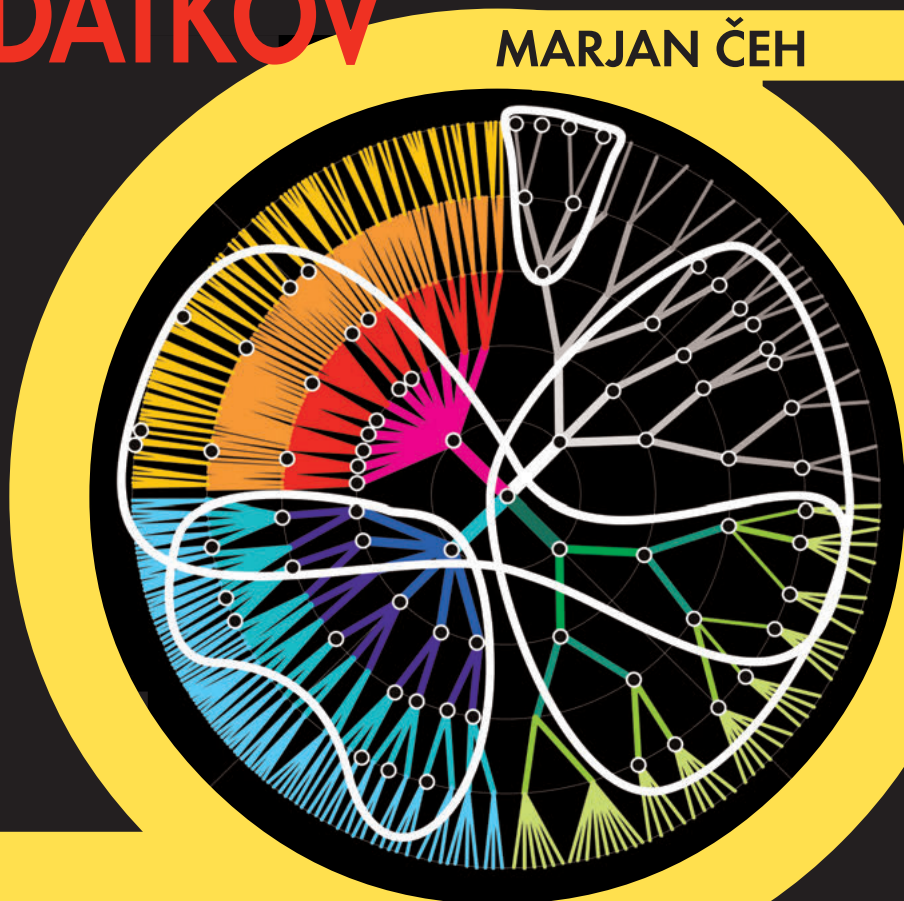


**SEMANTIČNA  
INTEGRACIJA  
ZBIRK  
PROSTORSKIH  
PODATKOV**

**MARJAN ČEH**





Marjan Čeh

•

SEMANTIČNA INTEGRACIJA  
ZBIRK PROSTORSKIH PODATKOV



ZALOŽBA  
Z R C

Marjan Čeh  
SEMANTIČNA INTEGRACIJA ZBIRK PROSTORSKIH PODATKOV

© 2003, Založba ZRC, ZRC SAZU

Recenzenta: Tomaž Podobnikar, Drago Perko  
Jezikovni pregled: Marjeta Humar  
Oblikovanje in likovno-grafična ureditev Milojka Žalik Huzjan

Izdajatelj: Prostorskoinformacijska enota ZRC SAZU Za izdajatelja: Ivan Šprajc

Založnik: Založba ZRC, ZRC SAZU  
Za založnika: Oto Luthar  
Glavni urednik: Vojislav Likar

Tisk: Littera picta d.o.o., Ljubljana

Digitalna različica (pdf) je pod pogoji licence CC BY-NC-ND 4.0 prosto dostopna:  
<https://doi.org/10.3986/9616500031>

CIP - Kataložni zapis o publikaciji  
Narodna in univerzitetna knjižnica, Ljubljana

659.2:681.3  
316.772.4:111

ČEH, Marjan  
Semantična integracija zbirk prostorskih podatkov / Marjan  
Čeh. - Ljubljana : Založba ZRC, ZRC SAZU, 2003

ISBN 961-6500-03-1  
126015744

# SEMANTIČNA INTEGRACIJA ZBIRK PROSTORSKIH PODATKOV

MARJAN ČEH

LJUBLJANA 2003



# VSEBINA

Predgovor .....	7
1 UVOD .....	9
1.1 Integracija informacij .....	9
1.2 Semantična medopravilnost .....	12
2 SEMANTIČNA INTEGRACIJA PODATKOV .....	15
2.1 Integracija shem zbirk podatkov .....	17
2.2 Semantična bogatitev shem .....	22
2.3 Integracija znanj v zbirkah prostorskih podatkov .....	32
3 KONCEPTUALNO MODELIRANJE .....	35
3.1 Klasifikacija in abstrakcija v konceptualnem modeliranju .....	36
3.2 Formalno modeliranje konceptov .....	39
3.3 Konceptualno modeliranje in informacijski sistemi .....	42
4 FILOZOFSKE IN SIMBOLNE TEORIJE PROSTORA .....	45
4.1 Kognitivno kartiranje prostora geografskih razsežnosti .....	48
4.2 Semiotika .....	51
4.3 Ontologija in informacijski sistem .....	59
4.4 Ontologija v geoinformatiki .....	63
5 ONTOLOGIJA PROSTORA GEOGRAFSKIH RAZSEŽNOSTI .....	67
5.1 Prostorski objekti geografskega obsega .....	67
5.2 Pravila za modeliranje ontologije prostora geografskih razsežnosti .....	69
5.3 Opredelitev ontologije prostora geografskih razsežnosti .....	70
6 OCENJEVANJE SEMANTIKE PROSTORSKIH PODATKOV .....	77
6.1 Analiza semantične ustreznosti OPGR za povezovanje .....	79
6.2 Semantična globina .....	82
6.3 Semantične razdalje .....	83
6.4 Semantična razpršenost .....	86
6.5 Analiza semantičnih razmerij aplikativnih konceptov v okviru semantičnega referenčnega omrežja .....	88

7	MEDOPRAVILNA INTEGRACIJA ZBIRK PROSTORSKIH PODATKOV V PORAZDELJENIH INFORMACIJSKIH SISTEMIH.....	91
7.1	Porazdeljeni informacijski sistemi in semantična medopravilnost .....	92
7.2	Standardi za povezovanje zbirk prostorskih podatkov .....	96
7.3	Deduktivni objektno usmerjen pristop za medopravilno upravljanje zbirk prostorskih podatkov .....	100
7.4	Strategija razvoja medopravilnih državnih zbirk prostorskih podatkov .....	102
7.5	Opredelitev metod medopravilnega povezovanja zbirk podatkov in GIS .....	104
8	ZAKLJUČEK .....	109
	Priloge.....	115
	Terminologija .....	119
	Besednjak .....	119
	Seznam uporabljenih kratic .....	135
	Seznam uporabljenih oznak .....	135
	Stvarno kazalo .....	137
	Kazalo slik .....	142
	Kazalo preglednic .....	143
	Literatura .....	144



# Predgovor

Vsakodnevno se hote ali nehote srečujemo z najrazličnejšimi podatki. Pri tem uporabljamo sposobnosti komuniciranja ali pa si pomagamo z najrazličnejšimi knjigami, učbeniki, enciklopedijami, pravopisi, imeniki, slovarji ter drugimi »zbirkami podatkov«. Vse večjo korist nam prinaša medmrežje, na katerem lahko z iskalniki najdemo informacije o (skorajda) vsem, kar nas zanima. Za iskani ali obravnavani izraz lahko dobimo več pomenov ali pa se istemu pomenu pripiše več izrazov. Omenjena zadrega nam v lingvističnem smislu lahko popestri vsakdanje življenje, če pa se takšni nepredvidljivi rezultati iskanja pojavljajo na primer med iskanjem informacij v prostoru po podatkovnih shemah v medmrežju, pa nam ga lahko občutno otežijo. Za primer vzemimo iskanje besede »baza«. V računalništvu si pod tem izrazom največkrat predstavljamo, da gre za »bazo podatkov«. Vendar lahko ta izraz v drugačnih razmerah (na pomenski oz. semantični ravni) razumemo tudi kot »osnovo« ali »podlago«, lahko gre za »vojaško bazo« ali pa morda za kemično »bazično snov«. Če se vrnemo k računalništvu, lahko izraz »baza podatkov« brez spremembe pomena zamenjamo z »zbirka podatkov«.

Pri semantičnem prevajanju in opredelitvi nekega predmeta, pojma ali aktivnosti si lahko pomagamo z ontologijo kot filozofsko disciplino, ki obravnava vzroke in najsplošnejše lastnosti stvarnosti. Ontologija omogoča semantično opredelitev zbirk prostorskih podatkov. Podoben problem je pred štiristo leti rešil René Descartes za geometrične podatke, ki jih je umestil v koordinatni sistem kot osnovo prostorskih referenčnih sistemov. Geometrični podatki, ki so lahko predstavljeni kot karte, načrti, posnetki in podobno, omogočajo uporabnikom geografskih informacijskih sistemov opredelitev prostorskih podatkov glede na njihov položaj in geometrijo, ne pa tudi glede na pomen. Za semantično uvrstitev, opredelitev, prevajanje in integracijo posameznih podatkov ter zbirk bi bilo treba vzpostaviti t. i. semantični referenčni sistem v domeni prostorskih podatkov. Omenjenega problema se je za izhodišče svoje raziskave na filozofsko-logični in tehnično-organizacijski ravni lotil avtor dela Marjan Čeh.

Avtor se v poglavjih od dve do štiri ukvarja z že znanimi teorijami, ki so izhodišče za semantično analizo prostorskih podatkov. Najprej na teoretični ravni podrobno opisuje obstoječe metode semantične integracije podatkov. Nato obravnava metode integracije konceptualnih shem, semantične bogatitve shem, reševanje problemov spajanja shem ter ocenjevanja kakovosti integracije podatkov. Kot orodje mu služijo principi konceptualnega modeliranja. V ta namen podrobno opredeli klasifikacijo in abstrakcijo v konceptualnem modeliranju, formalno modeliranje konceptov ter konceptualno modeliranje z vidika informacijskih sistemov. Uporaba ontologij je v geoinformatiki predmet živahnih raziskav, zato obravnava nekatere filozofske teorije prostora, kot so ontologija, merologija, semiotika ter kognitivno kartiranje.

Peto in šesto poglavje sta v teoretskem smislu najpomembnejši. Avtor je razvil neodvisno metodologijo za izvedbo semantičnega referenčnega sistema, ki jo imenuje »ontologija prostora geografskih razsežnosti«, skrajšano OPGR. Njegova univerzalna ontologija predstavitve znanja o prostoru geografskih razsežnosti je zasnovana kot sistem človeških aktivnosti in prostorskih pojavov. Pri tem uvaja temeljne mere za ocenjevanje semantičnih značilnosti zbirk podatkov, kot so semantična razdalja, globina in razpršenost (horizontalna, vertikalna) zbirke podatkov. Na ta način poda osnove za semantično integracijo zbirk prostorskih podatkov.

V zadnjem poglavju se ukvarja s praktičnimi vidiki pri medopravilni integraciji zbirk prostorskih podatkov. Avtor se nagiba k celostnemu pristopu, ki vključuje geografski prostor in geoinformacijsko tehnologijo. Pomembno je vedeti, da že obstajajo orodja – jeziki, kakršna sta GML in XML, ki omogočajo zapis t. i. semantičnih koordinat zbirk podatkov. S tem so ustvarjeni pogoji za analizo, prevajanje in kasnejšo integracijo prostorskih podatkov. Predlagani koncept organizacije, pri katerem za vsak podatek poznamo semantične značilnosti, kot je globina, omogoča na primer podporo pri odločanju pri samodejni generalizaciji, kar bi lahko s pridom uporabili kartografi. Pri vsakdanjih praktičnih opravilih z zbirkami prostorskih podatkov je treba izpostaviti vlogo skrbnika podatkov. Ta mora najprej sprejeti t. i. ontološko obveznost glede uporabe OPGR. S tem se naveže na semantični referenčni sistem, v katerega uvrsti svojo zbirko podatkov. Sledi t. i. semantično kartiranje, pri katerem dobi vsak podatek semantično koordinato, ki se v obliki oznake vodi kot metapodatek. Enak postopek se izračuna za vse podatke v zbirki. Za zbirko kot celoto se nato izračuna mere semantičnih značilnosti (semantična razdalja, globina, razpršenost ipd.), ki se shranijo kot metapodatek zbirke. Sledi objava omenjenih metapodatkov v porazdeljenem prostorskem informacijskem sistemu. Objava je pomembna za morebitno nadaljnjo integracijo zbirk, skrbništvo ali trženje. Po opisanem postopku se poveča organiziranost podatkov in s tem zmanjša entropija, uporabniki pa se lahko glede na objavljene mere semantičnih značilnosti lažje odločajo o morebitni uporabi zbirke podatkov.

Aktualnost teme o komunikacijah ali medopravilnosti prostorskih zbirk podatkov se z naraščanjem števila digitalnih podatkov veča. To delo morajo poznati skrbniki (načrtovalci) in tudi uporabniki zbirk prostorskih podatkov. Z obravnavano tematiko morajo biti seznanjeni vsi, ki se ukvarjajo s prostorsko informatiko in odločajo o uporabi posameznih zbirk podatkov. K dodelavi enotnega koncepta standarda semantične opredelitve prostorskih podatkov na državni, evropski in svetovni ravni pa bi nenazadnje pripomogla tudi politična podpora. Semantična opredelitev podatkov je z neprimerno bolj raziskano in standardizirano geometrično opredelitvijo pomembna in koristna, saj omogoča lažjo in učinkovitejšo uporabo prostorskih podatkov.

Tomaž Podobnikar,  
ZRC SAZU,  
junij 2003

# 1

## UVOD

Integracija podatkov je z naraščajočimi potrebami in dostopnostjo podatkov iz različnih podatkovnih virov predmet številnih raziskav. Postopki integracije zbirk podatkov usklajujejo podatke iz različnih virov v enotno in skladno predstavitev. Cilj integracije zbirk podatkov je združen pogled na obravnavane zbirke, in sicer preko enotnega vmesnika za dostopanje do porazdeljenih podatkov. Medopravilnost na podatkovni ravni pomeni skladno interpretacijo pomena podatkov pri povezovanju zbirk podatkov – semantično medopravilnost (slika 1).

Vzroki za integracijo podatkovnih virov so raznovrstni. V primerjavi s preteklostjo, ko so se uporabniki osredotočali na vprašanje, kako zagotoviti želene podatke, se s sodobnimi pristopi integracije podatkov in informacij uporabniki lahko osredotočajo na opredeljevanje vprašanj o tem, katere podatke potrebujejo. Namesto iskanja ustreznih virov in vzpostavljanja medsebojne izmenjave podatkov s posameznim podatkovnim virom ter ročnim povezovanjem uporabnik lahko postavlja vprašanja o želenih podatkih na enoten način.

V tem delu bodo obravnavani podatki o prostoru s poudarkom na prostoru geografskih razsežnosti oziroma geografskem prostoru. Prostorski objekti geografskih razsežnosti so objekti, ki so večji od človeškega telesa in jih v celoti ni mogoče zaznati v okviru enega zaznavnega dejanja (Egenhofer in Mark, 1995). Intenzivne raziskave preiskovanja prostora in integracije podatkov namreč potekajo tudi na področju robotike, ki pa niso vezane na geografske razsežnosti prostora. Termina »semantično« in »pomensko« se v tem besedilu uporabljata kot sinonima. Sicer sta besedilu priložena besednjak in stvarno kazalo, ki lahko služita kot pripomoček pri razumevanju pojmov.

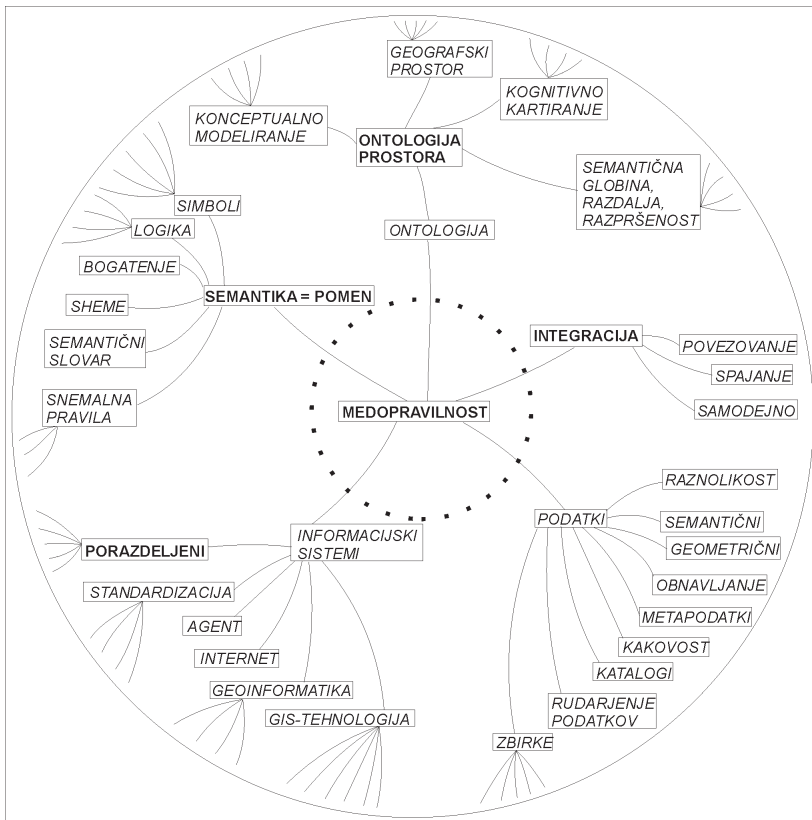
### 1.1 INTEGRACIJA INFORMACIJ

Osnove za povezovanje – integracijo podatkovnih nizov – temeljijo na modelu smiselno povezanih razredov členitve sveta – ontologij. Beseda ontologija se vse pogosteje pojavlja v kontekstu geoinformacijske znanosti. Ontologija se uveljavlja kot sredstvo za izboljšanje dostopa in porazdeljevanja obstoječih geografskih informacijskih virov. Semantika zbirke prostorskih podatkov vključuje kontekst zbirke podatkov – strukturne informacijske ontologije – kot tudi opredelitev ontologije geografskih struktur in procesov – ontologije prostora. Da bi bilo mogoče posledične povezave med podatkovnimi objekti nadgraditi in vnaprej določene povezave, je bilo razvitih več strategij. Integracija shem, ki je bila razširjena na geografske podatkovne nize, je

najpomembnejša metodologija za integracijo zbirk podatkov in je v domeni računalništva. Izgradnja ontologij o geografskih informacijah je zapleten postopek, povezan s porazdeljevanjem znanja o prostoru geografskih razsežnosti in večkratno uporabo znanja. Ontološko raziskovanje presega konceptualno modeliranje. Neformalno se ontologije že dolgo uporabljajo v obliki objektnih oziroma atributnih katalogov. Vendar obstoječe ontologije geografskih informacijskih sistemov vsebujejo pogled na svet, ki se osredotoča na obstoječe vsebine zbirk podatkov, in ne upoštevajo povezav s človeškimi aktivnostmi.

Razvoj konceptualnega modeliranja zbirk podatkov, porazdeljenih informacijskih sistemov in raznovrstnih zvezno porazdeljenih zbirk podatkov je odprl pomembno teoretično vprašanje: »Kako upravljati medsystemsko izvedljivost, medopravilnost heterogene podatkovne semantike?« Zaradi potrebe po porazdeljevanju podatkov med informacijskimi sistemi, katerih znanje o njihovih konceptualnih shemah je poznano, je potrebno odgovoriti na vprašanje: »Kako razvozlati pomensko neskladnost – terminološko, strukturno, vedenjsko in neskladnost razmerij?« Odgovore na ta vprašanja dajejo v tem delu podan pregled in raziskovanje integracije shem ter razvoj metod za pridobivanje in predstavitev semantičnega znanja iz konceptualnih shem obstoječih zbirk podatkov.

Razvoj in prenova zbirk prostorskih podatkov sta z razvojem informacijske tehnologije doživela znaten napredek. Z integracijo geografskih in računalniških tehnologij sta zajem in obdelava



Slika 1: Prikaz osnovnih pojmov. Hiperbolični brkljalnik omogoča dinamično prikazovanje hierarhičnih zasnov na hiperbolični ploskvi. Struktura hierarhije se ohranja tudi pri premiku gledišča opazovanja.

prostorskih podatkov postala globalna po svojem obsegu in univerzalna po uporabnosti informacij. Od povezovanja prenovljenih zbirk podatkov in obnavljanja prostorskih podatkov imajo lahko koristi posamezniki, vladne in nevladne organizacije ter gospodarski in negospodarski državni sektorji. Ker so začetni stroški prenove zbirk prostorskih podatkov in razvoja prostorskih informacijskih sistemov visoki, koristi pa raznovrstne, je smiselno, da strategijo razvoja prostorsko-podatkovne infrastrukture izdela država. Država, ki na najvišji ravni usmerja in koordinira različne skrbnike zbirk podatkov, organizira tudi finančne aranžmaje za izvajanje investicij v razvoj podatkovne infrastrukture.

Zaradi zmanjševanja stroškov javne uprave se iščejo nove možnosti optimizacije tudi v sektorjih državnih prostorskih podatkovnih infrastruktur. Obnavljanje oziroma ažuriranje prostorskih podatkov je ena najpomembnejših postavk pri obravnavi stroškov poslovanja organizacij, ki skrbijo za obnovljenost zbirk prostorskih podatkov. S povezovanjem povezljivih prostorskih tematik se odpirajo možnosti za enkratno izvedbo obnovitve podatkov več različnih zbirk podatkov kakor tudi možnosti zmanjšanja redundance povezljivih podatkov z vidika celotne državne podatkovne infrastrukture.

Za podatkovno povezovanje na vseevropski ravni je treba upoštevati enotno prostorsko referenčno osnovo in primerljive podatkovne strukture. Zasnove zbirk prostorskih podatkov so v evropskih državah različne. Z vzpostavljanjem digitalne tehnologije za upravljanje zbirk podatkov se razvojne usmeritve posameznih držav približujejo, še posebej državam, ki so včlanjene v Evropsko unijo, ter pridruženih članic. Informacijske sestavine, ki naj bi štele za referenčne zbirke podatkov bodoče vseevropske informacijske infrastrukture, so: geodetski referenčni sistem tako položajni kot tudi višinski, administrativne prostorske enote, enote zemljiških pravic, naslovi, izbrane topografske teme, na primer modeli višin, hidrografija, prometna omrežja, in ortorektificirani letalski oziroma satelitski posnetki.

Geodetski podatkovni viri, kot so geodetski referenčni sistem, administrativne prostorske enote, naslovi in topografske teme služijo vsem naštetim podatkovnim zbirkam kot osnova za zajem podatkov in kot osnova za položajno umeščanje tematskih podatkov v prostoru, tako imenovano georeferenciranje. Temeljne strategije geodetske stroke za približevanje potrebam uporabnikov geodetskih podatkov v smislu povezovanja zbirk podatkov so v oblikovanju povezovalnega semantičnega referenčnega sistema kot hrbtenice za povezovanje podatkov. V tem delu sta predstavljeni semantično omrežje – ontologija geografskega prostora za povezovanje prostorskih podatkov kakor tudi metoda za povezovanje podatkov zbirk prostorskih podatkov. Izdelano semantično omrežje prostora omogoča analize obstoječih zbirk podatkov za namene integracije v okolju porazdeljenih informacijskih sistemov.

