003. 237 str.
44. Vrančič Martin. 1998. Problematika heroina s policijskega aspekta. Policija. 1998, [URL: <http://www.policija.si/si/uks/heroin.html>], 22.03.2003
45. Žorž Jaka: OLAP ni samo za direktorje. Finance, Ljubljana, 2001, 97, str. 18
46. Žurej Jurij: Ob novem zakonu o varstvu osebnih podatkov. Novinar, Ljubljana, 2004. [URL: <http://>], 05.02.2005, str. 30-33

## 10. VIRI

1. CVI: E-poslovanja v javni upravi RS za obdobje od leta 2001 do leta 2004 (september – 2004). Center vlade za informatiko, 2001
2. KZ: Kazenski zakonik. (Uradni list RS, št. 63/94)
3. RRC. Zadovoljen direktor ima zagotovo dobrega informatika. RRC Računalniške storitve, 2004, [URL: <http://www.rrc.si/novitrendi/> ], 28.12.2004
4. SRC. Kaj razlikuje uspešna podjetja od neuspešnih?. SRC, 2004, [URL: <http://www.src.si/ponudba/prospekti/podatsklad.pdf>], 28.12.2004.
5. Vlada RS: Nacionalni program na področju drog 2002 - 2008. Vlada R Slovenije, 2002, 61. str.
6. ZKP: Zakon o kazenskem postopku. (Uradni list RS, št. 63/94)
7. ZPol: Zakon o policiji. (Uradni list RS, št. 49/98)
8. ZTP: Zakon o tajnih podatkih. (Uradni list RS, št. 87/01)
9. ZVOP-1: Zakon o varstvu osebnih podatkov. (Uradni list RS, št. 86/04)
10. ZZPPZ: Zakon o zbirkah podatkov s področja zdravstvenega varstva. (Uradni list RS, št. 65/00)