Ker je pravilno razumevanje podatkov in informacij temeljna paradigma za učinkovito povezovanje zbirk podatkov, je semantična redundanca njihova zaželena lastnost. Redundanca pomeni, da se določeno obvestilo prenaša z več znaki, prvinami, kot je nujno, zato semantična redundanca podatkov povečuje verjetnost pravilnega razumevanja informacije. Za namene povezovanja zbirk podatkov se jo spodbuja z metodami bogatenja podatkovnih shem in z dodajanjem podatkov o veljavnih razmerjih med podatki. Problem ni v redundanci podatkov, ampak v večkratnosti obnavljanja podatkov za isti stvarni objekt. Za nekatere vrste objektov velja, da se podatki o njih vzdržujejo in obnavljajo v različnih zbirkah prostorskih podatkov, ki so v pristojnosti različnih skrbnikov – organizacij, oddelkov. Z metodami odkrivanja semantične podobnosti oziroma skladnosti podatkov je mogoče ugotoviti sorodne ali celo enakopomenske koncepte v različnih zbirkah podatkov. Sorodne koncepte je treba podrobneje analizirati glede na možnosti integriranega obnavljanja podatkov.

Integrirano obnavljanje podatkov enakovrednih konceptov v porazdeljenih informacijskih sis-

temih bi povečalo učinkovitost informacijskih sistemov z vidika zmanjšanja stroškov za zajem novih podatkov. Več organizacij skrbnic zbirke podatkov bi s svojimi storitvami zajemanja svežih podatkov prispevalo k večji obnovljenosti integrirane zbirke podatkov. Drugi vidik zmanjšanja stroškov za zajem novih podatkov bi bil ta, da bi po ugotovljeni večkratnosti obnavljanja podatkov istega koncepta v različnih organizacijah vse organizacije, razen ene, lahko opustile obnavljanje teh podatkov in preko porazdeljenega informacijskega sistema prešle na njihovo integrirano uporabo.

V primerih pomensko sorodnih a neenakovrednih prostorskih konceptov bi bila potrebna usklajevanja za doseganje čim večje skladnosti, saj je to predpogoj za vzpostavitev učinkovitega integriranega sistema za obnavljanje podatkov. Za predstavitev in utemeljevanje različnih vidikov posameznih prostorskih pojavov v geografskem prostoru je obravnavana ustreznost formalnih teorij: ontološka ustreznost, epistemološka ustreznost, ustreznost abstrakcije, ustreznost predstavitve, ustreznost utemeljevanja, razlaganja.

Študij ontologij z vidika filozofije in računalniških znanosti lahko pomembno prispeva k združevanju različnih konceptualizacij geografskega prostora v temeljno ontologijo prostora geografskih razsežnosti. Postavljajo se vprašanja, ali ontologija omogoča reševanje ugotovljenih problemov semantične medopravilnosti ali ni le nova fraza za obravnavo semantičnih omrežij. Ontologije geografskega prostora vsebujejo raznolikosti, kot so razlike v ravni podrobnosti, razlike razmerij med konceptnimi tipi, profislovja v opredelitvah konceptov in nepopolni opisi značilnosti konceptov.

Za izmenjavo informacij med različnimi geografskimi podatkovnimi zbirkami je potrebno razviti primerno metodo za formalno opredeljevanje in predstavljanje znanja o prostoru geografskih razsežnosti. Preoblike in raznolikost podatkovnih standardov in terminologije, ki predstavljajo različne geografske koncepte in podatkovne koncepte, še povečujeta težave pri porazdeljevanju in večkratni uporabi geografskih informacij.

## 1.2 SEMANTIČNA MEDOPRAVILNOST

Semantična medopravilnost med računalniškimi sistemi se doseže s prevajanjem pomena med različnimi metodami modeliranja, jezikov in programskih orodij. Pri tem se kot izmenjevalni format uporabi ontologija. V tem kontekstu je ontologija opis konceptov porazdeljevanja informacij in razmerij med njimi, torej razmerij, ki obstajajo za agenta ali skupnost agentov. Ta opredelitev ontologije se od opredelitve v filozofskem smislu razlikuje v tem, da ne poudarja, komu je namenjena. Koristnost uporabe ontologije v sistemskem inženiringu se izraža v naslednjih značilnostih: ponovna uporabnost, pospešeno iskanje podatkov in informacij, povečana zanesljivost programske opreme, podrobnost opisa zahtev, enostavno vzdrževanje dokumentacije, hitrost pridobivanja znanja. Da bi bilo mogoče predstaviti znanje o določeni domeni, mora biti ta podprta z ontologijo, ki opredeljuje kategorije stvari v tej domeni in opredeljuje izraze, ki jih ljudje uporabljajo v pogovoru o kategorijah. Integracijo je mogoče izvršiti le, če so posamezne ontologije vstavljene v bolj splošno visokonivojsko ontologijo, ki omogoča zanesljivo osnovo za posebne uporabe (Sowa, 2000b). Postopek integracije več ontologij je razdeljen v dva glavna koraka: semantično razstavljanje na faktorje – faktorizacija – in izdelava omrežja konceptov. Formalni semantični podatkovni model prostorskih objektov sestavljajo formalne opredelitve geometričnih struktur geografskih pojavov in metode za upravljanje geometrične strukture zbirke

podatkov. Razvoj tovrstnega podatkovnega modela izhaja iz teorij o značilnostih prostorskih podatkov. Ko je za opis geometričnih struktur razvit dosleden matematični formalizem, je mogoče razviti posamezne koncepte podatkovnega modela, vključno s strukturo kompleksnih objektov (Kemppainen, 1995). Na osnovi natančnega pregleda in analize obstoječih zbirk prostorskih podatkov, od njihovih najstarejših različic pa do sodobnih zbirk prostorskih podatkov, je mogoče trditi, da se integracija podatkov v zbirkah prostorskih podatkov pojavlja že od njihovega nastanka naprej. Potencial integracije podatkov in zbirk podatkov pa je narasel predvsem z nastankom digitalne tehnologije in – še bolj specifično – z uveljavljanjem prostorske informacijske tehnologije, katere najznačilnejši predstavnik je geografski informacijski sistem.





# 2

---

## SEMANTIČNA INTEGRACIJA PODATKOV

Semantična integracija podatkov in informacij v spoznavnem, kognitivnem smislu je poenotenje (Anderson, 1981). Postopek spoznavanja besed je rezultat integracije različnih gesel, miselnih zvez, namigov. Na višji ravni je razumevanje pomenov terminov tudi rezultat integracije miselnih zvez. Integracija se povezuje z vrednotenjem v verigo integracijskih postopkov. Integracijski postopek na eni ravni zasnuje postopek vrednotenja za drugo raven obravnave. Pri zagotavljanju združenega pristopa do obstoječih zbirk podatkov se na osnovi lokalnih shem gradijo ustrezne podsheme, ki se povezujejo v enotno integrirano shemo (Ceri in Pagelatti, 1984).

Podatkovni viri, ki naj bi se združili, so lahko v pristojnosti ene organizacije ali pa so to poljubni viri, dosegljivi preko spleta (Convey et al., 2001). Posamezni podatkovni viri so neodvisno zasnovani in predani v samostojno uporabo. Podatkovni viri niso nujno zbirke podatkov. To so lahko tudi podatki v zastarelih informacijskih sistemih, ki jih je težko prenesti v sodobne računalniške arhitekture, ali pa podatki v posameznih manj urejenih datotekah (Embley, 1998). Za podatkovno integracijo je treba razrešiti razlike v modelih, semantiki, možne nedoslednosti in razlike v sposobnosti podatkovnih virov za integracijo.

Integracija zbirk podatkov je pomemben gradnik aplikacij za podporo odločanja, kot sta neposredno analitično procesiranje in rudarjenje podatkov, ter za podporo prostorskemu odločanju z geografskimi informacijskimi sistemi (GIS). Neposredno analitično procesiranje s postavljanjem večjega števila združenih vprašanj in analiziranjem podrobnih podatkov, zbranih iz več različnih virov, podpira poslovno odločanje. S tehnikami podatkovnega rudarjenja se odkriva znanje iz obsežnih podatkovnih virov. Iz neobdelanih podatkov se samodejno odkrivajo pravila in vzorci v podatkih. Podpora prostorskemu odločanju se izvaja na osnovi geolociranih zbirk podatkov in z orodji geografskih informacijskih sistemov.

Podatkovne vire, za katere se načrtuje integracija podatkov, je smiselno obravnavati ob upoštevanju njihovih značilnosti, ki so pomembne v postopkih povezovanja oziroma združevanja. Kriteriji za klasifikacijo podatkovnih virov za namene integracije podatkov so neodvisnost, raznovrstnost, porazdeljevanje in odkritost (Convey et al., 2001).

Neodvisnost se nanaša na stopnjo, s katero posamezni podatkovni vir lahko neodvisno deluje: oblikovna neodvisnost pomeni, da je podatkovni vir neodvisen od podatkovnega modela, poimenovanja podatkovnih elementov, pomenske razlage podatkov in omejitev; komunikacijska neodvisnost pomeni, da je vir neodvisen v odločanju glede zagotavljanja podatkov drugim sestavnim delom integriranega sistema, ki jim posreduje odgovore na poizvedovanja; izvršilna neodvisnost pa pomeni, da je vir neodvisen pri izvrševanju in razporejanju prihajajočih povpraševanj. Raznovrstnost se nanaša na stopnjo različnosti med podatkovnimi viri, ki so sestavni deli

integriranega podatkovnega sistema. Raznovrstnost se odraža na različne načine: tehni na raznovrstnost izvira iz različnih strojnih osnov, operacijskih sistemov, mrežnih protokolov in podobnih zasnov; konceptualna raznovrstnost izvira iz različnih podatkovnih modelov, programskih jezikov, različnega razumevanja in konceptualnega modeliranja istih stvarnih objektov. Porazdeljevanje podatkov se nanaša na predmetno mnogovrstna prizorišča: odjemalec/strežnik – informacijska arhitektura, kjer strežnik izvaja upravljanje podatkov in odjemalec zagotavlja uporabniški vmesnik. Odkritost se nanaša na ločevanje višje ravni semantike sistema od nizke ravni izvajalnih postopkov sistema: transparentni sistem skriva izvajalne postopke pred uporabnikom.

Ločujemo tri temeljne pristope k integraciji zbirk podatkov (Barja et al., 1998): pristop navideznega pogleda – pri tem pristopu ni integralno shranjenih podatkov, ti so na zahtevo dostopni iz lokalnih zbirk podatkov; pristop snovnega pogleda – podatki so vnaprej shranjeni v integralnem podatkovnem skladišču in so pripravljene na poizvedovanje brez komunikacije z lokalnimi podatkovnimi zbirkami; hibridni pristop – značilno je, da se glede na pogostost uporabe določijo podatki, ki se vnaprej shranijo v integralni zbirki podatkov.

Navideznemu pristopu integracije zbirk podatkov sta lahko osnova dva različna informacijska sistema, in sicer zvezni informacijski sistem ali vmesni, posredovalni sistem. Posredovalni sistem omogoča uporabniku, da pridobi informacije iz različnih vrst podatkovnih virov brez poznavanja identitete, lokacije nahajanja podatkov, shem, dostopnih postopkov in pomenov podatkovnih virov (Barja et al., 1998). Posredovalni sistem, ki skrbi za posredovanje podatkov, vsebuje enotno posredovalno shemo. Sprejema povpraševanja, ki so izražena skladno s posredovalno shemo, in upravlja vse postopke za pridobivanje primernih podatkov iz ustreznih podatkovnih virov. Vsak od podatkovnih virov deluje v skladu s samosvojo, lokalno shemo. Za prevajanje povpraševanj iz oblike, ki je razumljiva posredovalni shemi, v obliko, ki jo razumejo posamezni lokalni podatkovni viri, se uporabljajo programski paketi, imenovani ovojniki (wrapper). To so programska orodja, ki so del posameznega podatkovnega vira. Ovojniki prikrivajo tehnične in pomenske raznolikosti zbirk podatkov pred uporabniki. Prevajajo povpraševanja v obliko, ki je skladna s lokalno shemo, in pretvarjajo rezultate povpraševanj v obliko, razumljivo posredovalnemu sistemu.

Pristop snovnega pogleda integracije zbirk podatkov temelji na podatkovnem skladišču, v katerem so shranjeni podatki, ki so bili pridobljeni iz različnih podatkovnih virov. Podatki so usklajeni z enotno shemo. Pristop snovnega sistema z uporabo podatkovnega skladišča se uporablja zaradi velikih količin podatkov, ki jih je treba pregledati ob povpraševanju, in zato, ker so povpraševanja po informacijah obsežna, sestavljena in zapletena. Takšen pristop omogoča neposredna obdelava poslov (On-Line Transaction Processing – OLTP). Hibridni pristop združuje posamezne prednosti navideznega in snovnega pogleda. Tako obstaja podatkovno skladišče za del pogosto obravnavanih podatkov. Preostale podatke posreduje uporabniku posredovalni sistem povpraševanja preko posredovalne sheme in lokalnih shem.

V raziskavah samodejne integracije informacij z vidika semantike je cilj v ohranitvi pomena informacijskih entitet v različnih informacijskih okoljih. Lokalni model semantike zbirke podatkov predstavlja formalno osnovo za zvezno porazdeljene zbirke podatkov. Zvezno porazdeljena zbirka podatkov je sestava porazdeljenih, redundantnih, delnih in delno avtonomnih zbirk podatkov, ki jih upravlja zvezno porazdeljeni informacijski sistem. Porazdeljenost v tem primeru pomeni, da so zbirke podatkov zvezno porazdeljene, shranjene na različnih informacijskih sistemih, vsak od sistemov pa vsebuje njemu lastni del znanj. Redundanca v tem primeru pomeni, da je lahko isti del znanja predstavljen v več kot eni zbirki podatkov (Ghidini in Serafini, 1998). Za doseganje virov podatkov je mogoče razlikovati med operacijami porazdeljevanja in operacijami izmenjave. Pri porazdeljevanju se agenti – ljudje ali računalniki – sklicujejo na skupne

dele virov podatkov. Skupni podatki ostajajo izven aplikacij, ki si podatke porazdeljujejo. Izjema je primer, ko se interni podatki določene aplikacije preko sklicevanja zunanjih aplikacij nanje, porazdeljujejo v zunanje aplikacije. Ta postopek poteka preko standardnega vmesnika za doseganje podatkov (Standard Data Access Interface – SDAI); (ISO 10303-21, 1994). Izmenjava virov podatkov poteka s prenašanjem podatkov med aplikacijami.

## 2.1 INTEGRACIJA SHEM ZBIRK PODATKOV

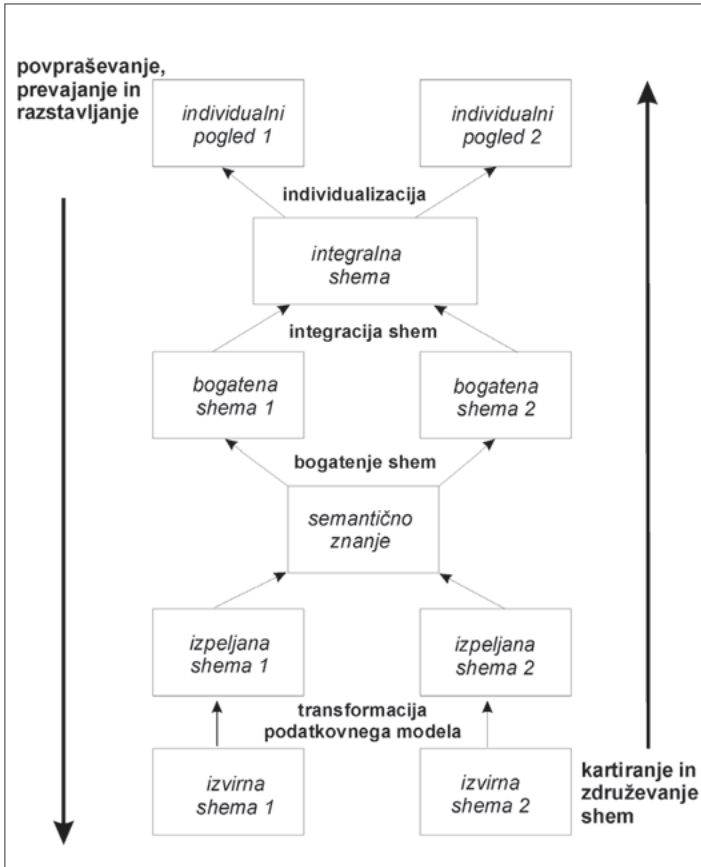
Pri zagotavljanju združenega pristopa do obstoječih zbirk podatkov se na osnovi lokalnih shem gradijo ustrezne podsheme, ki se povezujejo v enotno integrirano shemo (Ceri in Pagelatti, 1984). Integracija shem je opredeljena kot aktivnost združevanja shem obstoječih ali bodočih zbirk podatkov v enotno, integrirano shemo (Battini et al., 1986). Cilj raziskav integracije shem je odkrivanje semantičnih in strukturnih podobnosti in protislovij med shemami, podpora uporabnikom za reševanje protislovij in izdelava skladne globalne sheme. Raziskovanje integracije shem se sooča s problemi pridobivanja in predstavitev semantičnega znanja. Pridobivanje semantičnega znanja je povezano z odkrivanjem semantike simbolov, s katerimi so označeni stvarni objekti, z odkrivanjem semantične podobnosti in protislovij, z izdelavo arhitekture za orodja, namenjena integraciji shem in izdelavi semantičnih slovarjev (Song, 1995).

Integracija shem je tesno povezana s konceptoma semantične medopravnosti in raznovrstnosti – heterogenosti zbirk podatkov. Semantična medopravnost je sposobnost usklajenega sodelovanja med semantično raznovrstnimi sistemi zbirk podatkov (slika 2). Raznovrstnost sistemov zbirk podatkov se sestoji iz razlik v podatkovnih modelih in iz semantične raznovrstnosti. Razlike med podatkovnimi modeli izvirajo iz posebnosti struktur zbirk podatkov, ki jih zagotavljajo podatkovni modeli in konceptualni formalizem.

Za doseganje medopravnosti informacijskih sistemov se uporabljajo tehnike prevajanja shem za povezovanje shem (Sheth in Larson, 1990). Ključni problem pri integraciji shem je določitev pogojev, izraženih s skladnostnimi trditvami, pri katerih sta lahko dve shemi smiselno združeni. To pomeni določitev načina za spajanje dveh shem v integrirano globalno shemo, ki bo imela enako informativno sposobnost kot izvorni shemi. Integralna shema mora predstavljati toliko pomena kot izvorni shemi. To je mogoče doseči s postavitvijo omejitev shemam, namenjenim integraciji. Omejitev, ki se imenuje odsotnost nesoglasij, postavlja pogoj, da pravila, ki veljajo v eni shemi, skupaj z odgovarjajočimi trditvami ne smejo omejevati pravil in trditev v drugi shemi (Ekenberg, 1995).

Semantična relativistična teorija (Song, 1995) predpostavlja, da je mogoče z uporabo pripadajočih semantičnih razmerij identificirati koncept in koncepte, ki so s tem konceptom v odnosu. Teorija predstavlja čvrsto osnovo za semantične analize konceptualnih shem. Semantični relativizem podpira mnogokratne poglede ali tolmačenje istih entitet objektnega sistema.

Semantična sorodnost združuje semantično enakovrednost in združljivost, pri kateri se dva ali več konceptov v shemah nanašata oziroma nanašajo na enake ali podobne razrede. Semantična sorodnost je pri integraciji shem zbirk podatkov opredeljena s klasifikacijo semantičnih razmerij. Song (1995) navaja štiri stopnje semantičnih razmerij: šibko, znosno, enakovredno in spojitveno. Med koncepti obstaja več tipov semantičnih razmerij: istovetno (natančno enak), enakovredno (uporaba neenakih, a enakovrednih sestavin ter uporaba istih spoznanj), skladno (uporaba neenakovrednih, a ne nasprotujočih si sestavin), neskladno (vsi elementi nasprotujoči).



Slika 2: Omrežje za doseganje semantične medopravilnosti zbirk podatkov (privzeto po Fankhauser, 1997)

Razmerja med shemami so pri upoštevanju ravni integracije lahko štirih različnih vrst (Sheth in Larson, 1990):

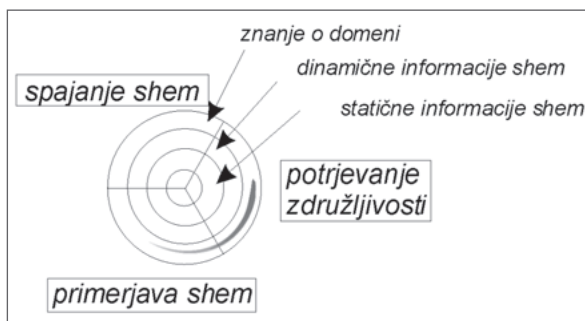
- identične – popolnoma prekrivajoče se,
- podobne – delno prekrivajoče se,
- sorodne – neprekrivajoče se vendar sorodne,
- različne – neprekrivajoče se in nesorodne.

V primerih podobnih in sorodnih shem je potrebno odkriti obstoječe podobnosti v razmerjih med shemami ter različnosti in neskladja v omrežju povezav med shemami. Za primere identičnih in različnih razmerij med shemami je odkrivanje podobnosti in neskladij enostavno opravilo. Razlikovati je mogoče med dvema vrstama integracije shem, ki sta: integracija pogledov in integracija zbirk podatkov. Integracija pogledov je del oblikovanja zbirke podatkov, kjer se izdelata integralni konceptualni opis bodoče zbirke podatkov. Integracija zbirk podatkov je značilnost porazdeljenih informacijskih sistemov, ko se ustvari globalna shema več obstoječih zbirk podatkov (Ozsu, 1990).

Za reševanje problemov navedenih vrst integracije shem so razviti ločeni pristopi. Prevladujoči pristop pri integraciji pogledov je bolj ali manj samodejen postopek izvedbe integralne sheme iz niza integracijskih trditev, ki se nanašajo na ustrezne elemente v pogledu. Pri integraciji

zbirke podatkov je prevladujoči pristop zagotovitev operacij za razstavljanje, ki omogoči gradnjo integralne sheme iz izvornih shem (Motro, 1987). Z operacijami razstavljanja je mogoče spremeniti hierarhijo konceptov ali pa sestavo razmerij. Za zniževanje stroškov razvoja shem in povečanje kakovosti integracije shem je pomemben razvoj iskanja vzorcev v konceptualnih shemah (Fowler, 1997). Tehniko za pridobivanje vzorcev shem na osnovi semantičnih informacij z metodo kopičenja v umetnih nevronske mrežah je razvil Wen (1998).

Konceptualne sheme so strukture za opis konceptualnega modela. Pri tem vsebujejo več informacij glede objektov in odnosov med njimi kot podatkovne sheme. Ustvarjanje konceptualne sheme iz podatkovne sheme se imenuje obrnjeno inženirstvo, ki ga navaja Šumrada (1993). V tem postopku se semantične informacije pridobivajo na osnovi podatkovne sheme. Manjkajoče semantične informacije se delno pridobijo na samodejen način in delno s postopki miselnega sklepanja, ki jih izvaja izkušen analitik. Zaželeno je visoka stopnja samodejnosti obrnjenega inženirstva, vključenega v opravila integracije shem in kartiranja globalnih shem. Veliko sodobnih objektno usmerjenih orodij za analizo in oblikovanje (Object Oriented Architecture and Design – OOAD; v preteklosti Computer Assisted System Engineering – CASE) že vsebuje postopke obrnjenega inženirstva in omogoča ročno ali povsem samodejno izvedbo obrnjenega inženirstva. Za samodejno integracijo konceptualnih shem zbirke podatkov je potrebno zgraditi semantični slovar in opredeliti kriterije za presojo semantične podobnosti. Za izdelavo semantičnega slovarja je potrebna osnova, ki se imenuje semantično omrežje. Semantično omrežje je sestavljena predstavitev semantičnih razmerij med objekti ali pojavi v obravnavanem stvarnem svetu. Pri semantični bogatitvi shem za namene primerjave shem se jasno poudarijo različne vrste »skritih« informacij, ki zadevajo semantiko modeliranih objektov/pojavov stvarnega sveta. Metode za semantično bogatitev shem vključujejo uporabo dodatnih virov znanja, poleg znanja, ki ga že vsebujejo statični deli konceptualnih shem. Vire dodatnega znanja za semantično bogatitev shem je mogoče prikazati glede na osnovne faze integracije shem.



Slika 3: Ravni znanja pri integraciji shem (privzeto po Hakkarainen, 1999)

Sektorji kroga se lahko razlagajo kot faze integracije shem: primerjave shem, prilagajanje shem in spajanje (slika 3). Krožni pasovi predstavljajo tipe znanja na različnih ravneh: statične, dinamične informacije in znanje o področju obravnave, ki ni del sheme.

### 2.1.1 Reševanje problemov pri integraciji konceptualnih shem

Konceptualne sheme so izdelek konceptualnega modeliranja. Sheme so formalne predstavitve uporabniških abstrakcij v modelih stvarnosti kakor tudi odslikave zbirke podatkov na pojavnih ravni.

Semantična neskladnost obstaja, kadar dva koncepta, ki predstavljata isto entiteto v stvarnem svetu ali imata enako opredelitev – simbol, dveh različnih shem vsebujeta protislovne podatke – različne podaljšane izjave. Nesoglasja med shemami v podatkih povzročajo napačen odraz stvarnosti in odpirajo različna problemska področja (Song, 1995): odkrivanje podobnosti in nesoglasij konceptov je problem, katerega vzrok se skriva v težavnosti odkrivanja semantike simbolov, ki naj bi se primerjali; nerazločnost konceptov in njihovih predstavitev uvaja problem določitve količinskih kriterijev semantične podobnosti za premagovanje nerazločnosti; neskladnost predstavitev konceptualnih shem je problem, ki nastopi zato, ker različni izdelovalci shem privzamejo za modeliranje istih entitet stvarnega sveta različne načine opisovanja in simbolov. Podporne tehnike reševanju problemov integracije konceptualnih shem so sledeče: pridobivanje dodatnega semantičnega znanja, ki se izvaja zaradi nepopolnosti in ohlapnosti uporabniških zaznav stvarnosti; izgradnja semantičnega slovarja obravnava vprašanja določitve strukture in vsebine semantičnega slovarja; bogatitev shem – semantično obogatena shema izraža več semantike objektnega sistema kot izvorna shema, kar omogoča nove načine reševanja problemov raznolikosti semantike; upoštevanje dinamičnih vidikov stvarnosti; obstajajo formalizmi, ki že podpirajo opredeljevanje vedenjskih in dinamičnih vidikov, ki bi morali biti vključeni v vsak semantično kakovosten opis kateregakoli področja obravnave; ocenjevanje kakovosti integracije izpostavlja vprašanje načina testiranja kakovosti integracije konceptualnih shem.

### 2.1.2 Odkrivanje podobnosti in nesoglasij konceptov

Pomenska neskladnost med simboli različnih konceptualnih shem je neizogibna. Zato je temeljna naloga integracije shem odkrivanje podobnosti in nesoglasij. Odkrivanje semantične podobnosti in nesoglasij simbolov sta dva vidika istega problema. Splošni postopek odkrivanja podobnosti in nesoglasij je sledeč:

- odkrivanje pomena simbolov z uporabo informacij o objektih, ki jih simboli predstavljajo, o njihovih poimenovanjih, lastnostih, razmerjih in z uporabo informacij o drugih objektih, ki so z njimi povezani;
- določitev semantičnih razmerij med primerjanima simboloma, kar pomeni določanje, ali sta simbola med seboj semantično istovetna, enakovredna, skladna ali neskladna.

Nerazločnost konceptov in njihovih predstavitev uvaja problem določitve količinskih kriterijev semantične podobnosti za premagovanje nerazločnosti. Kvantitativne meritve semantične podobnosti so predmet raziskav semantične integracije (Bouzeghoub in Comyn-Wattiau, 1990). Metode merjenja semantične podobnosti so metode merjenja semantične razdalje med primerjanimi simboli. Za predstavitev semantične bližine oziroma oddaljenosti se vsakemu primerjanemu paru simbolov pripiše izmerjena oddaljenost z intervala (0, 1). Vrednost 1 pomeni semantično istovetnost, vrednost 0 pomeni popolno semantično različnost. Slabost te metode je, da meritve niso nepristranske. Za zmanjšanje pristranosti meritev semantične razdalje je bila predlagana metoda, pri kateri se miselne povezave uvrstijo v več skupin. Vsaka skupina povezav vsebuje povezave konceptov, ki imajo enake lastnosti. Za vsako skupino se prešteje število konceptov z enakimi lastnostmi, ki jih vsebuje skupina, in na tej osnovi se skupinam pripišejo ustrezne uteži. Semantična raznovrstnost ima vzrok v različnosti in zapletenosti človeškega zaznavanja in izražanja stvarnosti oziroma področja obravnave. Semantična raznovrstnost se pojavlja v različnih oblikah, kot so terminološka, strukturna, vedenjska in relacijska nesoglasja. Nesoglasje poime-

novanja (Battini in Ceri, 1992) nastopi zaradi napačnega razumevanja in napačne uporabe simbolov za predstavitev konceptov. V zvezi z neskladnostjo predstavitev konceptualnih shem obstajajo trije jezikovni viri semantičnih nesoglasij: sopomenke (sinonimi) – dva različna simbola predstavljata isti koncept; enakozvočnice (homonimi) – dva različna koncepta sta predstavljena z istim simbolom; ciklične generalizacije – generalizacijska razmerja tipa (X je Y) in nato (Y je X), v shemah, ki naj bi bile integrirane.

Strukturna neskladnost predstavitev konceptualnih shem nastopi, kadar je isti pojav področja obravnave modeliran v dveh shemah z različnimi konstrukcijskimi elementi shem (objekti, razredi, atributi, razmerja – odnosi). Vedenjska neskladnost nastopi, kadar je isti pojav področja obravnave v dveh shemah modeliran z različnimi dogodki, dinamičnimi pravili in različnimi omejitvami. Neskladnost razmerij nastopi, kadar dve shemi opisujeta isto področja obravnave, vendar z različnih vidikov.

### 2.1.3 Pridobivanje semantičnega znanja

Obstaja več problemov pridobivanja semantičnega znanja: nepopolnost in ohlapnost uporabniških zaznav stvarnosti, nasprotujoče predstavitve konceptov, monogopomenskost simbolov, uporabljenih v shemah, oporečne razlage in opredelitve simbolov. Najpogosteje uporabljen pristop za pridobivanje semantičnega znanja je preučevanje sveta obravnave oziroma ustrezne ontologije domene. Pridobivanje semantičnega znanja neposredno iz obstoječih podatkovnih shem je zelo intuitiven proces.

Za pridobivanje semantičnega znanja se razvije integracijsko orodje. Zaželeno je, da je samodejno integracijsko orodje čimbolj »razumno«, tako da lahko samodejno odkrije in razreši čim večje število semantičnih nesoglasij, poleg tega naj orodje v primerih nerazrešenih nesoglasij ponudi uporabniku čim več informacij za podporo odločanju. Integracijsko orodje mora vsebovati določeno znanje, na primer slovar sopomenk. Za izdelavo razumnega integracijskega orodja obstaja več uporabnih metodologij: izdelava integracijske zbirke podatkov semantičnega znanja iz različnih virov (lokalne konceptualne sheme, integracijske izkušnje, uporabniške trditve); pridobitev znanja in izkušenj s strani področnih strokovnjakov s pomočjo tehnik ekspertnih sistemov (Hayne, 1990); določitev kriterijev semantične podobnosti in količinska predstavitev semantike simbolov z uporabo teorije mehkih množic in teorij dokazovanja v metodah umetne inteligence; izgradnja semantičnega slovarja.

Izgradnja semantičnega slovarja je glavni del zbirke podatkov semantičnega znanja. Semantični slovar je podpora integracijskemu orodju za pregledovanje simbolov: terminov, imen, besed, alfanumeričnih znakov, ki predstavljajo objekte v lokalnih shemah, ki se povezujejo. Semantični slovar predstavlja tudi zgodovinski pregled že izvedenih integracij, saj lahko beleži uporabniške trditve in rezultate razmišljanj in ugotovitve, ki jih je mogoče uporabiti v prihodnjih izvajanjih integracije. Pomembno je določiti strukturo in vsebino semantičnega slovarja. Najpogostejši so naslednji tipi strukture semantičnega slovarja: razširjeni slovar sopomenk – opredeljuje dejstva, ki jih dva sopomenska simbola predstavljata v istem ali podobnih konceptih; struktura sistematike – taksonomija – vsebuje dejstva, ki opredeljujejo, da imata dva simbola, ki predstavljata dva koncepta, v isti navedbi posploševalno – generalizacijsko razmerje (na primer knjiga, publikacija); zbirka pravil vsebuje trditve, ki podajajo napotke za sklepanje pri integraciji shem (na primer: če je možna združitev dveh entitet, potem naj se združijo tudi njune istovetne lastnosti). Brez poznavanja miselne povezave je z ontologijo težko presojati, ali dva simbola predstavljata isti koncept ali ne. V pomoč je lahko semantični opis, ki poleg opredelitve simbola

podaja natančnejši opis koncepta, ki ga predstavlja simbol, predvsem dovoljene operacije nad simboli. Semantični opis se uporablja pri semantični bogatitvi shem.

## 2.2 Semantična bogatitev shem

Predpostavlja se, da bogatejši, kot so opisi sveta obravnave v določenem časovnem trenutku, bolj natančna znanja o svetu obravnave je v določenem poznejšem časovnem trenutku mogoče pridobiti s tolmačenjem opisov. Ni dokazov, da so izvedena tolmačenja modelov edine obstoječe oziroma edine popolne zamisli stanj razmerij v stvarnem svetu. En od mogočih načinov reševanja problemov raznolikosti semantike pri integraciji shem je semantična bogatitev shem (Hakkarainen, 1999).

Semantično obogatena shema izraža več semantike objektnega sistema kot izvorna shema. To se doseže z uvajanjem dodatnih podatkov, pridobljenih iz informacij o obravnavanem področju – domeni. Tehnike za semantično bogatenje shem so raznolike, saj izvirajo iz različnih teorij: iz koncepta hierarhije abstrakcij, porazdeljene ontologije, jezikovnega znanja, mehkega izrazoslovja, semantičnega sosledstva, abstrakcij domen in razširitvene analize. Metode za semantično bogatitev shem so – glede na obseg nameravane avtomatizacije integriranja shem – uvrščene v dve vrsti: enoglasne, monotonične, in sistematične obogatitve shem. V postopku enoglasne obogatitve shem gre za samodejno pretvorbo semantično nezadostne sheme zbirke podatkov v semantično bogatejši jezik konceptualnega modeliranja (Metais et al., 1996, Mirabel, 1995). Pri sistematični bogatitvi shem gre za podporo postopku specifikacije z dodatnimi metodami in jeziki za modeliranje (Sutcliffe, 1994). Če so sheme obdelane z enako tehniko oziroma metodo bogatitve, potem semantična bogatitev lokalnim shemam navidezno vsili globalni pogled. Uporaba dodatnih znanj obravnavanega področja zahteva natančne napotke za pridobivanje dodatnih informacij za neposredno izvajanje bogatitve shem. Splošnih pravil za pridobivanje dodatnih znanj o objektnem sistemu ni, vendar so uporabna dognanja raziskav predstavitev znanja v kognitivni znanosti, splošni formalizmi za predstavitev znanja v ekspertnih sistemih, razvoj konceptualnih grafov ter jezikovni instrumenti.

Semantične bogatitve shem za integracijo pogledov in integracijo zbirk podatkov imajo pomembne razvojne možnosti v združeni uporabi konceptualnih grafov kot jezikovnih instrumentov. Semantična omrežja ponujajo intuitivno predstavitev splošnih znanj. Predstavljajo linearno semantiko logike in naravnih jezikov kot sestavo vozlišč in povezav. Sodobni tip semantičnega omrežja je s teorijo konceptualnih grafov razvil Sowa (1984, 1986). Pri semantičnem bogatenju shem je pomembno, da se uporabi tako splošen formalizem za predstavitev objektnega sistema, da je mogoče semantiko shem opisati na različnih konceptualnih ravneh z različnimi tipi shem in za različne podatkovne modele. Semantično bogatenje shem je pomembno tudi za druge faze integracije shem. Uporabno je pri izboljšavi kakovosti integracije shem kakor tudi pri njihovem spajanju. Obogatene lokalne sheme so namreč pogosto povezane s splošnim modelom oziroma shemo in tako skupaj tvorijo enotno navidezno globalno shemo.

Hakkarainen (1999) predpostavlja: »Več kot je eksplicitno navedene semantike o dinamičnih sestavah sheme, bolj izvedljiva postajata primerjava shem in spajanje shem« Za izvedbo semantične bogatitve shem z upoštevanjem dinamičnih vidikov shem predlaga tehniko jezikovnih instrumentov. Tako je mogoče zagotoviti sistematični način za izražanje povečane semantike objektnega sistema, kot ga vsebuje izvorna shema. Informacije o dinamičnih vidikih vključujejo



vidike vedenja, kot so aktivnosti v objektnem sistemu, izražene v obliki akcij, dogodkov in pravil dinamike.

Pomanjkanje znanja o vedenjskih navzkrižjih in načinih vplivanja teh navzkrižij na rezultate primerjave shem med drugimi navaja tudi Bouguettaya (1998). Na področju konceptualnega modeliranja (Mylopoulos, 1998) obstoječi formalizmi že podpirajo opredeljevanje vedenjskih in dinamičnih vidikov, ki bi morali biti vključeni v vsak semantično dovršen opis kateregakoli področja obravnave. Hakkarainen (1999) za podporo medopravilnosti v informacijskih sistemih uvaja metodo primerjave dinamičnih struktur shem v konceptualnem modelu.

Strukture shem implicitno opisujejo dinamično sestavino znanja v shemah, in to v obliki notranjih jezikovnih lastnosti, kot sta slovnični čas in časovna zaporednost, kar se imenuje časovni vidik shem. Časovni vidik shem označuje časovna razmerja med različnimi tvorniki dogodkov. Časovni vidik je opredmetena povezava med določenimi stanji, situacijami in odločitvami, ki jih je mogoče sprejeti z ozirom na stanja. Informacije o vedenju in dinamiki v konceptualnih modelih v grafičnih predstavitvah shem niso vidne. Cilj raziskovanja je prikazati dinamične informacije tako, da jih je mogoče uporabiti pri analiziranju shem.

Pri podrobnejšem opisovanju, specifikaciji dogodkov se opisujejo možne spremembe v stanjih mirujočih, statičnih sestavov. V konceptualnem modeliranju se v ta namen uporablja koncept s tremi sestavnimi deli: glava dogodka, pogojni del in sklepni del. V glavi dogodka se navajajo povezani elementi dogodka. Pogojni del določa pogoje za veljavnosti dogodka. V sklepnem delu pa se opredelita učinek dogodka in način za opredmetenje učinka dogodka (Kangassalo, 1993). Tehnika opredmetenja razmerij v konceptualnem modeliranju uvaja časovno oznako razmerij kot jasno določeno lastnost konceptov. Opredmetenje razmerij omogoča zgodovinske analize shem (Wohed, 1997). Sodobni pristopi primerjave shem torej upoštevajo statične, strukturne, terminološke informacije shem kakor tudi dinamične vidike shem.

### 2.2.1 Ocenjevanje kakovosti integracije shem

Lindland (1994) razvršča osnovne tehnike za izboljšavo kakovosti shem kot so preiskovanje, sistematični pregled, računalniško podprto vpogledovanje v shemo, razlaganje, simulacijske tehnike. Assenova (1996) predlaga sistematični pristop za izboljšavo kakovosti shem, ki temelji na postopnem preurejanju s transformacijo shem. Transformacija sheme je funkcija, ki kartira konceptualno shemo v drugo, novo shemo. Primeri transformacij shem so transformacije prve normalne forme, tretje normalne forme, Boyce-Coddove forme, četrte normalne forme in projekтивne razgradnje normalnih form (Troyer, 1993).

Kakovost integracije shem izpostavlja vprašanje načina testiranja kakovosti integracije konceptualnih shem. Battini et al. (1986) za ocenjevanje kakovosti integracije shem predpostavlja kriterije popolnosti, minimalnosti in razumljivosti. Popolnost integracije, kompletnost, pomeni, da globalna oziroma integralna shema vsebuje in obsega tudi značilnosti lokalne sheme (koncepte, razmerja med koncepti, lastnosti konceptov, omejitve in odvisnost). Minimalnost integralne sheme pomeni, da mora biti integralna shema minimalna, četudi mora obsegati vse informacije, ki obstajajo v integriranih lokalnih shemah. Ponovljene, redundantne informacije, ki izvirajo iz spajanja shem, morajo biti odstranjene.

Razumljivost integralne sheme je vezana na obširnost ali zgoščenost predstavitve konceptov in razmerij med njimi. Če je integralna shema predstavljena z grafičnim diagramom, je mogoče opredeliti elemente razumljivosti grafičnega diagrama integralne sheme, kot so oblika diagrama, dolžina razmerij, število križanj in podobno. Kakovost integracije in minimalnost integralne

sheme sta odvisni predvsem od integracijskega pristopa (pristop navideznega pogleda, snovnega pogleda, hibridni pristop). Ta mora zagotavljati zanesljiv prenos podatkov o konceptih in razmerjih le-teh med lokalnimi shemami in integralno shemo ter zagotavljati najmanjši zgornji obseg integralne sheme (Buneman et al., 1992). Razumljivost integralne sheme je odvisna predvsem od zahtev, ki veljajo pri izgradnji konceptualnega modela.

### 2.2.2 Organizacija opravil pri integraciji shem

Osnovne faze integracije shem so sledeče: primerjave shem, prilagajanje shem in spajanje shem (Hakkarainen, 1999). Primerjava shem vključuje analiziranje shem; pri analizi shem se išče obstoj razmerij med shemami in ko so ta razmerja odkrita, se odkrivajo podobnosti zaradi določitve skladnosti, torej ustreznosti oziroma neustreznosti različnih predstavitev istih pojavov stvarnega sveta; v postopku primerjave shem odkrita razmerja predstavljajo skladišče oziroma zalogo izmenjevalnih povezav, ki služijo prilagajanju in spajanju shem; v postopku analize shem se določajo formalne skladnostne, integracijske trditve. Prilagajanje shem obsega spreminjanje ene ali več shem, dokler ni doseženo stanje, ko so vsi pojavi iz objektnega sistema predstavljeni na enak način v vseh obravnavanih shemah. V fazo prilagajanja sodi tudi postopek bogatenja shem. Spajanje shem vključuje sestavljanje shem za pridobitev integrirane, globalne sheme. Larson in Navathe (1989) predlagata naslednji vrstni red nalog integracije shem: pretvorba shem – notranje prilagajanje, transformacija shem, ki naj bi bile integrirane, se izvede s preoblikovanjem shem v obliko, primerno za integracijo; podroben opis skladnosti trditve; pri specifikaciji trditve se natančno opišejo trditve, ki se tičejo razmerij skladnosti med koncepti v dveh ali več shemah; spajanje, integracija shem; z integracijo shem se združijo razredi dveh ali več shem na osnovi specifikacije trditve; izdelava diagramov povezav. Izdelajo se diagrami povezav, ki jih je mogoče uporabljati pri integraciji pogledov ali integraciji zbirk podatkov ter pri kartiranju spajanja shem.

### 2.2.3 Postopek integracije shem

Postopek integracije shem se sestoji iz petih iterativnih korakov, ki so: predintegracija posameznih lokalnih shem, primerjava konceptov v posameznih lokalnih shemah, ureditev posameznih lokalnih shem, integracija več lokalnih shem, ureditev integralne sheme. V postopku predintegracije posameznih lokalnih shem oziroma standardizacije sheme je potrebno izvesti preoblikovanje sheme v obliko, primerno za integracijo. To se izvede z upoštevanjem predpisanih pravil. Odkrijejo se neskladja in protislovja znotraj lokalne sheme z upoštevanjem vnaprej oblikovanih kriterijev in zahtev.

Pri koraku primerjave konceptov v posameznih lokalnih shemah se med seboj primerjajo koncepti lokalne sheme zaradi odkritja objektnih parov, ki so semantično podobni, ali pa odkrivanja konceptov, ki so v navzkrižju z lokalno shemo. To je najpomembnejši korak postopka integracije shem, ker se v njem koncepti razpoznavajo glede na pomene, semantiko, pridobljeno iz lokalnih shem kot tudi iz semantičnega slovarja, če ta obstaja. V tem koraku se oblikujejo priporočila za reševanje odkritih neskladij med koncepti lokalne sheme. Pridobivanje semantičnega znanja obsega več različnih pristopov: analiziranje shem, odkrivanje semantičnih podobnosti, odkrivanje strukturnih neskladij, določanje razmerij med lastnostmi primerjanih konceptov.

Posamezne lokalne sheme in neskladja med njimi se poskuša urediti z določenimi integracijski-

mi orodji, če so le-ta na razpolago. Odkrita strukturna in semantična neskladja so združena v skupine, na primer močno enakovredno, skladno, tesno razmerje. Neskladja, ki jih ni mogoče urediti z integracijskimi orodji, se predstavijo uporabnikom, ti pa potem sami sprejmejo odločitve o njihovem reševanju.

Na osnovi lokalnih shem se izdelata skupna, globalna shema. Pri tem se uporabljajo uporabniške trditve in znanje, shranjeno v zbirki podatkov integracijskega znanja (Song in Johannesson, 1992). Globalna shema je običajno predstavljena grafično. Nato se začnejo ustvarjati povezave med sestavinami lokalnih in globalne sheme.

V koraku ureditve globalne sheme se globalna shema preizkusi v odvisnosti od pogojev ponovljivosti in doslednosti. Shema se preoblikuje s postopki odstranjevanja ponavljajočih se lastnosti, razmerij in odvisnosti konceptov. Postopki morajo biti v skladu z vnaprej opredeljenimi preoblikovalnimi pravili kot tudi z odločitvami uporabnikov. Izvede se stvarna povezava med izvirnimi shemami in globalno shemo. Operacija razširitve izvedbe integracije shem se vpelje na primerna mesta v porazdeljenih sistemih zbirk podatkov. Podatkovna zbirka integracijskega znanja je razširjena z vključitvijo novega semantičnega znanja, pridobljenega iz uporabniških trditvev. Vsakič, ko so dodana nova znanja in pravila, se izvede preizkus doslednosti, konsistentnosti.

#### 2.2.4 Primerjava shem

Primerjanje shem zadeva tolmačenje, interpretacijo modelov objektnega sistema (Hakkara-inen, 1999). Modele ustvarijo različni ljudje z različnimi nameni. Zato bi bilo neutemeljeno popolnoma avtomatizirati postopke primerjave shem (Johannesson, 1993). Vendar je mogoče razviti metode in orodja za podporo pri odkrivanju ujemanja med shemami in za preverjanje verodostojnosti predlaganih povezav. Z vidika primerjave shem bodo obravnavane štiri metode: metoda semantičnega poenotenja, metoda mehke terminologije, metoda enakovrednosti atributov in metoda porazdeljenih ontologij.

Vsem metodam primerjanja shem je skupno, da informacije, ki so vsebovane v lokalnih shemah, niso zadovoljive za izdelavo zanesljivega niza povezovalnih trditvev (Hakkara-inen, 1999). Zato se v vse novejših tehnikah primerjanja shem uvajajo metode semantičnega bogatenja lokalnih shem. V vseh obravnavanih metodah primerjave shem je očitna potreba po obravnavi dinamičnih vidikov konceptualnih shem.

#### 2.2.5 Primerjava shem z metodo semantičnega poenotenja

Metoda semantičnega poenotenja shem obsega logično poenotenje, poenotenje strukturiranih objektov in poenotenje semantičnih informacij (Bouzeghoub in Comyn-Wattiau, 1990). Logično in strukturno poenotenje je sintaktičen postopek. Postopek poenotenja ni uspešen, če se združuje dva objekta, ki nista sintaktično enotna, četudi sta semantično ustrezna. Interna semantična predstavitev objekta je izvedena v obliki diagrama. Entiteta je sestavljena iz spiska atributov in omejitev.

Pri integraciji pogledov je semantično poenotenje razširjeno na ujemanje vzorcev entitet, oziroma razmerij. Pogled je zapleten objekt sestavljen iz entitet in razmerij. Tolmačenje sestavin sega od enakovrednosti do nepodobnosti oziroma je lahko še nedorečena. Sklepanje po podobnosti – analogiji, temelji na primerjavi podobnostnih vektorjev. Mehanizmi za primerjavo so sledeči: primerjava izraza z izrazom, dedukcija z dedovanjem, indukcija (generalizacija/specializacija)

preureditev pogleda ter sodelovanje z uporabnikom.

### 2.2.6 Primerjava shem z metodo mehke terminologije

Metoda mehke terminologije je pristop k integraciji podatkovnih shem z uporabo nejasnih znanj o stvarnosti (Fankhauser in Neuhold, 1993, Fankhauser, 1997). Semantična dvoumnost opredelitve imen razredov se rešuje z določanjem največjega razpona drevesa za določeno ime razreda, njegove attribute in razmerja v terminološkem drevesu, omrežju nejasnih znanj. Stopnja podobnosti in najprimernejši način integracije dveh raznovrstnih razredov se določata s primerjavo in poenotenjem dreves obeh razredov. Terminološko drevo je način prikaza semantičnega slovarja.

Združevalno omrežje povezuje izraze z dvema načinoma binarnih razmerij (generalizacija/specializacija in pozitivna povezava). Pozitivna povezava se uporablja za izraze s prekrivajočim se zapisom simbolov ter za redno povezane simbole. Stopnja podobnosti razredov se izraža v numerični obliki na intervalu  $(0, 1)$ . Stopnja podobnosti se izračuna na osnovi razdalje med dvema izrazoma ali pa jo izkustveno določi izvajalec primerjave. Iz površnosti znanja, izražene v omrežju nejasnih znanj izvirata dve vrsti nezaželenih stanj. Modelna dvoumnost se pojavi, kadar niso opredeljeni osnovni gradniki modela, kot so na primer osnovna razmerja. Vendar Fankhauser in Neuhold (1993) trdita, da modelna nejasnost poenostavlja pridobivanje znanja in zahteva manj intelektualnega napora za kartiranje shem v zbirko znanja. Modelna dvoumnost ne zmanjšuje sposobnosti metode za nudenje integracijskih priporočil. Semantična dvoumnost izvira iz dovoljenih večpomenskih, polisemičnih izrazov v zbirki znanja, kjer je en izraz povezan z več pomeni. Večpomenski izrazi se pojavijo celo, če se klasifikacija uporablja dosledno. Reševanje dvopomenski izrazov je mogoče razrešiti z doslednim preverjanjem pomena sheme, z ontologijo.

Za določanje semantike sheme se uporabljajo podatki o atributih in razmerjih. V primeru metode mehke terminologije pomeni določanje največjega razpona drevesa za določen razred semantično bogatenje sheme. Spajanje zgradb dveh shem je lahko izvedeno z generalizacijsko integracijo, in sicer v primeru, da se prekrivajo ekstenzije dveh raznovrstnih razredov. Če se prekrivajo intenzije dveh razredov, se na osnovi skupnih atributov razredov izvede navidezna povezava. Z združitvijo zgradb dreves razponov za razrede se glede na stopnje podobnosti med razredi pridobi informacija o integracijski moči. Glede na informacije o integracijski moči se pripravi predlog integracijskega zaporedja za spajanje shem.

### 2.2.7 Primerjanje shem z metodo enakovrednosti atributov

Pristop primerjave shem z metodo enakovrednosti atributov obravnava primerljivost atributov različnih razredov. Predpostavlja se, da če je mogoče integrirati attribute dveh razredov, potem je mogoče integrirati tudi same entitete. Rezultati analize primerljivosti atributov se uporabijo kot informacije za izvajanje integracije shem. Atributno enakovrednost je glede na enakovrednostne lastnosti mogoče razdeliti v tri stopnje: močno, slabotno in razdeljeno. Močna atributna enakovrednost je razdeljena v dve kategoriji: atributi, ki so močno a enakovredni, in atributi, ki so močno b enakovredni. Obstajajo štiri strategije spajanja glede na atributno enakovrednost (Larson et al., 1989): integracija vseh nerazdeljenih atributov; integracija atributov, ki so a enakovredni; integracija atributov, ki so b enakovredni in kažejo razmerja med nezdruženimi

podobnimi atributi; integracija vseh nerazdeljenih atributov in prenos vrednosti med atributi. Pristop primerjave shem z metodo enakovrednosti atributov opredeljuje enakovrednost med razredi in med nizi razmerij.

### 2.2.8 Primerjanje shem z metodo porazdeljenih informacijskih ontologij

Pristop primerjave shem z metodo porazdeljenih ontologij se uporablja kot računski model podobnosti. Podobnost se izračunava na osnovi konceptualnih opisov (Spanoudakis in Constantopoulos, 1994). Objekti so razdeljeni glede na dva osnovna kriterija: klasifikacijska raven objekta in vloga objekta. V podatkovnem modelu, ki je uporabljen v metodi porazdeljenih ontologij, obstajajo štiri kategorije objektov: entitetni simboli, atributni simboli, entitetni razredi in atributni razredi.

Vsako objektno kategorijo označuje podniz sistemskih identifikatorjev za objekte, razrede, domene, attribute in vrednosti atributov. Računski model preiskuje semantiko štirih abstrakcij konceptualnega modeliranja: identifikacijo, klasifikacijo, generalizacijo in atributiranje. Za ocenjevanje podobnosti obstaja zbirka načel: načelo ontološke enotnosti, načelo normalizacije in načelo delno enotne predstavitve. Načelo ontološke enotnosti omejuje primerjavo dveh objektov le na tiste objekte, ki so entitete ali atributi iste klasifikacijske ravni. Ocene semantičnih razdalj med objekti različnih ontologij niso semantično razločljive. Načelo normalizacije opredeljuje podobnost kot monotono padajočo funkcijo »mere normalizirane semantične razdalje«, ki omogoča prepuščanje sicer podobnih, a po strukturi različnih konceptov.

Načelo delno enotne predstavitve opredeljuje podobnost kot funkcijo, ki združuje delne razdalje, ki so semantične razdalje med objekti glede na štiri že navedene abstrakcije konceptualnega modeliranja. Dva primerjana simbola se upoštevata kot semantično neraznolika, homogena, če sta oba identifikatorja objektov oziroma razreda objektov ter če sta oba atributa objektov, ki so specializacije istega podedovanega atributa, superrazreda. Semantična razdalja identifikacije med identifikatorjema obravnavanih objektov se meri s funkcijo identitete, ki prireja vrednost nič za identične simbole in vrednost ena za neidentične simbole. Semantična razdalja klasifikacije in generalizacije se meri s približno oceno semantične raznolikosti atributov dveh objektov. Različnost razredov iste ravni ali razredi na različnih ravneh (nadrazredi, podrazredi) se obravnavajo kot vir atributnih raznolikosti. Izračuna se utež kot inverzna vrednost globine med razredi v diagramu. Semantična razdalja atributov dveh primerjanih objektov je merjena kot seštevek razdalj med semantično homogenimi in semantično nehomogenimi atributi objektov glede na njihove faktorje posebnosti, ki jih objektom pripiše uporabnik.

Atribut dveh objektov se šteje za skupnega, če je podedovan po istem razredu ali superrazredu. Odkrivanje semantično homogenih atributov temelji na predpostavki, da ljudje pri razlaganju podobnosti, v smislu kartiranja med objekti, težijo k obliki 1 : 1. Atributi, za katere se ne šteje, da so skupni, homogeni ali identični, se obravnavajo kot raznoliki. Pristop primerjave shem z metodo porazdeljenih ontologij ponuja enostavne načine prekrivanja sestavnih delov shem v eno integrirano shemo, ko so zagotovljene informacije o podobnostnih razmerjih med podatkovnimi strukturami v lokalnih shemah.

### 2.2.9 Integracijske trditve in analiza skladnosti shem

Skladnostna, integracijska trditev je objektivizacija razmerja med primerjanima strukturama dveh shem. Integracijska trditev je enakovredno pripisana obema primerjanima strukturama

dveh primerjanih shem. Skladnostno razmerje med dvema strukturama dveh shem ni popolno, ampak delno. To pomeni, da so pogoji za skladnostni tip povezave lahko le delno izpolnjeni, vendar bo ta tip povezave vseeno izbran v integracijsko trditev. Določanje moči oziroma stopnje skladnosti omogoča rangiranje rezultatov, ki jih dobimo pri primerjavi shem. Relativnost skladnostnega razmerja v postopku integracije shem omogoča navajanje elementa negotovosti. Posamezne strukture shem so lahko vključene v več skladnostnih trditvah, v katerih se skladnostni tip, stopnja skladnosti in vir informacij za neko trditev razlikujejo glede na žarišče vidika primerjalne analize. Skladnostna trditev opisuje razmerje med strukturama dveh primerjanih shem in vsebuje štiri sestavine: par struktur primerjanih shem, skladnostni tip, stopnja skladnosti in niz virov trditve. Trditvi sta izrazno podobni, če imata isto poimenovanje oziroma simbole. Trditvi sta konceptualno podobni, če izražata znanje statičnih tipov informacij. Trditvi sta funkcionalno podobni, če izražata znanje dinamičnih tipov informacij.

Za strukturi par primerjanih shem je mogoče zapisati eno ali več skladnostnih trditvah. Vendar se ena skladnostna trditev lahko istočasno uporabi le za opis enega tipa skladnosti za določen par struktur primerjanih shem. Skladnostni tipi trditvah so naslednji: istoveten, enakovreden, podoben, znosen/združljiv, različen. Stopnja skladnosti trditve določa, v kakšni meri je določena skladnostna trditev med strukturama dveh shem veljavna. Upoštevati je treba, da je pri primerjavi istega para struktur dveh primerjanih shem mogoče odkriti več skladnostnih trditvah. Stopnja skladnosti je v tem primeru prirojalec povezave na enega od splošnih skladnostnih tipov. Stopnja skladnosti služi spremembi skladnostnega tipa trditve, ki je pripisana razmerju med dvema strukturama sheme.

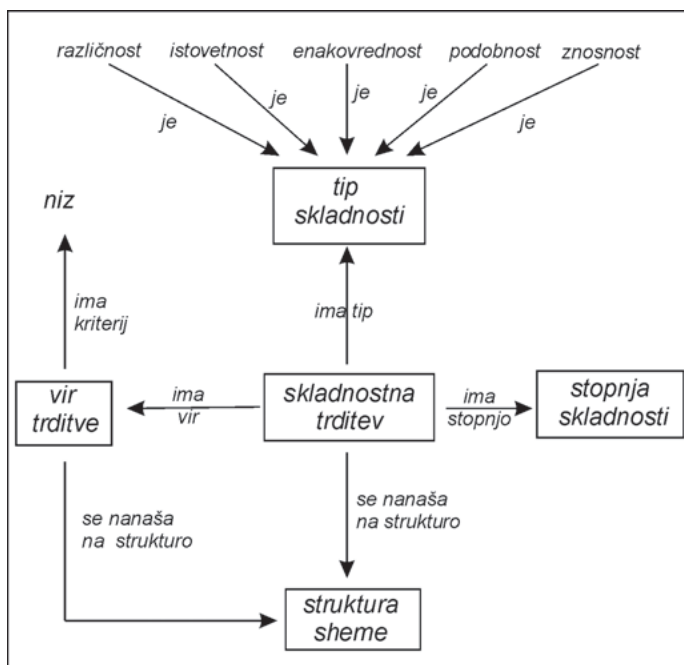
Vir trditve označuje strukturo tretje sheme, ki je vključena v pogojno vodenje odločanja o skladnostnem tipu trditvah, in to v primeru, da skladnost obstaja. Razlog za navajanje vira trditve je v zagotavljanju pojasnila, glede vzrokov za predlaganje določenih trditvah. Skladnostna trditev vsebuje sklicevanje na vir trditve tudi zato, da se omogoči opis in klasifikacija metod za vpletanje metod semantičnega bogatenja.

Opredelitve za skladnostne tipe, torej istovetnost, enakovrednost, podobnost, združljivost, navajamo v nadaljevanju (Hakkarainen, 1999). Dve strukturi sheme sta istovetni, če imata enako sintakso, ob tem da imata enotno predstavitev in isto ekstenzijo. Dve strukturi sta enakovredni, če imata isti pomen, ob tem da imata enoten učinek ali aplikacijo in isto intenzijo. Dve strukturi sta podobni, če sta istovetni ali enakovredni ter klasificirani kot izraz, koncept, vedenje oziroma funkcionalna podobnost, odvisno od tega, kateri tip informacij opredeljujeta. Dve podobni strukturi shem sta izrazno podobni, če imata isto poimenovanje oziroma simbole, sta konceptualno podobni, če izražata znanje statičnih tipov informacij, sta funkcionalno podobni, če izražata znanje dinamičnih tipov informacij.

Dve strukturi shem imata lahko dva tipa podobnosti, in sicer izrazno in funkcionalno. Izrazno sta podobni celo, če nimata enake sintaktične oblike. Strukture shem imajo lahko poleg istovetnosti ali enakovrednosti tudi konceptualno podobnost. Dve strukturi shem sta združljivi (tip skladnosti), če sta: obe povezani/podobni s strukturo tretje sheme in klasificirani kot abstraktni ali opisni tip povezave s strukturo tretje sheme. Pri tem sta dve strukturi shem abstraktno združljivi, če sta obe strukturi podobni strukturam tretje sheme, ki izraža znanje o tipu domene, in opisno združljivi, če sta obe strukturi shem opisani s strukturo tretje sheme, ki izraža znanje o tipu domene.

Pri tem tipu skladnosti je pozornost osredotočena na skupno strukturo tretje sheme. Oblika, pomen in tip informacij struktur dveh shem niso obravnavani. Predpostavlja se, da obravnavane sheme izražajo znanje istega tipa informacij. Pri tipu skladnosti »združljivost« je mogoče opredeliti lastnosti združevalnih razmerij znotraj posamezne sheme, kot so »vsebovan-v«, »je-del-od« ali »to-je-razmerje«.

Namen opredelitve je uporaba bogatitvene strukture za odkrivanje lastnosti združevalnih razmerij znotraj posamezne sheme kakor tudi razmerja med strukturami shem, kot so istovetnost, enakovrednost in podobnost.



Slika 4: Metamodel za skladnostne trditve (povzeto po Hakkarainen, 1999)

## 2.2.10 Analiza skladnosti shem in njihovo spajanje

Analizo skladnosti shem je mogoče prikazati v omrežju odvisnosti elementov v skladnostni analizi. Omrežje omogoča natančno predstavitev odvisnosti med tipi skladnosti oziroma opisi vrst znanja, ki se pojavljajo v različnih postopkih primerjave shem. Omrežje navaja postopke odkrivanja skladnosti, analize nesoglasij in semantične bogatitve shem v okviru faze primerjave shem pri njihovi integraciji (slika 4).

Spajanje shem je problem povezovanja več shem zbirk podatkov v enoten pogled oziroma integralno shemo. Ta problem nastopa med izgradnjo novih, obsežnih zbirk podatkov ali pa pri zagotavljanju integriranega pristopa do že obstoječih zbirk podatkov. Del postopkov podatkovne integracije je sestavljanje sistema, ki kartira pomene, semantiko podatkov različnih shem in omogoča pregled nad spajanjem shem. Pri posredovalnem sistemu in pri sistemu podatkovnega skladišča obstaja enotna shema, ki zagotavlja enoten, integriran pogled na podatke iz različnih

virov. V podatkovnih sistemih, ki imajo n različnih shem podatkovnih virov, je potrebno ustvariti in izvesti najmanj n različnih kartiranj shem. Pri obravnavi podatkovnega skladišča je kartiranje shem enosmerno. Kartiranje mora predstaviti le način za predstavitev informacij podatkovnega vira, in to na tak način, da je v skladu z integriranim pogledom enotne sheme (Kosky, 1994). Posredovalni sistem kot možni sistem za pristop navideznega pogleda integracije zbirke podatkov mora zagotavljati dvosmerno kartiranje shem. Prva smer je taka kot pri kartiranju shem za podatkovno skladišče, in sicer je to predstavitev informacij podatkovnega vira v skladu z integralno, posredovalno shemo. Vendar pri posredovalnem sistemu nastopajo povpraševanja v skladu z integralno, posredovalno shemo, ki jo je potrebno prevesti v obliko, ki je razumljiva za shemo podatkovnega vira.

Pregled različnih pristopov za združevanje shem navaja Battini et al. (1986). Pristopi imajo razpon od orodij za interaktivno upravljanje spajanja shem pa do samodejnih algoritmov spajanja shem. Za splošno rabo je najprimernejši sestavljeni pristop, ki omogoča določene uporabniške posege za reševanje nesoglasij med shemami in dodatno omogoča samodejno spajanje shem, kar je še posebej pomembno pri obsežnih shemah. Spajanje zbirke shem je opredeljeno z naročanjem informacij iz shem do najmanjšega zgornjega obsega spojitve, kar je komutativna in asociativna matematična operacija. To pomeni, da je spajanje neodvisno od vrstnega reda, v katerem se sheme spajajo. Tako je vseeno, ali se sheme spojitijo naenkrat ali v iterativnem postopku, rezultat spajanja je vedno enak. Pomen jasnih matematičnih lastnosti operacije spajanja shem je večplasten.

Semantika podatkovnih shem je lahko skrajno zapletena, zato obstaja velika verjetnost, da se bodo interaktivno uvedene trditve izkazale kot napačne, kar ima lahko nezaželene posledice. Iskanje pravih trditve je kočljiva naloga. Zato je željeno, da model spajanja omogoča operacijo razveljavljanje trditve, ki je odvisna le od trditve samih, ne pa od zaporedja njihove predstavitve. Pomanjkanje asociativnosti modela spajanja bi v navedenem primeru pomenilo ročno razveljavljanje vseh sprememb, uvedenih z napačno trditvijo. Drugi vidik pomembnosti jasnih matematičnih lastnosti operacije spajanja shem je enostavnost trditve, ki povezuje sheme. Spojitveni postopek je mogoče opisati z enostavnim potekom: »predstavi obravnavane sheme vključno z uporabniškimi trditvami in vse skupaj združi«. Za spojitveni postopek je potrebno uporabiti formalni model, ki pojasnjuje semantiko spojitvenega postopka. Model mora biti tako splošen, da je z njim mogoče predstavljati sheme različnih podatkovnih modelov.

Pri modeliranju shem zbirke podatkov za namene kartiranja spajanja shem gre za modeliranje razmerij višjega reda, to je razmerij med razmerji shem. Pri tovrstnem modeliranju se skuša zadržati število različnih simbolov čim manjše zato, da bi bil model čim bolj enostaven in enoten. Model vključuje razmerje specializacije to-je (is-a) in predstavitve cikličnih opredelitev entitet in razmerij, ki se pojavljajo v nekaterih objektno usmerjenih podatkovnih sistemih (Kosky, 1994). Za formalne predstavitve shem zbirke podatkov in njihovih sestavnih delov se v modelu uporabljajo usmerjeni grafi. Razredi, ki so začetne oziroma končne točke usmerjenih grafov, odgovarjajo razmerjem, nizom entitet in osnovnim tipom entitet. Puščice odgovarjajo atributom entitet in vlogam entitet in razmerij. Koncepti razmerje/entitete/osnovni tipi ter koncepti atributi/vloge so poenoteni v dva koncepta: razredi in puščične oznake.

### 2.2.11 Reševanje problemov spajanja shem

Prvi problem pri oblikovanju integralne podatkovne sheme je določanje odgovarjajočih povezav med razredi različnih zbirke podatkov, ki se obravnavajo. Te povezave mora zagotoviti oblikovalec



integriranega sistema. Ker so odvisne od razlag stvarnega sveta, morajo biti obravnavane za vsak primer posebej. Oblikovalec integriranega sistema mora sam razjasniti imenovalna nasprotja, kot so homonimi in sinonimi. Ekberg (1995) je dokazal, da za splošni primer ni algoritma za odločanje o tem, ali sta dve shemi lahko smiselno povezani ali ne.

Prevlada sheme in enakovrednost shem nista absolutni razmerji med shemami. Ekberg (1995) opredeli prevlado ene sheme nad drugo kot relativno razmerje glede na niz integracijskih trditev, ki zagotavljajo intuitivna razmerja med shemami. Šibka prevlada med shemami je prehajajoča, tranzitivna: dve shemi sta enakovredni, če druga drugo šibko prevladujeta. V splošnem se spajanje shem izvaja kot unija shem. Integralna shema šibko dominira posamezno izvirno shemo, če je izpolnjen pogoj odsotnosti nesoglasij.

Cilj spajanja shem je poiskati integralno shemo, ki predstavlja vse informacije obravnavanih shem, vendar naj integralna shema ne predstavlja dodatnih informacij. Integralna shema naj bi imela najmanjši zgornji obseg, kar predstavlja problem. Ob trditvi, da ena shema predstavlja več informacij kot druga – prevlada sheme, se predpostavlja tudi, da je lahko vsaka instanca prve sheme tudi instanca druge. Tako mora vsaka instanca zbirke podatkov vsebovati najmanj vse potrebne podatke, da bi lahko bila instanca druge sheme. Za iskanje rešitev za spojitve natanko dveh podatkovnih shem se uporablja pojem iskanja binarnega operatorja, ki je asociativen in komunikativen ter omogoča binarni spoj shem, opredeljen kot najmanjši zgornji obseg združitve. Binarna spojitve, bisekcija, predstavlja enakovredno razmerje med shemama (Assenova, 1999). Spajanje dveh shem lahko poleg običajnih razredov shem vsebuje posebne implicitne razrede, ki se spajajo. Uvajanje takšnih razredov predstavlja problem, ker jih je potrebno obravnavati drugače kot običajne. Če se obravnavajo kot običajni razredi, se ugotovi, da binarni spoj ni več asociativen, neodvisen od vrstnega reda združevanja. Spajanje take integralne sheme s tretjo shemo bi lahko vsebovalo dodatne implicitne razrede, ki bi lahko bili specializacije razredov prve sheme. Vrstni red shem v postopku združevanja bi s tem postal pomemben, kar je nezaželeno. Dodatni problem pri domnevnem naročanju informacij o shemah je možnost, da ena od shem predstavlja več informacij kot druga, ob tem, da ima manj implicitnih razredov. Če se želi z eno shemo predstaviti vse informacije druge sheme in še dodatne informacije, je očitno potrebno, da prva shema vsebuje najmanj vse razrede druge sheme. Zaradi naštetih razlogov se je treba izogniti zapletenosti uvajanja implicitnih razredov. Razredi postanejo resnično implicitni. V ta namen se popusti strogost opredelitve podatkovne sheme, tako da postanejo implicitni razredi nepotrebni. Tako opredeljene podatkovne sheme se imenujejo ohlapne sheme.

Dokazano je, da najmanjši zgornji obseg zbirke ohlapnih shem vedno obstaja, kadar specializacijska razmerja shem zadostijo določenim združitvenim omejitvam. Takšna zbirka ohlapnih shem je opredeljena kot združljiva zbirka shem. Obstaja način izražanja ohlapne integralne sheme, ki vsebuje vse informacije, vsebovane v katerikoli združljivi zbirki obravnavanih shem, vendar ne predstavlja nobenih dodatnih informacij. Tak najmanjši zgornji obseg zbirke ohlapnih shem se imenuje spoj ohlapnih shem, tako dobljena integralna shema pa se imenuje ohlapna integralna shema. Ko je iskani spoj ohlapnih shem odkrit, je v ohlapno integralno shemo potrebno uvesti neizogibne implicitne razrede. S povratnim postopkom iskanja razredov se preuredi ohlapna integralna shema v pravo integralno shemo.

### 2.3 INTEGRACIJA ZNANJ V ZBIRKAH PROSTORSKIH PODATKOV

Semantična integracija geografskih podatkovnih nizov je postopek vzpostavljanja razmerij med odgovarjajočimi primeri objektov za različne, neodvisno izdelane geografske podatkovne nize določenega območja (Uitermark et al., 1993, Uitermark, 1996, 2001). Semantični vidik geografskih podatkovnih nizov je zelo odvisen od namena aplikacije, za katero so bili izdelani. Poleg zbirk podatkov, ki so namenjene porazdeljeni uporabi med različnimi uporabniki, obstajajo tudi podatkovni nizi, ki so v osnovi enonamensko uporabni in popolnoma aplikativno zasnovani. Isti geografski pojav/objekt je lahko različno poimenovan v različnih izvedbenih področjih. Bruegger (1996) je razvil prostorsko teorijo za integracijo mozaičnih prostorskih podatkov, rastrov. V razviti prostorski teoriji geografskih informacijskih sistemov navaja tri sestavine: predstavitve enostavnih in zapletenih atributov, meta podatke in transformacije prostorskih podatkov, podrobnih v manj podrobne ločljivosti, ki jih primerja z drugimi generalizacijskimi postopki. Za razvoj teorije je uporabil Sintonovo (1979) taksonomijo prostorskega znanja. Sintonov model stvarnosti geografskega prostora opredeljuje tri sestavine prostorskih podatkov: položaj, tematiko in čas. Sinton je predpostavil, da je mogoče modele predstavitve geografskega prostora razviti tako, da se učvrsti ena sestavina, nadzira druga in meri tretja. Bruegger je v svoji raziskavi predpostavil čas kot učvrščen in preizkusil dve metodi za integracijo modelov geografskega prostora, in sicer:

- nadzor teme, meritve položaja in
- nadzor položaja in meritve teme.

Namen semantične integracije geografskih podatkovnih nizov je neovirano porazdeljevanje informacij med različnimi viri geografskih informacij. Še posebej je pomembno preučevanje integracije geografskih podatkovnih nizov v smislu razširjanja obnovljenih podatkovnih nizov (update propagation), to je večkratne uporabe obnovljenih podatkov v več različnih zbirkah podatkov (Wijngaarden et al., 1997; Kim, 1999). Zajem tradicionalnih analognih načrtov in kart se v državah Evropske unije zaključuje, vse pomembnejše pa postaja vzdrževanje zajetih podatkovnih nizov ter njihova integracija. V analognih načrtih in kartah so objekti povezani samo posredno, s skupnim referenčnim sistemom, to je z državnim koordinatnim sistemom. Prostorska integracija podatkov temelji na zanesljivem geometričnem referenčnem sistemu in na razumni ravni ujemanja podrobnosti oziroma na ločljivosti različnih podatkovnih nizov. To pomeni, da mora biti za izdelavo vseh različnih zbirk podatkov, ki so vključene v integracijski postopek, uporabljen isti koordinatni sistem in vse integrirane zbirke podatkov morajo imeti primerljivo raven ujemanja podrobnosti vsebin, tematik podatkovnih nizov ter ločljivosti – če gre za mozaične podatkovne strukture.

Semantika zbirke prostorskih podatkov vključuje namenski vidik zbirke podatkov, strukturno informacijsko ontologijo kot tudi opredelitev ontologije geografskih struktur in procesov – prostorsko ontologijo. Tudi opis informacijske strukture, kodiranja podatkov na fizični objekt, pojasnjuje pomen podatkov (Nyegres, 1992). To pomeni, da je mogoče povečati semantičnost geografske zbirke podatkov, če je vanjo vključenih več relacij med podatkovnimi strukturami pojavov, ki so predstavljeni v zbirki podatkov. Da bi te posledične povezave med podatkovnimi objekti bilo mogoče nadgraditi v vnaprej določene povezave, je bilo razvitih več strategij. V domeni računalništva je najpomembnejša metodologija za integracijo zbirk podatkov integracija shem (Spaccapietra, 1992). Ta metodologija je bila razširjena na domeno geografskih podatkovnih

nizov (Devogle, 1998).

V geoznanostih je prišlo do privzemanja metod teorije komunikologije, na primer metode primerjanja razmerij (Sester, 1998), ter do privzemanja ontologij s področij umetne inteligence (Vet in Mars, 1998). Izdelava in uporaba ontologij za geografske podatkovne nize omogoča preverjanje neskladnosti rezultatov integracije za geografske podatkovne nize. Pri integraciji geografskih podatkovnih nizov za namene porazdeljevanja obnovljenih podatkovnih nizov ima velik pomen njihova časovna usklajenost, kar je mogoče upoštevati z uporabo časovnih atributov podatkovnih nizov.

Pri preučevanju pomenkega povezovanja zbirk podatkov je potrebno upoštevati različne znanstvene discipline: nauk o bivajočem oziroma ontologijo kot del filozofije, pomenoslovje, logiko in konceptualno modeliranje. Pri specializiranem preučevanju prostora in prostorskih pojavov pa tudi kognitivno kartiranje in prostorsko modeliranje.

### 2.3.1 Odkrivanje znanja v zbirkah prostorskih podatkov

Število zbirk prostorskih podatkov, novi viri informacijske, merske in mobilne tehnologije in njihov obseg sčasoma naraščajo, posebej na področjih, kot so geomarketing, nadzor prometa in okoljske študije. Rast zbirk prostorskih podatkov presega človeške sposobnosti analiziranja, zato postaja pomembnejše samodejno odkrivanje znanja, ki se imenuje odkrivanje znanja v zbirkah podatkov (Knowledge discovery in databases – KDD). Odkrivanje znanja v zbirkah podatkov je zapleten postopek odkrivanja neobičajnih, a veljavnih vzorcev v podatkih, ki imajo uporabne možnosti (Fayyad et al., 1996).

Metoda za odkrivanje znanja v zbirkah prostorskih podatkov oziroma rudarjenje prostorskih podatkov (Spatial Data Mining – SDM) je predmet številnih raziskav (Lu, 1993, Keim, 1994, Knorr, 1997, Easter et al., 1997, Leung, 2000). Rudarjenje prostorskih podatkov je odkrivanje uporabnih vzorcev in razmerij med podatki v zbirkah prostorskih podatkov. Ta pristop omogoča obravnavo dinamičnih prostorskih pojavov. Za rudarjenje prostorskih podatkov je treba ustvariti konceptualno omrežje za sistematično preiskovanje struktur v obsežnih, raznolikih in dinamičnih zbirkah prostorskih podatkov.

Najprej se izvede površni pregled zbirke podatkov zaradi odkrivanja verodostojnih vzorcev v smislu trendov, serij, linearnosti in nelinearnosti, čemur sledi podrobno preiskovanje topoloških, urejenih in vzročno povezanih struktur. Zaradi nelinearnosti v vzorcih prostorskih podatkov se za obravnavo zapletenih in dinamičnih prostorskih pojavov uporabljajo umetne nevronske mreže in razvojne modele prostorskih struktur in dogajanj (Openshaw, 1997, Fischer in Leung, 1998, Leung, 1997).

Večina algoritmov za rudarjenje prostorskih podatkov temelji na preučevanju razmerij sosedstva, ker imajo lahko nekateri atributi sosednjih objektov vpliv na obravnavane objekte. Metode za odkrivanje prostorskega znanja iz podatkov so tesno povezane z osnovnimi značilnostmi sistemov za upravljanje zbirk prostorskih podatkov (SDBMS), zato obstaja težnja po integraciji prostorskih algoritmov SDM s sistemi SDBMS (Ester et al., 1997).

### 2.3.2 Grafi in poti sosedstva prostorskih razmerij

Koncept graf sosedstva je določna predstavitev vpletenih sosedskih razmerij, ki so pomembna pri odkrivanju znanja in strojnega učenja. Sosedski graf prostorskega razmerja je opredeljen z vozlišči, ki predstavljajo objekte v zbirkah podatkov, in s povezavami med njimi. Med dvema vozliščema se lahko izvede povezava, če je razmerje sosedstva med njima veljavno. Veljavna sosedska razmerja so lahko sledečih tipov: topološko razmerje (stikanje, prekrivanje, prekrit\_z, vsebuje, znotraj, enak), metrično razmerje (razdalja  $< d$ ), smerno razmerje (določena smer: sever, jug, vzhod, zahod ali azimut). Na osnovi sosedskega grafa je opredeljen koncept sosedska pot, in sicer kot spisek zaporednih vozlišč tega grafa. Na novo opredeljene operacije nad sosedskimi grafi in potmi za namene odkrivanja znanja v zbirkah prostorskih podatkov so sledeče: vzemi\_nGraf (db, rel), vzemi\_sosedstvo (graf, o, pred), oblikuj\_nPoti (objekt, graf, pred, i), razširi (niz\_poti, graf, pred, i), (Ester et al., 1997).

Odkrivanje razmerij ujemanja oblik meja (Boundary Shape Matching – BSM) prostorskih objektov ali skupin objektov sosedstva je ena od standardnih nalog podatkovnega rudarjenja v zbirkah prostorskih podatkov. Odkrivati je mogoče razmerja ujemanja oblik meja med skupinami točk, linij in poligonskih predstavitev pojavov prostorskih objektov v birki podatkov. Pred izvedbo odkrivanja razmerij med oblikami meja je treba razrešiti nekatera vprašanja: »Kaj je meja oziroma oblika meje skupine točk v prostoru geografskih razsežnosti? Kaj je veljavno ujemanje odkritih meja, ki so v sosedskem razmerju? Kako je mogoče določiti stopnjo veljavnosti ujemanja oblike meja?« Odkrivanje veljavnih vzorcev ujemanja oblik meja ima pomen v različnih aplikacijah, kot so tržno vrednotenje posesti, kmetijstvo, gozdarstvo, prostorsko planiranje in prostorska demografija.

# 3

---

## KONCEPTUALNO MODELIRANJE

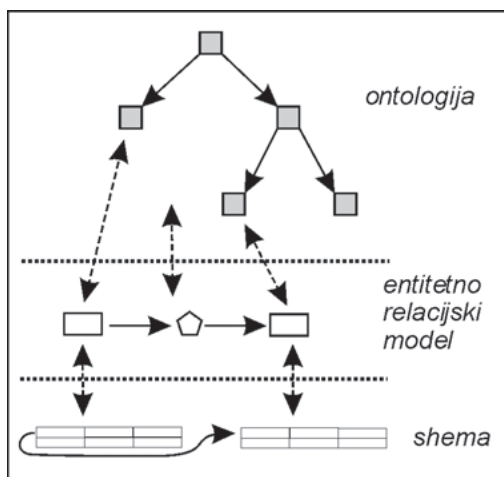
Člani različnih kultur živijo v kognitivno različnih svetovih. Različnost človeških kognitivnih sistemov povzroča različnost ontologij (Simons, 1987). Različno mišljenje je prilagojeno, določene razlike človeških kognitivnih sistemov pa so posplošene tako, da je mišljenje splošno veljavno. Splošno, zdravorazumsko mišljenje torej obstaja neodvisno od različnosti človeških kognitivnih aktivnosti (Smith, 1992, 1993). Dva konceptualna sistema, ki izvirata iz dveh kognitivno različnih svetov, sta ontološko primerljiva, če je vsakega od njiju mogoče obravnavati kot sistem miselnih zasnov objektivnega sveta, ki obstaja neodvisen od človeških konceptualizacij in zasnov. Lakoff (1987) navaja pet različnih načinov primerljivosti dveh konceptualnih sistemov: prevedljivost iz enega v drug sistem, razumljivost – posameznik lahko razume obe alternativni, splošnost uporabe – uporaba istih konceptov v obeh sistemih, usklajenost okvirjenja – situacije v obeh sistemih so okvirjene in obstaja usklajenost med okvirji sistemov, organiziranost – v obeh sistemih se pojavljajo isti koncepti, ki so organizirani na enak način.

Za razumevanje in analiziranje zapletenih pojavov obstaja učinkovit pristop, kot je izdelava modela. Model je v primerjavi s stvarnostjo enostavnejša struktura oziroma mehanizem, s katerim je mogoče opisati določene dele stvarnosti. Primeri modela so karte prostora, nizi diferencialnih enačb, miniaturne makete in modeli informacijskih sistemov. Namen modela je zagotovitev okolja, v katerem se je mogoče učiti dejstev o stvarnosti z analiziranjem modela namesto z neposrednim opazovanjem stvarnosti. Na primer: razdaljo med dvema izbranima točkama v geografskem prostoru je mogoče namesto z merjenjem dolžine neposredno po zemeljskem površju določiti z uporabo karte. Model je vedno lažje proučevati kot stvarnost, ker vsebuje le določene vidike zapletenih pojavov stvarnosti. Na primer: s karto sveta se je mogoče osredotočiti na politični, na topografski ali na kakšen drug vidik sveta.

Pri ustvarjanju modela je pomembno razjasniti, kateri del stvarnosti bo z njim predstavljen. Temu delu stvarnosti se reče tudi nominalna osnova ali področje obravnave (Universe of Discourse – UOD). V nadaljevanju bodo obravnavane miselne zasnove modeliranja, ki so uporabne pri ustvarjanju modelov za informacijske sisteme, še posebej prostorske.

Koncept podatkovnega modela se v splošnem razume kot splošen opis in podroben opis strukture podatkov. Entiteta je stvarni pojav, ki ni naprej deljiv na pojave istega tipa. Modelna entiteta je osnovni nosilec podatkov. V podatkovnem modelu predstavlja entiteta abstraktni koncept stvarnega pojava z ozirom na njeno namembnost (Šumrada, 1993); (slika 5).

Podatkovni model predstavlja percepcijo in miselno razlago stvarnega sveta z ozirom na neko določeno uporabo. Podatkovni model zagotavlja mehanizem abstrakcije, ki omogoča skrivanje strukture strojnega shranjevanja podatkov. Uporaba formalnih metod pri podatkovnem modeliranju omogoča uporabo konceptov, katerih pomen je formalno opredeljen. Matematična



Slika 5: Razmerje med ontologijo, entitetno-relacijskim modelom in shemami (privzeto po Zhan et al. 2000)

formulacija omogoča, da so vsebine predmeta obravnave mehanizirane v strojih, računalniških sistemih (Finkelstein, 1982). Objektivno usmerjeno podatkovno modeliranje je tehnika modeliranja podatkov, s katero je mogoče v zbirke podatkov vključiti več semantičnih podatkov kot v tradicionalne, relacijske podatkovne modele (Zdonik in Maier, 1990). Eden od formalno opredeljenih semantičnih podatkovnih modelov je IFO (Abiteboul in Hull, 1987), ki poudarja strukturo objektov in razmerja med njimi. Model IFO omogoča opis stvarnosti s konceptualnim grafom (Conceptual Graph - CG) ter zagotavlja mehanizem za opazovanje učinkov vzdrževanja podatkov.

V osemdesetih letih dvajsetega stoletja so bili razviti semantični podatkovni modeli, ki so imeli v primerjavi s tedaj prevladujočim relacijskim podatkovnim modelom nove ter povečane podatkovne pomene (Šumrada, 1993). Semantični podatkovni modeli so bili primarno razviti kot orodja za oblikovanje shem. Shema naj bi bila najprej ustvarjena z uporabo semantičnega modela visoke ravni, nato pa za namene udejanjanja prevedena v enega od tradicionalnih podatkovnih modelov. Prednost uporabe semantičnega podatkovnega modela je ta, da poenostavi integracijo vidikov različnih uporabnikov modela in ustvari skladen pogled na podatke in informacije v informacijskem sistemu (Ekenberg, 1995).

Semantični podatkovni model temelji na objektih in funkcijah. Funkcije povezujejo objekte med seboj in jim dodeljujejo vrednosti. Relacije med objekti, atributi objektov in objektom pripadajoči postopki so vsi ponazorjeni s funkcijami objektov (Šumrada, 1993). Relacije med podatki semantičnega podatkovnega modela so v zbirki podatkov predstavljene eksplicitno, kar ne velja za relacijski podatkovni model (Hull in King, 1987).

### 3.1 KLASIFIKACIJA IN ABSTRAKCIJA V KONCEPTUALNEM MODELIRANJU

Ljudje skušajo organizirati sestavo svojih zaznav in dojemanja celo na področjih, kjer stvarnost ne more biti naravno razdeljena na ločene objekte in razrede. Zato je temeljno filozofsko vprašanje, kako so koncepti in razredi sestavljeni. Kljub temu dejstvu ljudje zaradi obstoja splošnega, zdravorazumskega mišljenja klasificirajo z upoštevanja vredno lahko. Izkaže se, da so razredi oblikovani z osredotočenjem na določen vidik obravnave, vsi ostali vidiki pa so neupoštevani. V

znanstvenih klasifikacijskih shemah so meje razredov natančno določene. Nejasna in skrivnostna stvarnost se strukturira z uporabo nasilnega prilagajanja konceptom, kot so objekt, lastnost in razred. Različni jeziki na isto stvarnost vsiljujejo različne strukture. Jeziki torej močno vplivajo na človekovo zaznavanje stvarnosti. Različnost struktur med jeziki pri izbiri jezika za konceptualno modeliranje odpira vprašanja natančnosti in ekonomike.

### 3.1.1 Hierarhije abstrakcij

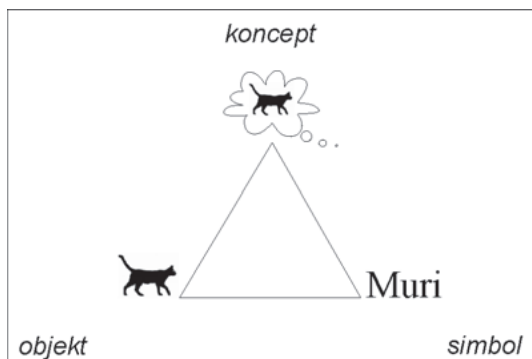
Za opisovanje niza konceptov in razmerij med njimi je prikladna uporaba hierarhije. Hierarhija abstrakcij je koncept sadnega drevesa z vcepljenimi različnimi sortami sadja (različne veje) na isto debelno podlago. Vozlišča vej drevesa so posamezni koncepti in veje opisujejo posploševalna, generalizacijska razmerja med njimi. Generalizacija je splošen koncept, ki nastane kot rezultat iskanja skupnih lastnosti velikega števila konceptov (na primer »sadna rast«). Specializacija pa je bolj svojevrsten koncept, ki ima nekatere posebne lastnosti (na primer plodovi jabolane), različne od lastnosti splošnega koncepta. Vse lastnosti splošnega koncepta generalizacije je mogoče pripisati tudi svojevrstnemu konceptu, specializaciji. Obratna operacija ni veljavna (Boman et al., 1997).

Pri konceptualnem modeliranju je vedno mogoče uporabiti različne koncepte in klasifikacije za opis istega dela stvarnosti. Zato bi bilo zelo uporabno, če bi obstajal splošni kriterij za presojo kakovosti klasifikacije. Za hierarhijo abstrakcij obstaja vsaj en splošno sprejet kriterij kakovosti: svojevrstni koncept ustreznega splošnega koncepta mora biti sestavni del splošnega koncepta torej mora vsebovati del lastnosti splošnega koncepta. Pri praktičnem izvajanju konceptualnega modeliranja obstaja problem, da na želeni ravni abstrakcije naravni jeziki ne zagotavljajo ustreznih besed za koncepte.

### 3.1.2 Opredeljevanje simbolov

Za učinkovito komunikacijo pomena simbolov se uporabljajo opredelitve. Opredelitev je govorni ali kako drugače izvedeni postopek, v katerem se posamezniki med seboj obvestijo o pomenu obravnavanega simbola. Opredelitev navaja, na kateri koncept – miselno zasnovo – se nanaša simbol (znamenje, termin, alfanumerični znak, grafični znak). Opredelitev vedno opredeljuje simbol. Opredelitev nikoli ne opredeljuje koncepta (Devitt in Sterelny, 1987). Opredelitve se v odvisnosti od namena delijo v dve vrsti: opisne in dogovorne.

Slika 6: Richards-Ogdenov semantični trikotnik, ki predstavlja posredno razmerje med simbolom in stvarnim objektom (povzeto po Sowa, 2000b)



Namen opisnih opredelitev je pojasnilo načina, na katerega določeni ljudje uporabljajo določen simbol. Opisna opredelitev je oblika zgodovine, saj se dejansko nanaša na preteklost. Pojasnjuje namreč, kaj so določeni ljudje z določenim simbolom mislili v določenem času in določenem prostoru (slika 6). Opisna opredelitev vedno vključuje tri osebe. Prva oseba je opredeljevalec, ki pojasnjuje pomen simbola. Druga oseba je prejemnik, za katerega je pomen simbola opredeljen. Tretjo osebo predstavlja več tistih ljudi, ki so simbol uporabljali in mu z uporabo dali pomen. Tretja oseba je nujno potrebna, ker je pomen simbola to, kar pomeni ljudem. Če se vprašuje po opisni opredelitvi, je neustrezno vprašanje: »Kaj ta simbol pomeni?« Vprašati je treba: »Za kaj so tisti ljudje uporabljali ta simbol?« Za opisno opredeljevanje simbolov se uporablja opisna logika (Description Logic – DL).

Dogovorne opredelitve imajo namen vzpostaviti, uzakoniti pomen simbola. Ljudje uporabljajo dogovorne opredelitve, kadar želijo uvesti v uporabo nov simbol, ki se ni uporabljal v preteklosti, ali kadar želijo začasno spremeniti pomen uveljavljenega simbola. Kadar nekdo izvede dogovorno opredelitev, potem ta sam izbere pomen simbola in ne zabeleži pretekle uporabe simbola. Dogovorna opredelitev vključuje le dve osebi: opredeljevalca, ki pojasnjuje pomen simbola, ter prejemnika, za katerega je pomen simbola opredeljen. Druge osebe niso vpletene v dogovorno opredeljevanje. Če se vprašuje po dogovorni opredelitvi, je ustrezno vprašanje: »Kaj ta simbol pomeni?« ali »Kaj si določil za pomen tega simbola?« Neustrezno vprašanje bi bilo: »Za kaj ti ljudje ta simbol uporabljajo?« (Boman et al., 1997)

Opredelitve je mogoče razvrstiti tudi glede na metodo opredeljevanja. Z njo je mišljeno sredstvo, s katerim se doseže namen. Enostavna metoda opredeljevanja je metoda sopomenk, sinonimov. Pri tej se učencu ponudi sopomenko, katere pomen mu je domač. Tako se učencu pove, da simbol, katerega pomena se uči, pomeni isto, kot njemu že poznani simbol.

Druga metoda opredeljevanja simbolov je metoda analize, poznana kot Aristotelova metoda opredeljevanja rodov in razlik. Pri tej učitelj imenuje simbol višjega razreda, generalizacijo, v katero spada opredeljevani objekt. Nato poda še nekaj, kar opredeljevani objekt ločuje od preostanka višjega razreda, ter tako opredeli simbol za opredeljevani objekt. Torej se stvar, ki je mišljena s simbolom, pojasni s predstavitvijo analize te stvari.

Metoda sinteze opredeli simbol z navajanjem razmerij med stvarjo, ki se opredeljuje, in drugimi stvarmi. Sistem razmerij je združen, sintetiziran v celoto. Medtem ko analitična metoda pri opredeljevanju simbola pokaže na stvar, ki je mišljena s simbolom, kot na celoto, sestavljeno iz delov, pa metoda sinteze nasprotno pokaže stvar kot del celote.

Vse opisane metode imajo skupno lastnost, da opredeljujejo besede z drugimi besedami. Pri teh metodah se predpostavlja, da učenec že pozna pomen nekaterih drugih besed. Če temu ni tako, so navedene metode neuporabne. V takšnem primeru se uporablja razkazovalna metoda, pri kateri se učencu pokaže stvarni primer objekta, obenem pa se izgovori simbol, ki se opredeljuje.

### 3.1.3 Enostavnost in zapletenost konceptualnih modelov

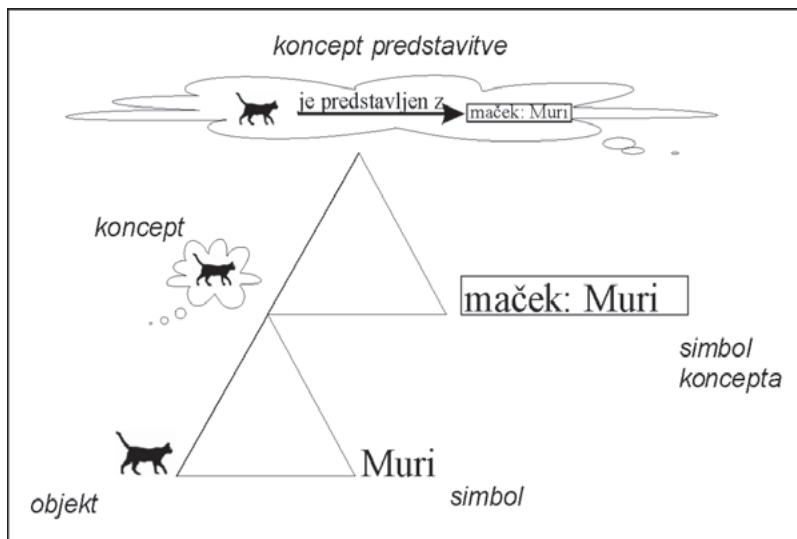
Večina uporabnikov tehnike in orodij za konceptualno modeliranje nima zanesljivega znanja o formalnih znanostih, kot sta logika in matematika. Zato je zaželeno, da se obdrži enostavnost predstavitev znanja. Zaželeno je, da bi uporabnik že s površnim pregledom vsebine informacijske osnove in konceptualne sheme lahko dobil pregled nad objektnim sistemom in tudi na enostaven način upravljal znanje v tem sistemu.

Ena najenostavnejših in najmanj formalnih metod predstavljanja znanja je kartiranje mišljenja, misli, spomina. Miselne karte so popolnoma neformalni grafi, ki se uporabljajo v začetnih korakih



modeliranja. Miselne karte, miselne vzorce je težko primerjati med seboj, nenazadnje zato, ker nimajo enotnega sistema simbolov. Zato je težko odkrivati in odstranjevati napake v miselnih kartah, še posebej v nizih miselnih kart, ki predstavljajo isto področje obravnave. Celovitost predstavitve znanja je logični koncept, pri katerem je za vsako dejstvo, ki je v obravnavanem delu stvarnosti pravilno, mogoče potrditi pravilnost tudi v konceptualnem modelu.

Integracija dveh miselnih kart je skoraj nemogoča, ker je običajno uporabljena različna raven abstraktnosti (slika 7) in ker ne obstaja način opredelitve postopka za njuno integracijo. Tako pri konceptualnem modeliranju vedno obstaja kompromis med enostavnostjo predstavitve znanja in celovitostjo – temeljitostjo in pravilnostjo predstavitve znanja.



Slika 7: Predstavitev koncepta s simbolom koncepta (povzeto po Sowa, 2000b)

### 3.2 FORMALNO MODELIRANJE KONCEPTOV

Za izvedbo formalnega opisa področja obravnave obstajajo osnovna načela in miselne zasnove modeliranja kot so objekt, objektni identifikator, razmerja med objekti, atributi, lastnosti objektov, vrednosti atributov in druge. Beseda objekt se uporablja v dveh kontekstih oziroma miselnih povezavah, v stvarnosti in v modelih. Objekt je stvar ali pojav v stvarnem svetu. Objekti v stvarnem svetu so lahko predmetni – konkretni ali pojmovni – abstraktni. Včasih se kot sopomenka za stvarni objekt uporablja beseda entiteta. Objekt je tudi osnovni modelni gradnik, okrog katerega so opredeljeni ostali modelni gradniki.

Jezik je formalen, če slovnica jezika opredeljuje vrste stavkov, ki spadajo jeziku. Jezik je sestavljen, strukturiran, če je mogoče dejstva stvarnega sveta izraziti le na en način ali izjemoma na več načinov. Različne načine izražanja istega dejstva stvarnega sveta je treba nedvoumno opredeliti. Področje obravnave je mogoče opisati tudi v naravnem jeziku. Za izognitev pomanjkljivostim naravnega jezika, na primer dvoumnosti in podvajanju, se je pri konceptualnem modeliranju treba osredotočiti na uporabo umetnih jezikov.

Za utemeljevanje objektnega sistema so potrebni jezikovni simboli, ki se nanašajo na določene stvarne objekte. Takšni simboli se imenujejo objektni identifikatorji. Objektni identifikatorji zagotavljajo prijemališča objektov v objektnem sistemu. Razliko med stvarnostjo in jezikom je mogoče predstaviti s sledečo trditvijo: objekti pripadajo svetu obravnave – delu stvarnosti, objektni identifikatorji pa pripadajo jeziku, ki je uporabljen za utemeljevanje objektnega sistema. Objekti so na različne načine v razmerju drug z drugim. Jezikovni simboli, ki predstavljajo razmerja med objekti, se imenujejo atributi. Za jezikovno predstavitev določenega razmerja med dvema objektoma se odgovarjajoča objektna identifikatorja s primernim atributom med seboj povežeta v atributni stavek. Objekti imajo lahko lastnosti, ki se modelirajo z jezikovnimi simboli, ki se prav tako imenujejo atributi. Lastnosti objektov so v jeziku predstavljene z atributnimi stavki in atributnimi vrednostmi (Boman et al., 1997).

Objekti so lahko med seboj podobni z določenega vidika. V takšnem primeru se pogosto uvrstijo v isto skupino oziroma razred. Za predstavitev razredov v jeziku se uporabljajo objektni tipi. Utemeljitev pripadanja določenega objekta v določen razred se izvede v jezikovnem stavku objektnega tipa. Stavek je sestavljen iz objektnega identifikatorja, ki se v stavku objektnega tipa imenuje dokaz – argument stavka.

Če se povzamejo zgornja načela konceptualnega modeliranja, je mogoče objektni sistem na kratko opisati kot niz objektov, ki pripadajo različnim razredom, imajo različne lastnosti in so na različne načine povezani med seboj. Objekti v stvarnem svetu niso nespremenljivi ali neuničljivi. Ustvarjeni so v določenem časovnem trenutku, pridobivajo lastnosti in jih izgubljajo, se povezujejo in razdružujejo in na koncu v določenem časovnem trenutku lahko preminejo. Smiselno je razmišljati o objektnem sistemu, ki se v času neprestano spreminja, kot o nizu posnetkov sistema v različnih trenutkih v času.

Vsak posnetek predstavlja stanje objektov, njihovih lastnosti in povezav v določenem časovnem trenutku. Kar povzroči spremembo stanja objektnega sistema, je dogodek. Ob dogodku se ustvarijo ali prenehajo novi objekti, lastnosti in razmerja med objekti. Za jezikovni prikaz stanja objektnega sistema se uporablja opis stanja, ki vsebuje niz stavkov objektnih tipov in atributnih stavkov. Obseg objektnega tipa v določenem stanju objektnega sistema vsebuje vse objektno identifikatorje, ki so primeri objektnega tipa. Obseg atributov v določenem stanju objektnega sistema je niz vseh parov objektnih identifikatorjev, povezanih med seboj z atributi.

### 3.2.1 Konceptualni grafi

Zasnova konceptualnih grafov temelji na Peirceovi logiki stvarnih grafov in teoriji kazalcev. Konceptualni grafi so po obliki izomorfni – enaki ugnezdenim strukturam predstavitev področja obravnave (Discourse Representation Structures – DRS), ki jih je za namene predstavitve in razreševanja kazalcev v naravnih jezikih razvil Hans Kamp (1981). Neodvisno razvite teorije, ki se stekajo v semantično enakovredne predstavitve, so dokazale veljavnost in so tako bolj zanesljive osnove za standardizacijo kot soglasja, dosežena v različnih odborih za standardizacijo (Sowa, 1984; 1987). Neformalno je konceptualni graf (Conceptual Graph – CG) struktura konceptov in konceptualnih razmerij, kjer vsaka povezava povezuje konceptualno vozlišče in vozlišče konceptualnega razmerja. Za oblikovanje konceptualnih grafov pripravlja tehnični odbor standard ISO/JTC1/SC 32/WG 2 (Sowa, 1998).

### 3.2.2 Formalno modeliranje in geografski prostor

Splošni namen uporabe formalnih teorij za modeliranje geografskega prostora je dvoplasten (Bittner in Frank, 1999): namen predstavitve človeškega znanja o objektih in pojavih v geografskem prostoru ter predstavitve pravil, ki uravnavajo te objekte in pojave na formalen način; namen uporabe formalnih metod utemeljevanja za izpeljavo novih znanj o dejstvih o geografskih objektih, razmerjih med njimi ter nova znanja o pravilih, ki uravnavajo geografske objekte in njihova razmerja.

Ker so znane omejitve formalnih metod, je očitno, da ne more obstajati enotna, splošna teorija, ki bi zaobjela vsa dejstva in pravila prostora geografskih razsežnosti in ki bi bila lahko uporabljena za računalniško obdelavo. Posledično je potrebno uporabiti posamezne teorije o posameznih pojavih in posameznih vidikih. Ontološka ustreznost formalne teorije se nanaša na izbiro posameznih enot, razmerij in operacij, in to na tak način, da se vnaprej ohranja razlikovanje entitet sveta (Guarino, 1994). Epistemološka ustreznost se nanaša na predstavljanje in razlaganje znanja o geografskem prostoru (McCarthy, 1977). Samo tisto, kar je znano, je mogoče predstaviti nedvoumno določeno. Ločeno ontološko in epistemološko obravnavanje geografskih objektov omogoča ustreznost različnih tehnik abstrakcije (geometrične, lokacijske, kvalitativne).

Uporaba formalnih metod za predstavitev geometrije in lokacije prostorskih objektov je povezana z izdelavo končne in ločene jezikovne predstavitve, formalne teorije neskončnih in nepretrganih področij, geografskega prostora (Bittner in Frank, 1999). Jezikovne predstavitve neskončnih in nepretrganih domen se dosežejo z abstrakcijo in ločevanjem. Eden od načinov predstavljanja neskončnih in nepretrganih domen je način predstavitve ustreznih razredov posameznih enot namesto predstavitve posameznih enot. Posamezne enote istega razreda so obravnavane kot enakovredne. Med njimi ni razlik. Vsaka posamezna enota pripada izključno enemu ustreznemu razredu.

V postopku predstavljanja razredov posameznih enot (namesto predstavitve posameznih enot) se pojavlja problem opredelitve ustreznih razredov na tak način, da se sestavne lastnosti izvirne domene posameznih enot ohranijo. To je mogoče doseči tako, da se strukture, ki uravnavajo domeno posameznih enot, uporabijo za opredelitev ustreznih razredov. Obstaja več skupin struktur, ki jih je mogoče uporabiti za opredelitev ustreznega razreda posameznih enot. Posameznim tehnikam abstrakcije odgovarjajo posebne skupine struktur. Obravnavane skupine struktur (Bittner in Frank, 1999) so naslednje: lastnosti in razmerja, ki pod vplivom določenih skupin transformacij ostajajo nespremenljivi; referenčni okvirji in lokacije posameznih enot znotraj okvirja; vrednosti posamezne enote, ki predstavljajo občutne spremembe, ter raven razmerij, ki kvalitativno strukturirajo domeno.

Tehnike abstrakcije geografskega prostora, ki uporabljajo prednosti navedenih struktur, so: geometrijske abstrakcije – temeljijo na opredelitvi ustreznosti/enakosti glede na nespremenljivost razmerij in konfiguracij posameznih enot; lokacijske abstrakcije – temeljijo na opredelitvi ustreznih razredov posameznih enot, utemeljenih na lokaciji v referenčnem okvirju; kvalitativne abstrakcije – temeljijo na opredelitvi ustreznosti glede na vrednosti posameznih enot in red razmerij med njimi.

### 3.3 KONCEPTUALNO MODELIRANJE IN INFORMACIJSKI SISTEMI

Osnovni namen informacijskega sistema je zagotovitev modela področja obravnave, ki ljudem s preučevanjem modela omogoča pridobivanje podatkov in informacij o stvarnosti. V tem smislu deluje informacijski sistem kot nedejavno skladišče podatkov, ki odraža sestavo in obnašanje področja obravnave. V posebnih primerih lahko informacijski sistem izvede aktivnosti, ki vplivajo na stvarnost področja obravnave. Informacijski sistem se z vidika konceptualnega modeliranja sestoji iz treh sestavnih delov. Konceptualna shema opisuje jezik, ki je uporabljen za razlaganje nominalne osnove, določa, kateri stavki so lahko vključeni v informacijsko osnovo in opisuje vedenje nominalne osnove preko dogodkovnih pravil. Podatkovno/informacijska osnova se sestoji iz stavkov (objektnih tipov in atributnih stavkov), ki pojasnjujejo trenutno stanje modela. Informacijski predelovalec – procesor je programski paket, katerega namen je uporabniku omogočiti poizvedovanje in obnavljanje konceptualne sheme in informacijske osnove (Boman et al., 1997).

#### 3.3.1 Konceptualna shema

Konceptualna shema ima tri sestavne dele: jezik, niz izpeljevalnih pravil ter niz dogodkovnih pravil. Določitev jezika z uporabo abecede in slovnice; abeceda je razdeljena v pet razredov simbolov; od teh se pri opredelitvi jezika navajata le dva, in sicer tipe ter attribute; simboli in slovnica se brez navajanja obravnavajo kot vsi, to so: objektni identifikatorji, spremenljivke in ločila. Niz izpeljevalnih pravil je sestavljen iz jezikovnih izrazov; namenjen je izpeljavi novih stavkov iz obstoječih stavkov informacijske osnove; niz integritetnih omejitev je namenjen selektivni izbiri stavkov, ki smejo biti v informacijski osnovi. Niz dogodkovnih pravil, ki opisujejo vedenje področja obravnave; dogodkovna pravila so ločena na zunanja in notranja; dogodki, ustvarjeni na pobudo dejavnikov, agentov v področju obravnave, se imenujejo zunanji dogodki; nekateri dogodki so lahko ustvarjeni na pobudo informacijskega sistema in se imenujejo notranji dogodki. Za pojav notranjih dogodkov je mogoče oblikovati pravila, ki določajo, ob katerih razmerah naj se dogodki pojavijo. V nekaterih primerih se lahko izvede notranji dogodek, ne da bi ta neposredno povzročil predmetne spremembe v področju obravnave. Sprememba v področju obravnave se zgodi v primeru, ko je notranji dogodek mogoče označiti kot izvrševalno uporabo jezika preko deklarativnega govornega nastopa, deklarativnega stavka. V takšnem primeru se dogodek sestoji iz izvedbe govornega nastopa s strani informacijskega sistema oziroma agenta informacijskega sistema. Ta izvedba stavka spremeni stvarnost področja obravnave. Ne spremeni se predmetna, ampak socialna stvarnost, in sicer se oblikuje nova pravica, obveznost ali dovoljenje, na primer posebna pravica.

#### 3.3.2 Informacijska osnova

Konceptualna shema ne opisuje stanj področja obravnave. Opis stanj je podan v informacijski osnovi. Ta je niz atributnih stavkov, izraženih v jeziku konceptualne sheme. Elementi informacijske osnove se imenujejo stavki osnove. Nekateri stavki osnove za opis stanja stvarnosti so lahko z uporabo izpeljevalnih pravil izpeljani iz informacijske osnove in se imenujejo izpeljani stavki. Izraz konceptualni model pomeni konceptualno shemo, skupaj z odgovarjajočo informacijsko

osnovo. Informacijska osnova krši integritetne zahteve, in to v primeru, da zahteva ni zadovoljena v opisu stanja področja obravnave – odgovarjajoči informacijski osnovi. Informacijska osnova, ki krši nujne integritetne omejitve, je nepravilna ali pa je nominalna osnova v nezaželenem stanju. V takšnem primeru se mora izvesti primerna aktivnost za ureditev stanja (Boman et al., 1997).

### **3.3.3 Informacijski procesor**

Informacijski procesor uporabniku omogoča poizvedovanje in obnovitev konceptualne sheme in informacijske osnove. Uporabnik izvaja poizvedovanje z vložitvijo vprašanja informacijskemu procesorju. Vprašanje je formula, ki je izražena v jeziku konceptualne sheme in običajno vsebuje spremenljivke. Odgovor informacijskega procesorja je odvisen od informacijske osnove kot tudi od izpeljevalnih pravil konceptualne sheme.

Ko se v stvarnosti pojavi dogodek, mora biti informacijskemu sistemu sporočeno, da se lahko ustrezno spremeni informacijska osnova. Uporabnik zato z vložitvijo dogodkovnega sporočila obvesti informacijski procesor o pojavu dogodka. Dogodkovno sporočilo je sestavljeno iz dogodkovnega tipa in spiska podatkov, vrednosti. Ko informacijski procesor prejme dogodkovno sporočilo, na informacijski osnovi izvede ustrezno spremembo. Sprememba se imenuje izvršitev, transakcija spremembe vrednosti. Za izvedbo spremembe informacijski procesor preudari dogodkovno sporočilo kot tudi ustrezna dogodkovna pravila.



# 4

---

## FILOZOFSKE IN SIMBOLNE TEORIJE PROSTORA

Medopravilnost prostorskih tematik je skladnost pomena kodiranih podatkov o prostoru, skladnost ontologij znanj o prostoru. Za računalniško integracijo znanj o prostoru je pomembno vedeti, kako ljudje upravljajo podatke o prostoru v svojem umu. Uporaba ontologij v geoinformatiki je predmet živahnih raziskav, zato so v tem poglavju na kratko navedene nekatere filozofske teorije prostora. Obravnavani so koncepti lokacije, koncepti predmetnih in nepredmetnih objektov ter tehnike abstrakcije geografskega prostora. Začetki ontologije segajo do Aristotela, ki je spoznal vrednost členitve sveta v smiselno povezane razrede. Svet je začel deliti na njegove sestavne dele in postopke. Prve poskuse je izvedel za živalski in rastlinski svet (taksonomija) in zaključil delo v pisnih delih Fizika in Metafizika. Razvrščanja v razrede in podajanje prioriteta sta temeljni operaciji pri gradnji ontologije.

Teorija nizov v povezavi s koordinatno geometrijo ima pri obravnavi prostora določene prednosti, vendar prinaša tudi določene probleme. Prazen prostor je v ontološkem smislu popolnoma homogen. Kar velja v enem predelu prostora, velja tudi v vseh drugih. Po teoriji nizov obstaja sporna metafizična teza, da je prazen prostor zgrajen iz točk. Če se sprejme teza, da so prostorske regije nizi, potem sledi, da so prostorske regije abstraktni objekti. Vendar se postavlja vprašanje, kako so lahko abstraktni objekti takšne narave, da jih lahko zasedajo trdne, predmetne stvari (Smith in Mark, 1998).

Bolj nepristranska teorija je mereologija, ki jo sestavljata atomizem in njegova negacija. Mereologija za osnovo privzema prostorsko nepretrganost, kontinuum. Skrajna različica mereološkega pristopa »od zgoraj navzdol« je aristotelski mereološki potencializem, po katerem »del celote nikoli ne more biti predmetna stvar, če je celota predmetna stvar«. Del celote je lahko le potencialno predmetna stvar; del celote bi postal predmetna stvar le s postopkom ločitve od celote. Tako je mogoče opredeliti podenote in nadenote. Mereološki potencializem se izogiba razsipanju; priznava samo en dejanski, predmetni objekt – univerzum.

Doktrina mereološkega aktualizma (Brentano, 1988) postavlja trditev, da »imajo vsi posamezni deli celote enake ontološke pravice kot njihova celota«. Iz tega izhaja, da so posamezni deli stvari ravno toliko stvarni, kot so dejanske stvari same po sebi. Mereološki aktualizem je mogoče razširiti na obravnavo prostorskih meja, notranjih in zunanjih. Potrebno je zavrniti tezo teorije nizov, ki pravi, da »meje kot izolirane točke lahko obstajajo neodvisno od entitet višjih razsežnosti, katere omejujejo«. Meje so dejanske stvari, vendar so odvisne entitete. Obstajajo lahko le v paru z večjimi stvarmi ali regijami, ki so njihovi gostitelji (Chisholm, 1993, Barry, 1993).

Pri obravnavi nepraznega prostora so razmere povsem drugačne kot pri obravnavi praznega prostora. Poleg zaloge homogenih prostorskih objektov, ki obstajajo v smislu osnovne geometrije prostora, je mogoče razlikovati različne vrste prostorskih objektov. Med njimi je mogoče

ločiti (Smith, 1994): naravne, bona-fide prostorske objekte (planeti, Luna, otoki, jezera, reke), ki obstajajo na osnovi notranjih fizičnih prekinitev v snovni sestavi Zemlje, ter umetne, fiat prostorske objekte (države, regije, lastniške zemljiške parcele). Umetna vrsta objektov nima notranjih fizičnih prekinitev, ampak so meje utemeljene na osnovi dogovora ali so na drugačen način izdelki človeškega geografskega delovanja.

Fiat prostorski objekti so, v primerjavi z zgolj geometričnimi objekti praznega prostora ter v primerjavi z naravnimi prostorskimi objekti, človeške kreacije, entitete, superdodatki svetu, posledica človeških kognitivnih aktivnosti in delovanja. V zgodovini metafizike obstaja večno nasprotje med stvarmi, ki so najdene oziroma odkrite, in med narejenimi oziroma ustvarjenimi stvarmi. Za izdelavo ontologije geografskega prostora je mogoče izbirati med več doktrinami. Skrajni idealizem je doktrina, ki trdi, da so vsi objekti ustvarjeni oziroma da obstajajo kot izdelek ali domislek človeškega poznavanja. Zmerni realizem je doktrina, ki trdi, da so nekateri objekti odkriti, nekateri pa ustvarjeni. Skrajni realizem (geometrični objektivizem, platonistična, Meinong metafizika) je doktrina, ki trdi, da so vsi objekti odkriti, natančneje: da so vsi objekti najdeni in ne izdelani.

Razprave je vredna primerjava skrajnega in zmernega realizma. V skrajnem realizmu, geometričnem objektivizmu, se razumejo politični in administrativni prostorski objekti, ki jih je sicer mogoče obravnavati kot ustvarjene na osnovi »fiat« dogovora, kot izključno logični sestavi geometrijskih objektov in le-tem je podeljeno ontološko popolno priznavanje. Na drugi strani se zmerni realizem sklicuje na naravnost. Tako priznava, da je določen politični oziroma politično-administrativni prostorski objekt, na primer država, prestal lastno zgodovino. Tako je v toku časa lahko spremenil svojo velikost, obliko ali celo lokacijo v primerjavi z drugimi prostorskimi objekti na površini Zemlje. Z vidika zmernega realizma je politično-administrativna entiteta primerljiva z živimi organizmi v smislu rasti in razvoja ter celo premikanja iz kraja v kraj, vendarle vedno na tak način, da se je ohranila njena veljavnost.

Za razliko od organizmov lahko administrativne entitete celo preživijo obdobja neobstoja, po katerem je njihova identiteta spet obnovljena na primer, Avstrija in Poljska. Prostorski objekti, kot so lastniške zemljiške parcele, politično-administrativne entitete, ki se lahko razdelijo ali združijo z istovrstnimi objekti na tak način, da so nove entitete ustvarjene iz delov prej obstoječih objektov (Cassati in Varazi, 1994). Gotovo je administrativna entiteta v vsakem trenutku skladna z določeno regijo, ki je zgolj geometričnega značaja, vendar ni nikoli identična s katerokoli geometrično regijo v prostoru.

Ontologija zgodovinsko obstoječih politično-administrativnih prostorskih objektov je izpeljana iz teorije večrazsežnostnih nepretgranosti (Brentano, 1988). Brentano osnuje domeno prostorskih objektov kot večplastno stavbo z domenami raznovrstnih drugotnih prostorskih objektov, zgrajenih nad osnovo homogene prostorske nepretgranosti nižje ravni. Teorijo večrazsežnostnih nepretgranosti je mogoče obravnavati tudi kot izpeljavo iz teorije izvrševanja, ki jo je razvil Adolf Reinach (1913) v delu Apriorne osnove civilnega prava. Razlikuje dva razreda govornih nastopov, izvrševalne uporabe jezika, ki povzročata dve različni vrsti posledic, in sicer abstraktno in konkretne.

Ukazi, zahvale, oprostitev, opozorila in grožnje so izvrševalna uporaba jezika s konkretnimi posledicami, kot so aktivnosti, vedenje, občutja ljudi ter entitete, ki so del stvarnega zgodovinskega sveta vzročnih sprememb. Obljubljanje, zakonodajalstvo, pogodbeništvo, priseganje, krščevanje, podeljevanje imen in plemiških naslovov so izvrševalna uporaba jezika z abstraktnimi posledicami, kot so terjatve, obveznosti, zakoni, pravice (vključno z lastninsko pravico), zaroke, viteštvo, imena, entitete, ki niso neposredno predmet vzročnih vplivov.

S tem opisom razlikovanja jezikovnih dejanj je mogoče opredeliti obstoj različnih posledic na



področju izvrševalnih uporab geografskih kart. Konkretno posledice izvrševalne uporabe kart so: akcije (iskanje poti, vojaška dejanja), občutki (ogroženosti, prestrašenosti, užaljenosti); (Monmoyer, 1991). Abstraktne posledice izvrševalne uporabe kart so: oblikovanje državnih, regionalnih in posestnih meja, ustanovitev mednarodne datumske linije. Tako kot abstraktne posledice izvrševalne uporabe jezika (terjatve, obveznosti in druge pravne entitete) spadajo tudi abstraktne posledice izvrševalne uporabe kart med platonske objekte, ki spadajo izven domene časa in sprememb, ter med stvarne objekte iz domene vzročnega toka (Twardovski, 1979).

Značilnost politično-administrativnih prostorskih objektov je torej odvisnost od posebnih človeških, fiat, dejanj, vendar nekateri objekti kažejo večjo stopnjo te značilnosti kot drugi. Obstajajo skrajno enostavni objekti, katerih meje so natančne geometrične oblike, običajno ravne linije. Politično-administrativni objekti s premočrtnimi mejami so posledica kolonializma. Zarisali so jih v vladah, še preden so poznali stvarno stanje na terenu (ZDA, Kanada, Mehika, Avstralija, Podsaharska Afrika). V Evropi so meje razčlenjene v odvisnosti od pazljivega opazovanja stanja na terenu, po ideji samoopredelitve narodov, in so tako bolj odvisne od posebnih človeških fiat dejanj. Za politično-administrativni objekt je značilno tudi dejstvo, da se tokom časa spremeni po obliki in lokaciji, vendar na tak način, da ostaja ločena entiteta, ki ohranja veljavnost.

Za narode, ki živijo na otokih, se zdi, da imajo kar najbolj naravne politično-administrativne meje, vendar so tudi te dejansko abstraktne posledice posebnih človeških dogovorov, fiat dejanj. Navidezna naravnost izgine, ko se upošteva status narodov kot politično-zgodovinsko stvaritev. Naravnost naravnih politično-administrativnih prostorskih objektov se zmanjšuje tudi ob dejstvu, da ni v obravnavi ena sama »zemljiška tvarina«, temveč gre za zapletene izdelke človeškega razmejevanja. Celo tam, kjer meje politično-administrativnih prostorskih objektov v veliki meri sovpadajo z rekami ali drugimi naravnimi geografskimi značilnostmi, te meje niso identične z omenjenimi geografskimi pojavi. Naravne geografske značilnosti se vedno nadomeščajo z določenimi nenaravnimi nadomestki, na primer meja vzdolž sredine reke.

Zdi se, da morajo biti meje politično-administrativnih prostorskih objektov neskončno tanke, ker bi bilo sicer mnogo sporov zaradi neolastninjenih zemljišč (Smith, 1995). Tej domnevi je mogoče oporekati, kajti kljub neskončno tankim mejam se netižno spori predvsem zaradi samega poteka meja. Teoretično neolastninjena zemljišča ne obstajajo, le vrste lastninskih pravic na njih se lahko zelo razlikujejo, na primer zasebna lastninska pravica, pravica do javnega dobra, skupna lastninska pravica – solastnina, državna lastninska pravica, lastninska pravica vaše skupnosti, kar lahko daje vtis neolastninjenosti. V nekaterih primerih meje različnih politično-administrativnih prostorskih objektov sovpadajo na celotni dolžini, zavzemajo identični prostorski položaj, vendar zaradi tega razloga obravnavani politično-administrativni prostorski objekti niso identični. Tudi v tem primeru vsak zase ohranjajo svojo veljavnost.

Obstaja tudi kategorija prostorskih objektov, ki jim manjka izrazita zunanja meja. Takšni objekti niso rezultat dogovornega, fiat kartiranja meja, kot je poznano v obravnavanih primerih meja socialnega izvora, kot so meje politično-administrativnih objektov. Meje za objekte te kategorije so pridobljene z opazovanjem, vzorčenjem in uporabo posebnih algoritmov, ki izvedejo meje območij pojavov. Pojavi – objekti z neizrazito zunanjo mejo so: območja politične moči iz zgodovine, meteorološke karte, puščave, oblaki, podvodni morski tok, jate ptic, jate rib. Grafični prikaz tovrstnih objektov je tematska karta. Tovrstni objekti so obstajali pred objekti socialnega izvora kot, so zgodovinske regije naselitve pred konstitucijo državnih meja. Meje politično-administrativnih objektov so socialnega in ne geografskega izvora – najprej zavedanje o pripadanju grupi, potem zavestna odločitev grupe za ustanovitev ali premik meje. V kategorijo prostorskih objektov, ki jim manjka izrazita zunanja meja, so uvrščeni vsi naravni, bona-fide objekti splošne prostorske spoznavne/kognitivne oblike naravnih prv in geografske razsežnosti, naravni elementi

v vseh svojih pojavnih oblikah.

Načela ontologije geografske stvarnosti podrobneje navaja Mark (2000). Geografske okolnosti in stvari niso zgolj locirane v prostor. Za geografske okolnosti je značilno, da so neločljivo umeščene v prostor, tako da privzemajo mnoge strukturne lastnosti prostora, kot so mereološke, topološke in geometrične lastnosti. Vloga in vrsta meja je pri obravnavanju ontologije geografske stvarnosti še posebej pomembna.

#### 4.1 KOGNITIVNO KARTIRANJE PROSTORA GEOGRAFSKIH RAZSEŽNOSTI

Kognitivno kartiranje in poenotenje ima za integracijo podatkov in informacij temeljni pomen, kajti izvedba analiz je zaradi integriranih podatkov in informacij mogoča na množični ravni, brez potrebe po takojšnji presoji podrobnosti. Standard XML obeta razvoj spleta od področja čiste skladnje, sintakse, dokazne teorije do področja pomena besed – pomenoslovja, semantike, v katerem bo pomen neločljivo povezan s predstavljenimi podatki (Koepsell, 2001). Pri tem se postavlja vprašanje, na kakšen način in s katerimi mehanizmi ljudje miselno zbirajo, organizirajo, priklicujejo in upravljajo podatke in informacije o prostoru. Na to vprašanje je mogoče odgovoriti s preučevanjem teorije kognitivnega kartiranja.

Kognitivno kartiranje je miselna obdelava prostorskih podatkov in informacij. Je miselna, mentalna sposobnost abstraktnega zbiranja, organiziranja, priklicevanja in upravljanja informacij o prostoru. Kognitivna karta je izdelek, shema v človekovem umu, in sicer miselna predstavitev nekaterih delov prostora v določenem časovnem preseku. Predstavitel nekaj simbolizira, je slikovit prikaz miselne slike oziroma miselnega modela v človekovih možganih. Predstavlja svet, za katerega oseba verjame, da je tak, četudi predstavitev ni nujno pravilna. Kognitivna karta je uspešna, če njena predstavitev ne vsebuje preveč neresničnosti (Downs in Stea, 1977).

Miselna predstavitev se nanaša na določen vidik človekovega okolja. V sklopu kognitivnega kartiranja prostora geografskih razsežnosti se miselna predstavitev nanaša na vidik vsakodnevnega niza aktivnosti v prostoru. V kognitivni karti se organizira in upravlja znanje o svetu z namenom osvojitve vedenja o položaju pojavov, stvari ali ljudi. S pomočjo kognitivnih kart, miselnih prostorskih predstavitev sveta, ljudje rešujejo prostorske probleme na abstraktne načine, z miselnim upravljanjem in kombiniranjem podatkov in informacij. Tako je mogoče dobiti odgovor na osnovni prostorski vprašanje: »Kje in kaj?«

Za rešitev kakršnega koli prostorskega problema sta potrebna dva osnovna tipa informacij, ki se nanašata na vprašanje: »Kje?« in »Kaj?«. Ta dva tipa informacij sta komplementarna, kajti pri iskanju odgovora na vprašanje »Kje?«, se ve »Kaj?« se išče.

Kako se informacije kognitivno obdelujejo? Človeška bitja imajo čutila – receptorje – in psihične sposobnosti za sprejemanje podatkov in ustvarjanje informacij. Svet, tak kot ga vidimo ljudje, je sinteza različnih vrst informacij, vizualnih, slušnih, vohalnih in tipnih. Kognitivne karte zato niso nujno le vizualne slike sveta. Posledično se kažejo občutne individualne razlike. Kognitivne karte reševanja istega problema so lahko funkcionalno podobne, a formalno različne.

### 4.1.1 Kognitivno kartiranje kot interaktivni postopek

Za reševanje prostorskih problemov s pomočjo miselnih predstav sveta je značilna neprestana interakcija z okoljem, ki se kognitivno kartira. V jedru tega procesa je povratna zveza učenja in izvajanja. Narava interakcije učinkuje na tip pridobljene prostorske informacije kakor tudi na podatkovni vir – akcijo opazovanja, s katerim so pridobljene informacije. Held in Rekosh (1963) sta s serijo poskusov dokazala, da je pravilnost zaznave sveta odvisna od neposredne interakcije med organizmom, ki sprejema informacije, in okolico oziroma prostorom, ki je opazovan. Motorična aktivnost sprejemnika v prostoru opazovanja je ključna za povezavo zunanjega okolja z notranjo predstavitvijo prostora.

### 4.1.2 Selektivnost pri kognitivnem kartiranju

Pri kognitivnem kartiranju je človeški um selektiven z dveh vidikov (Downs in Stea, 1977): selektivnost v smislu izbire informacij, ki jih človek miselno kodira in shrani; to vključuje tudi odločitve o tipu informacij, načinu simboliziranja informacij, ureditve shranjevanja in določitvi pomembnosti informacije; selektivnost v smislu postopka čitanja kognitivne karte, katere informacije bodo priklicane in kakšne strategije upravljanja priklicanih informacij bodo uporabljene. V jedru selektivnosti je koncept kriterijev. Za kognitivno kartiranje obstajata dva glavna kriterija: funkcionalna pomembnost informacij in razločevalnost oziroma slikovitost. Kriterij funkcionalne pomembnosti upravlja izbiro informacij v smislu splošne uporabnosti informacij v vsakodnevem vedenju, aktivnostih in individualnem odločanju v prostoru. Kriterij razločevalnosti oziroma slikovitosti se nanaša na obliko in vidnost – vizualno dominantnost, ki odražata kontrast in ureditev prostora samega po sebi.

Postopek selektivnosti se občasno spreminja, na kar vpliva predvsem vrednostni sistem osebe. Glede na vrednostni sistem se včasih zavestno ignorirajo določeni vidiki prostora in obravnavanega okolja. Tako obstaja še selektivni kriterij kognitivnega prenašanja elementov okolja v novo okolje. Prenašanje poteka s povezavo kognitivnih predstav prostora in niza pravil, ki izhajajo iz preteklih izkušenj v prostoru. Mogoče ga je obravnavati z vidika teorije kopiranja in konstruktivistične teorije. Po teoriji kopiranja se izkušnje v okolju odražajo kot neposredne kopije in odslikujejo podobo okolja, shranjeno v možganih. Kognitivno kartiranje je aktivni konstruktivni proces. Navidezna enostavnost selektivnih kriterijev je zavajajoča, kajti kognitivno kartiranje poleg selektivnosti tudi povezuje, organizira in sintetizira številne informacije. Brez sposobnosti reorganizacije nesistematičnega zbiranja prostorskih podatkov bi bile prostorske izkušnje nerazumljive za praktično uporabo. Tako se v nedoločenem organizacijskem postopku informacije urejajo v uporabno obliko urejenega znanja o prostoru.

Organizacijska narava kognitivnega kartiranja je v iskanju pomena stvari. Za organiziranje podatkov o izkušnjah v prostoru je potreben čitljiv okvir, ki mora biti nepretrgan in sestavljen iz razpoznavnih, a povezanih delov. Del organizacijskih povezav izkušenj v prostoru je odvisen od svojstvenih lastnosti prostora kot samoorganizacije. Preostali del okvira je prežet z osebnim pomenom. Kognitivne predstavitve prostora tako postanejo polnopomenske. Kognitivno kartiranje ni le selektivno, ampak tudi sintetično in organizacijsko. Pomembni vidiki kognitivnega kartiranja so skriti v vsebini in kognitivnih predstavitvah okolja. Vprašanji, kot sta »Kako se s sposobnostjo kognitivnega kartiranja kodirajo in dekodirajo informacije?« in »Kakšne oznake ljudje uporabljajo za predstavitev svojega prostorskega okolja?«, je mogoče pojasniti z obravnavo enot vsebine predstavitev ter organizacijo predstavitev.

Kategorije identitete in kategorije ustreznosti so osnovna gradiva za konstruiranje kognitivnih predstavitev. Vendar kategorije v kognitivnem kartiranju niso idealen odraz prostorskega okolja. Niso enolično ponovljive niti niso kartografske karte. So zelo generalizirane in poenostavljene rekonstrukcije izkustvenega okolja in rezultat istočasnih in ponavljajočih se procesov interakcije s prostorom ter selekcije in organizacije informacij o prostoru. Za pisno, ustno ali grafično komunikacijo je potreben slovar porazdeljenih kategorij.

Kakšne oblike imajo prostorske miselne kategorije, če je znanje o prostorskem okolju dosegljivo z razvojem in uporabo kategorij? Kategorij ni mogoče opazovati neposredno. Na njihov obstoj je mogoče sklepati le iz pripovedovanja ali obnašanja ljudi. Kakšno fiziološko obliko kategorije zavzemajo v možganih, ni mogoče ugotoviti. Mogoče pa je ugotoviti, da je kategorije lahko obnoviti kot zavestni priklic ali pa kot komunikacijo z drugimi osebami. Za izražanje znanja o prostoru je mogoče izbirati med dvema sredstvoma, in sicer med podobami in besedami. Ta možnost izbire odraža prožnost kognitivnega kartiranja. Besedno izražanje in izražanje s podobami je mogoče nadomeščati eno z drugim. Obe sredstvi se med seboj dopolnjujeta. Zamenjava med obema sredstvoma se zgodi v trenutku. Besedno izražanje ima v medosebni komunikaciji očitne prednosti pred slikovnim. Slikovno izražanje kognitivnega znanja o prostoru je bolj osebno in ima večjo izrazno moč.

Za zunanje predstavitve miselnega prostorskega znanja s slikovnim sredstvom izražanja se uporablja niz simbolov, ki niso enaki nizu kognitivnih simbolov v notranji predstavitvi v človeškem umu. Tudi za besedno izražanje obstaja omejen niz simbolov, ki omejuje izražanje internih kognitivnih simbolov navzven. Če oseba izvede miselni prevod, translacijo iz enega v drugo izrazno sredstvo, rezultat verjetno predstavlja spretnost osebe za izvedbo translacije, ne le načina predstavitve znanja o prostoru. Problem enolične translacije postane enostavnejši, če se obravnava interno prostorsko znanje – informacije – kot slovar (seznam besed), interne simbole pa kot slovnična pravila.

Vsaka predstavitev znanja, notranja ali zunanja, je rezultat uporabe kognitivnih simbolov nad prostorskim znanjem. Kuipers (1994) trdi, da temelji teorija kognitivnega kartiranja na hierarhiji predstavitev prostorskega znanja, ki se imenuje prostorska semantična hierarhija (Spatial Semantic Hierarchy). Vsaka raven hierarhije ima svojo lastno ontologijo, niz objektov in relacij za opisovanje sveta in svoj niz metod reševanja problemov in sklepanj.

### 4.1.3 Klasifikacije znanj o prostoru pri kognitivnem kartiranju

Znanje o prostoru in njegovo klasifikacijo je mogoče obravnavati glede na več klasifikacijskih razsežnosti (Mark, 2000): narava prostorskega znanja (deklarativna, postopkovna in konfiguratívna); viri prostorskih informacij za kognitivno kartiranje; človeška interakcija s svetom in prirejena uporaba jezikov. Klasifikacija znanj o prostoru, pri upoštevanju narave prostorskih znanj, opredeljuje relativno stabilne oblike in vidike prostorskih znanj, ki se nahajajo pri odraslih ljudeh, ne pa pri otrocih.

Trditveno, deklarativno prostorsko znanje imenujemo tudi geografska dejstva (Freundschun, 1991). Tipični primeri trditvenega prostorskega znanja so dejstva o lokaciji, velikosti in populaciji geografskih objektov. Trditveno prostorsko znanje je mogoče pridobiti iz izkušenj v stvarnosti, kot del postopkovnih znanj, iz geografskih kart (konfigurativna znanja) ali iz drugih medijev, kot so knjige, časopisi, filmi, televizijski programi in osebna komunikacija.

Postopkovno prostorsko znanje je posledica sposobnosti ljudi za izvajanje različnih nalog v prostoru, na primer naloge iskanja poti v geografskem prostoru – navigacije. Postopkovno pro-

storsko znanje lahko vsebuje informacije, ki niso dostopne zavestnemu mišljenju. Kuipers (1978) je modeliral postopkovno znanje o prostoru kot kolekcijo parov »pogled-akcija« (PA) in pogled, ki bo sledil akciji »pogled-akcija-pogled« (PAP). Znanje o prostoru je pogosto razdeljeno med »stvarnost« in »mišljenje«, kar pomeni, da so za doseganje nekaterih znanj potrebne izvedbe prostorskih akcij, vendar je določen del ključnih informacij v stvarnosti sami (Liben, 1988).

Konfiguracijsko prostorsko znanje geografskih razsežnosti je podobno geografski karti in ima približno ali popolno evklidsko geometrijo. Konfiguracijsko prostorsko znanje lahko obstaja po delih. Topološko znanje, ki je vrsta posrednega znanja med trditvenim in postopkovnim znanjem, običajno vsebuje povezave med objekti (MacEachern, 1992). Pri polno razviti obliki konfiguracijskega prostorskega znanja lahko oseba oceni absolutne razdalje in smeri tako natančno, kot lahko to stori, če gleda na geometrično pravilno geografsko karto.

Mark (1993) navaja tri temeljne vire prostorskih informacij za kognitivno kartiranje, zasnove, ki nastopajo v človeškem prostorskem spoznavanju, in sicer dotikalni – haptični prostor, prostor podob, ki se doživlja skozi gledanje, ter premikalno zaznavni prostor (trans-perceptualni), ki se doživlja pri iskanju poti skozi prostor. Navedeni prostori so hierarhično urejeni tako, da je vsak delno izgrajen na konceptu prejšnjih. Semantika oziroma pomenoslovje simbolov in besed je pomembna sestavina formalnega ontološkega specificiranja.

## 4.2 SEMIOTIKA

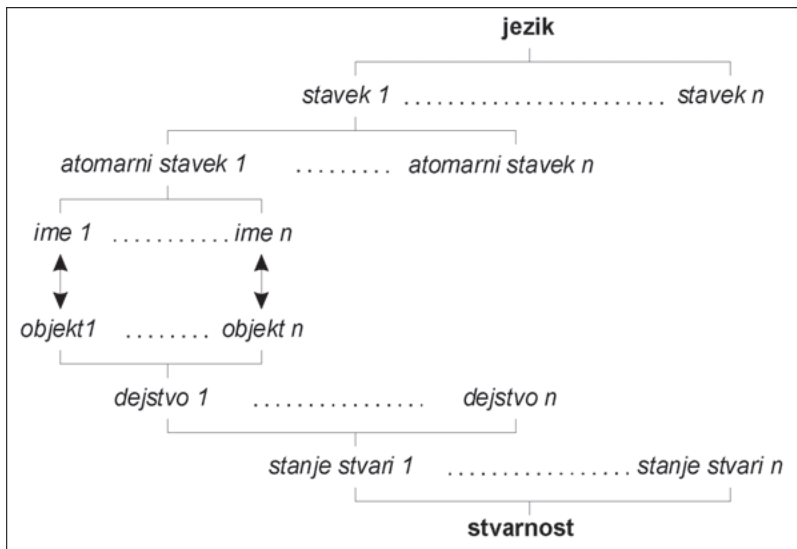
Semiotika (semiologija, splošna semantika) je veda o znakih in je del teorije informacij, ki obravnava komunikacijo med različnimi vrstami agentov preko jezika in drugih simbolnih sistemov. Semiotika je sorodna kognitivni znanosti v tem, da je interdisciplinarna veda, katere področje se nanaša na jezikoslovje, logiko in filozofijo (Sainsbury, 1991).

Semiotiko sta neodvisno razvila logik in filozof Charles Sanders Peirce in jezikoslovec Ferdinand de Saussure (1916). Izraz semiotika izvira iz grške besede za znak, *sēma*. Po Peirceu je semiotika disciplina, ki proučuje uporabo znakov vsake inteligence, ki je zmožna učenja na osnovi izkušenj. Po tem kriteriju bi se lahko računalniške tehnike za obdelavo zbirk znanja in zbirk podatkov imenovala računalniške semiotike. Semiotika obravnava tri področja jezika: skladnjo – sintakso, pomenoslovje – semantiko in uporabnost – pragmatiko. Sintaksa je po Peirceu čista slovnica; je preučevanje povezav med znaki; s sintakso se obravnavajo načini sestavljanja stavkov iz manjših sestavnih delov, kot so znaki abecede, besede in fraze; odkriva splošna pravila za sestavljanje stavkov. Semantika je logična primernost, ki je formalna znanost o pogojih pravilnih predstavitev; semantika je preučevanje povezav med znaki in stvarmi in povezav vzorcev znakov z vzorci stvari. Pragmatika je čisto govorništvo – retorika; pragmatika je preučevanje povezav med znaki in agenti, ki jih uporabljajo za sklicevanje na stvari v stvarnosti in za sporazumevanje o stvareh z drugimi agenti, ki imajo podobne ali različne namene glede istih ali različnih stvari. Struktura jezika, razmerja med stvarnostjo in jezikom ter dejanska uporaba jezika v smislu komunikacije so še posebej pomembni, zato so jezikoslovne teorije temelj konceptualnega modeliranja (Gamut, 1991).

#### 4.2.1 Jeziki in konceptualno modeliranje

Jezik je simbolni sistem, ki ga posamezniki in skupine uporabljajo za medsebojno komunikacijo. V človeški komunikaciji se uporabljajo tudi drugi simbolni sistemi, med drugimi tudi grafični simbolni sistemi prostorskih načrtov in geografskih kart. Naravni jeziki so neprimerni za strojno, računalniško sklepanje, zato je za tovrstno logično sklepanje treba razviti umetne jezike. Konceptualno modeliranje z umetnim jezikom je zelo učinkovito vendar se poraja vprašanje, kako uporabiti rezultate preučevanja umetnih jezikov za prilagoditev naravnega jezika. To vprašanje zadeva nalogo formalizacije. S formalizacijo pomena besed iz naravnih jezikov se ustvarjajo ontologije za raziskovanje informacijskih sistemov in njihovih aplikacij (Mark et al., 2003).

Razmerje med jezikom in stvarnostjo je predstavljeno z diagramom (slika 8). Diagram je simetričen – jezik je ogledalo stvarnosti. Poleg tega, da jezik služi komunikacijskim namenom, ga je mogoče uporabiti tudi kot podporo spretnosti mišljenja, dokazovanja in sklepanja. Uporaba jezika je najučinkovitejši način za izvedbo jasnega poteka razmišljanja. Z razumno uporabo jezika je mogoče natančno dokazovati in preverjati pravilnost dokazovanja. Pri preučevanju vloge jezika v dokazovanju se je uporaba naravnih jezikov izkazala kot neustrezna. Glavni razlog temu je neprimernost naravnih jezikov za izražanje logičnih lastnosti idej, ki bi jih radi izrazili (Boman et al., 1997).



Slika 8: Simetrični diagram kaže, da je jezik sestavljen iz stavkov, ki so sestavljeni iz atomarnih stavkov. Pomen stavkov je mogoče ugotoviti iz pomenov atomarnih stavkov. Atomarni stavki so sestavljeni iz kombinacij konstant in predikatnih razmerij, ki odgovarjajo objektom stvarnega sveta (povzeto po Boman et al., 1997)

Najpomembnejša lastnost naravnih jezikov je besedna dvoumnost. Isti jezikovni simbol ima lahko več pomenov, o čemer se lahko prepričamo v vsakem slovarju. Za samodejno pomensko obdelavo podatkov predstavlja besedna dvoumnost problem. Obstaja tudi nejasnost pomena stavkov, česar ni mogoče pripisati besedni dvoumnosti, zato se imenuje strukturalna nejasnost. Težavo naravnih jezikov predstavljata tudi besedna in strukturalna prekomernost, redundanca.

Besedna redundanca se pojavi, kadar različni besedi oziroma simbola predstavljata isti objekt ali idejo. Strukturna redundanca nastopa, kadar je lahko ista ideja izražena z več različnimi stavki, celo če stavki vsebujejo iste besede.

Vsem sistemom komunikacije je skupno, da uporabljajo simbole, to pomeni objekte, ki se lahko nanašajo na druge objekte. V človeškem jeziku se uporabljajo simboli, ki se v splošnem imenujejo besede. Jeziki se razlikujejo od večine drugih simbolnih sistemov po ključnih lastnostih, ki dajejo jeziku izrazno moč in raznovrstnost (Pinker, 1994): samovoljnost jezika – v splošnem ne obstaja resnična, notranja povezava med jezikovnim simbolom in tistim, kar simbol predstavlja; nejasnost, težka razumljivost – stavek je lahko povzetek bolj podrobnih zamisli in je osredotočen le na en vidik določene situacije; ustvarjalnost – za sestavljanje in razumevanje neskončnega števila izjav je potrebno le poznavanje končnega števila besed in pravil; izrazna moč – v primerjavi s komunikacijskimi sistemi živali je človeški jezik neomejen glede možnosti predstavitve.

Načelo kompozicije v jezikoslovju navaja, da je pomen stavka v popolnosti določen s posameznimi pomeni sestavnih delov stavka. Mogoče je dokazati, da pomen stavka poleg pomenov sestavnih delov določa tudi sobesedilo, kontekst stavka. Zato načelo sestavljanja ni vedno primerno za analiziranje pomena stavkov. Za razumevanje pomena posameznih besed in izjav je potrebno razlikovati med njihovimi podaljšanji in osredotočenostmi. Podaljšanost izjave se nanaša na objekt ali niz objektov v stvarnosti, na katere se izjava nanaša. Osredotočenost izjave se nanaša na pomen, ki ga osebe običajno razumejo v izjavi.

Za analiziranje pomenov stavkov je bila razvita trditvena – propozicijska logika. Atomarni deli jezika (besede, fraze) se predstavijo s trditvenimi simboli – črkami. Trditvam se pomen pripisuje z določanjem pogoja pravilnosti trditve. Takšen način podajanja pomena izrazom se imenuje semantika Tarskega (Tarski, 1993). V jeziku trditvene logike je mogoče izvajati le zelo grobe analize pomenov stavkov. Bolj podrobna analiza je mogoča z jezikom vzdevčne, predikatne logike, imenovane tudi logika prvega reda (First Order Logic – FOL). Predikatni račun podaja jezikovne sestave za označevanje posameznih objektov, lastnosti objektov in razmerij med objekti.

#### 4.2.2 Semantika

Semantika je študij pomena izrazov uporabljenega jezika – razmerij med besedami in stvarnostjo. Semantika se nanaša na pomen izrazov v jeziku ali logiki v nasprotju s sintakso, ki obravnava le obliko izražanja v jeziku. Semantika je druga veja semiotike, ali – kakor je semantiko imenoval Peirce – logična primernost. Prvi sistem logike je razvil Aristotel v logičnih razpravah, ki jih je sestavil v zbirko imenovano Organon. Zbirka obsega razprave o interpretaciji, apriorni analizi, posteriorni analizi, temah in sofističnih, navidezno resničnih ovrzbah trditvev.

Gottlib Frege (1879) je razvil popolni sistem logike, ki se imenuje logika prvega reda. Drugi popolni sistem logike je bil algebrائي zapis predikatnega računa, ki ga je neodvisno od Fregeja razvil Charles Sanders Peirce (1883, 1885). Ker je bil predikatni račun za večino uporabnikov preveč matematično obarvan, je Peirce (1897) izdelal tretji kompletni sistem logike, ki se imenuje stvarni oziroma eksistencialni grafi.

Peircovi stvarni grafi so semantično enakovredni obema prej razvitima zapisoma (logiki prvega reda in algebrائي zapisom predikatnega računa), vendar stvarni grafi na najbolj enostaven način kartirajo logično-semantične prvine. S stvarnimi grafi je Peirce ustvaril najbolj prvinske oblike za izražanje elementov logike. Četudi je razvil grafični zapis za izražanje prvin, jih je mogoče enakovredno izraziti v naravnih jezikih, algebrائي zapisih ali v drugih linearnih, grafičnih ali celo govorjenih predstavitevah.

Navedene semantične prvine so dosegljive v vseh naravnih jezikih in v vseh različicah logike prvega reda. Imenujejo se semantične prvine, ker segajo preko slovnično-sintaktičnih razmerij do znakov za semantična razmerja med znaki in stvarmi. Vsak zapis, s katerim je mogoče izraziti navedene prvine v vseh mogočih kombinacijah, vsebuje logiko prvega reda kot podniz. Različni jeziki lahko uporabljajo različne zapise za predstavitev prvin za obstoj, sklicevanje, razmerje, sovpadanje in razveljavitev (preglednica 1). Obstoj je v večini naravnih jezikov vključen pri omenjanju stvari. Za poudarjanje obstoja imajo jeziki tudi določene eksplicitne izraze obstajanja. V algebraini notaciji za logiko je obstoj izražen z eksplicitnim simbolom  $\exists$ . V umetnem jeziku SQL se uporablja beseda »exists«. Sklicevanje – dva različna znaka se nanašata na isto stvar. V naravnih jezikih za sklicevanje obstajajo implicitne in eksplicitne metode; določniki, zaimki in različne oblike glagola biti. V linearnih zapisih za logiko se uporabljajo spremenljivke in znak za enakost. V grafičnih zapisih za logiko se uporabljajo povezovalne črte ali povezovalni loki. V umetnem jeziku SQL se uporabljajo spremenljivke in znak za enakost »= $\equiv$ «. Razmerje – relacija – je v naravnih jezikih in linearnih zapisih za logiko izraženo z besedami. V umetnem jeziku SQL se razmerja imenujejo tabele ziroma razredi. Sovpadanje – konjunkcijo – je mogoče v naravnih in umetnih jezikih izraziti vpleteno, z navajanjem stavka za stavkom, ali določno, z uporabo besede in oziroma simbola » $\wedge$ «. V umetnem jeziku SQL se uporablja beseda »and«. Za razveljavitev – negacijo se v vseh naravnih jezikih in umetnih načinih izražanja logike uporabljajo besede oziroma simboli. V umetnem jeziku SQL se za razveljavitev uporablja beseda »not«.

<b>Prvina</b>	<b>Neformalni pomen</b>	<b>Primer</b>
obstoj	Nekaj obstaja.	<i>Tam je pes.</i>
sklicevanje	Nekaj je enako kot nekaj.	<i>Pes je moj ljubljeneček.</i>
razmerje	Nekaj je v razmerju do nečesa.	<i>Pes ima uši.</i>
sovpadanje	A in B.	<i>Pes teče in laja.</i>
razveljavitev	Ne A.	<i>Pes ne teče.</i>

Preglednica 1: Semantične prvine izražene v naravnem jeziku (privzeto po Sowa, 2000b)

Z uporabo petih prvin je mogoče opredeliti druge operatorje (preglednica 2). Najpogostejši operatorji so posploševalni (univerzalna mera), posledični (implikacija) in ločilni (disjunkcija). Našteti operatorji ne štejejo za semantične prvine, ker niso neposredno opazni.

<b>Operator</b>	<b>Primer</b>	<b>Prevod v semantične prvine</b>
posplošitev	<i>Vsak pes laja.</i>	$\text{not}(\text{tam je pes and not}(\text{laja}))$
posledica	<i>Če je pes, potem laja.</i>	$\text{not}(\text{tam je pes and not}(\text{laja}))$
ločevanje	<i>Pes laja ali mačka mjavka.</i>	$\text{not}(\text{not}(\text{pes laja}) \text{ and } \text{not}(\text{mačka mjavka}))$

Preglednica 2: Opredelitev logičnih operatorjev (privzeto po Sowa, 2000b)

Peirce je poleg petih semantičnih prvin opredelil tudi prvino odvisnosti od nedavne omembe – konteksta, ki jo je imenoval kazalec. V naravnih jezikih so kazalci predstavljeni z zaimki, kazalnimi



zaimki. V izražanju logike s konceptualnimi grafi je kazalec označen z znakom #. Tudi lastna imena so kazalci.

Algebraino izražanje logike, ki je običajno v matematiki, je le en način od mnogih semantično enakovrednih zapisov. Pet semantičnih prvin in mehanizme za opredeljevanje drugih operatorjev logike prvega reda je mogoče prirediti veliko različnim načinom zapisovanja, vključno z naravnimi jeziki in zapisoma XML (Extensible Markup Language) in RDF – omrežje opisovanja virov (Resource Description Framework). Logika je lahko predstavljena na veliko načinov grafičnih ali linearnih zapisov. Primera grafičnih zapisov sta stvarni in konceptualni graf. Primer linearnega zapisa je izmenjevalni format znanja (Knowledge Interchange Format – KIF).

Za boljšo berljivost zapisov je mogoče predstaviti logične operatorje v nadzorovanih naravnih jezikih, ki uporabljajo podniz sintakse in slovarja naravnega jezika. Tako je mogoče vsebino konceptualnega grafa prevesti v drugo obliko formalnega zapisa s stilizirano različico naravnega jezika. Logika predstavlja, skupaj s formalno opredeljeno semantiko, sredstvo za izvajanje semantično enakovrednih prevodov med jeziki s temeljito različnimi sintaksami, pod pogojem, da je opredeljena primerna ontologija.

#### 4.2.3. Združevanje logične primernosti z ontologijo

Čista logika je ontološko nevtralna. Ne postavlja nikakršnih predpostavk o tem, kaj obstaja oziroma bi lahko obstajalo v katerikoli domeni ali kateremkoli jeziku, s katerim se govori o domeni. Da bi bilo mogoče predstaviti znanje o določeni domeni, mora biti domena podprta z ontologijo, ki opredeljuje kategorije stvari v tej domeni in opredeljuje izraze, ki jih ljudje uporabljajo v pogovoru o kategorijah. Ontologija opredeljuje besede naravnega jezika, predikate predikatnega računa, koncepte in relacije konceptualnih grafov, razrede objektnousmerjenega jezika, ali tabele in stolpce – podatkovne attribute relacijske zbirke podatkov. Za ponazoritev opredeljevanja ontologije se pogosto uporabljajo konceptualni grafi.

Formule predikatnega računa in konceptualni grafi uvajajo ontološke predpostavke. Predstavitev dejavnosti z različnimi koncepti je mogoča s Peircovo ontologijo, ki se imenuje dogodkovna semantika in predstavlja akcijo z jasno ločenimi entitetami in relacijami tipa agent in tema.

Bolj popolna opredelitev koncepta mora vsebovati namen agentov. Relacija namen povezuje koncept akcije s konceptom situacije, ki bo nastopila po uspešnem zaključku akcije. Namen je triadna relacija, kar pomeni, da vsebuje dva eksplicitna argumenta in en implicitni argument (agent), katerega namen je doseči željeno situacijo. Koncept situacije je ugnuzden v blok konteksta situacije, ker je namenski status različen od konteksta akcije. Če se akcija ne zaključi uspešno, se namen ne uresniči – situacija ne nastopi. Ontologija situacij in njihovih predstavitev temelji na Peirceovi logiki, ki je kombinirana z idejami, razvitimi na področju umetne inteligence, jezikoslovja, filozofije in logike (Sowa, 2000b).

Kontekstni blok lahko obsega namenske situacije oziroma časovno ali prostorsko ločene dele večje situacije. Večje situacije so predstavljene z vrsto ugnuzdenih situacij. Za izogibanje večkratnemu križanju povezovalnih linij v konceptualnih grafih se uvedejo sklicevalne etikete. Metanivojsko razmerje med situacijo in predlogom je izvedeno z opisno trditvijo, ki je ustrezna kontekstualnim okvirom v grafični predstavitvi.

Semiotika s svojimi vejami, torej sintakso, semantiko in pragmatiko, zagotavlja usmeritve za organiziranje in uporabo znakov za predstavljanje nečesa (entitete) nekomu (agentu) z določenim namenom. Poleg predstavljanja podpira semiotika tudi metode za prevajanje vzorcev znakov namenjenih enemu namenu, v druge vzorce, ki so namenjeni za drugačno, vendar sorodno

uporabo. V ontologijah za medmrežne objekte se zanemarjajo fizični objekti, procesi, ljudje in njihove namere z neustreznim in premajhnim številom kategorij. Dvoumnost imen in njihovih nejasnih povezav s kontekstom je problem poimenovalnih dogovorov na spletu. Poimenovalni dogovori so semiotške pojave, ki so lahko predstavljene z metanivojskimi tipi in razmerji v konceptualnih grafih ali z drugimi zapisi, ki temljijo na logiki.

Ontologije vsebujejo razrede, vrste – kategorije, leksikoni vsebujejo pomene besed, terminološki slovarji vsebujejo izraze, spiski vsebujejo naslove, katalogi vsebujejo številke izdelkov, zbirke podatkov vsebujejo števila, črke in binarne velike objekte (BLOB). Vsi ti spiski, hierarhije in mreže so tesno povezane zbirke znakov. Vendar njihova primarna povezava ni v bitih, ki kodirajo te znake, ampak je povezava v mišljenju ljudi, ki zbirke znakov tolmačijo. Pomembno je spoznati, kako je mogoče uporabiti metode logike in ontologije za opredelitev, povezavo in pretvorbo znakov med besednjaki. Analizirati je treba razlike pomenov in njihov učinek na podatke v medmrežju. Računalniki so vse bolj zmogljive naprave, zato uporabniki od njih pričakujejo pomoč pri reševanju zapletenih nalog in problemov. Želeli bi, da bi bili računalniki bolj razumni, da bi lahko delovali bolj samostojno. Ker so razvijalci programov ljudje, katerih um ni dobro prilagojen reševanju zapletenih problemov – tako kot pri vseh ljudeh, si pomagajo tako, da povezujejo več enostavnih programov. Povezovanje programov pa pomeni, da programi kličejo drug drugega in si pri tem izmenjujejo podatke. Programi so lahko porazdeljeni v omrežju ali več omrežjih. Lahko so izdelani z različnimi programskimi orodji in tečejo na računalnikih z različnimi operacijskimi sistemi. Pred standardom XML je vsak program na svoj način shranjeval podatke – praviloma v datoteke. Če so bili ti podatki namenjeni ljudem, težav praviloma ni bilo, ker so naši možgani izredno prilagodljivi. Težava nastopi, kadar je treba raznovrstne podatke prebrati v drugem programu, ki ga je izdelal drug razvijalec. Torej XML lahko naredi prenos podatkov med programi bolj učinkovit. Razvoj programov, ki si izmenjujejo podatke, je lahko s pomočjo standarda XML hitrejši, cenejši in bolj zanesljiv, saj vsebuje manj skritih napak.

XML lahko igra tudi drugo pomembno vlogo. Je jezik za standardizacijo podatkov, s katerimi upravljajo ljudje. Z XML lahko opredelimo podatkovni slovar, besednjak na določenem področju. Natančno lahko določimo pomen besed. Podatkovne slovarje pa lahko procesirajo tudi računalniki. Standard XML rešuje tudi težave podvajanja podatkov. Iste podatke, ki jih predstavljamo na različne načine, imamo shranjene v več različicah. Na primer: pogosto potrebujemo iste podatke na papirju in na spletni strani (dve različni datoteki). Če želimo narediti spremembo, obnoviti podatke, moramo to storiti dvakrat. XML omogoča zapis podatkov v nevtralni obliki, ki jo lahko na standarden način predelamo v druge oblike.

Cilj različnih metapodatkovnih predlogov je izvedba miselnih povezav z označevanjem, etiketiranjem podatkov z več znaki. Ti metapodatkovni znaki imajo nadaljnje povezave, ki so lahko etiketirane z metanivojskimi znaki. Vendar brezpomenski podatek s pripenjanjem brezpomenskih znakov ne more pridobiti pomena. Prvobitni vir pomena je stvarni svet in zastopniki – agenti, ki uporabljajo znake za predstavitev entitet stvarnega sveta in namenov teh znakov – opisnih opredelitev.

Jezik SGML (Standard Generalized Markup Language) in njegov naslednik, standard XML, omogočata jasno ločitev med formatiranjem besedila in pomenom besedila. Za oba standarda so značilni označeni podatki. Oba jezika uporabljata za predstavitev pomenov oznake – etikete v besedilu. Oznake pišemo v zašiljenih oklepajih:

<oznaka>podatek</oznaka>

Standard XML določa pravila za pisanje označenih podatkov. Pravzaprav je XML metajezik. To je jezik, ki določa pravila za definiranje novih podatkovnih jezikov – na primer GML. Črka X v kratiki XML označuje razširljivost jezika. HTML je jezik za predstavitev podatkov na spletni strani. Tako kot XML tudi HTML uporablja oznake za podatke. Vendar oznake določajo, kako prikazati podatke, o samih podatkih pa HTML-oznake ne povedo ničesar. V XML pa oznake nič ne povedo o predstavitvi podatkov, ampak opisujejo podatke same. Standard XML omogoča uporabo oznak, ki opisujejo podatke.

Etikete morajo imeti jasno opredeljeno semantiko. Vendar večina priročnikov za uporabo XML ne vsebuje navodil za predstavitev pomena podatkov. Spremembe v terminologiji računalniških besednjakov povzročajo težave pri samodejni računalniški primerjavi sestavnih delov zbirk podatkov različnih ponudnikov podatkov (Sowa, 2000b). Internet je ogromen semiotski sistem. Je masivna zbirka treh vrst znakov (po Peircu): ikon, ki predstavljajo obliko nečesa, kazalcev, ki kažejo na nekaj, in simbolov, ki predstavljajo nekaj v skladu z določenim dogovorom. Nekateri predlogi za ontologije in meta podatke prezirajo najpomembnejše lastnosti znakov. Vsak znak ima tri vidike: entiteto, ki jo predstavlja znak, drugo entiteto in agenta, ki znak razlaga. Z ločenim obravnavanjem znakov samih se izgubita vidik entitete, ki jo predstavlja znak, in vidik agenta (človeka, stroja), ki razlaga pomen znaka.

Omrežje opisovanja virov (Resource Description Framework – RDF) je orodje, ki temelji na XML-notaciji (Bray, 1998). XML in RDF s standardizacijo zapisov predstavljata pomemben prvi korak k porazdeljevanju pomena podatkov, vendar ta korak ni zadosten. Za porazdeljevanje podatkov so potrebni načini za primerjanje, povezovanje in pretvorbo besednjakov. Vendar standardizacija besednjakov ne more samodejno zagotoviti medopravilnosti. Nasprotno, standardizacija besednjakov bi lahko celo povzročila več težav pri porazdeljevanju podatkov, in sicer z zakrivanjem zapletenosti za umetnimi dogovori. XML je le metajezik za izražanje razumevanja pomena podatkov. Potrebno pa je spoznati stvarnost in razlike pomenov pri izmenjavi podatkov (Phipps, 2000). Bolj pomembno od standardiziranja besednjakov je razvijanje metod za opredeljevanje in pretvorbo slovarjev. Z uglaseno semantiko in pragmatiko se morajo izrazi besednjakov povezati s stvarmi, na katere se nanašajo, in z ljudmi, ki izraze uporabljajo za sporočanje informacij o teh stvareh.

#### 4.2.4 Povzemanje logike iz naravnega jezika

Ker je mogoče vseh pet semantičnih prvin izraziti v vsakem naravnem jeziku, je mogoče predstaviti logiko prvega reda kot podniz vsakega naravnega jezika. Tak podniz se imenuje slogovni, stilizirani ali nadzorovani naravni jezik in ga lahko razume vsakdo, ki lahko bere nenadzorovani naravni jezik. Za večino ljudi je torej potrebna le vaja za pisanje v nadzorovanem naravnem jeziku. Za polsamodejno analiziranje pomenov v nenadzorovanih jezikih je bil razvit sistem za povzemanje znanja iz pisnih listin (knowledge extraction – KE, Document-Based Knowledge Management – DBKM), (Skuce, 2000).

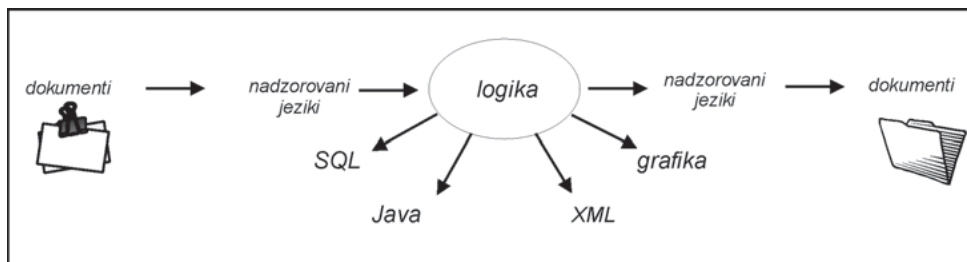
Sistem povzemanja znanja, ki se imenuje ClearTalk, uporablja nadzorovani angleški jezik. Stavki tega jezika so shranjeni v zbirki znanja ali pa so pripisani kot beležka izvornim listinam. Zelo bogat in nedvoumen, a nadzorovani angleški jezik, ki uporablja Kampovo (1981) teorijo razreševanja kazalcev, je ACE (Attempto Controlled English), (Fuchs et al., 1998; Schwitler, 1998), (slika 9).

If a copy of a book is checked out to a borrower  
 and a staff member returns the copy  
 then the copy is available.  
 If a staff member adds a copy of a book to the library  
 and no catalog entry of the book exists  
 then the staff member creates a catalog entry  
 that contains the author name of the book  
 and the title of the book  
 and the subject area of the book  
 and the staff member enters the id of the copy  
 and the copy is available.

Slika 9: Primer uporabe ACE (privzeto po Fuchs et al., 1998)

Najstarejši logični vzorci, izraženi v nadzorovanem naravnem jeziku, so štirje tipi stavkov sistema logičnega sklepanja, sistema silogizmov. Vsako pravilo logičnega sklepanja združuje glavno trditev in pomožno trditev za izpeljavo sklepa. Tipi stavkov logičnega sklepanja so: splošno pritrdilno, podrobno pritrdilno, splošno nikalno in podrobno nikalno. Navedeni tipi stavkov in logični sklepi, ki temeljijo na njih, so v uporabi v mnogih nadzorovanih naravnih jezikih in umetnih jezikih. Za izražanje dedovanja v ekspertnih sistemih in objektno usmerjenih sistemih je glavna trditev tipa splošna pritrditev, in to v stavku z glagolom »je«. Pomožna trditev je tudi tipa splošna pritrditev ali pa podrobna pritrditev z glagoli »je«, »ima« in drugimi. Za izražanje omejitev v zbirkah znanja in podatkov je največkrat uporabljena glavna trditev tipa splošno nikalno, ki preprečuje nezaželena sovpadanja – konjunkcije.

Stavke, izražene v naravnih jezikih, ki imajo polno izrazno moč logike prvega reda, je mogoče prevesti v jezik SQL. Izdelanih je bilo že mnogo sistemov povpraševalnih jezikov, temelječih na nadzorovanih naravnih jezikih, vendar nobeden od njih še ni postal tržno uspešen, ker je potreben znaten trud za opredelitev simbolov v slovarju in za kartiranje simbolov v primerna polja v zbirki podatkov. Če pa se orodja sistema za povzemanje znanja uporabijo za oblikovanje zbirke podatkov, je slovar ustvarjen kot stranski izdelek oblikovalskega postopka. Oblikovalski in specifikacijski jeziki imajo več metaravni. UML (štiri ravni: na metaravni jezik opredeljuje generični jezik, sintakso in semantiko zapisov UML; na meta ravni jezik opredeljuje splošne tipe UML; na ravni systemske analize so opredeljeni izvedbeni tipi kot primeri tipov UML; na izvedbeni ravni so izvedeni primeri izvedbenih tipov. Za prikaz enotnega pogleda vseh navedenih ravni so Gerbé (1997) in Gerbé et al. (1995, 1996, 1998a, 1998b) izdelali oblikovalno orodje, ki za predstavitev jezika na vsaki od ravni uporablja



Slika 10: Tok informacij od dokumentov do računalniške predstavitev (povzeto po Sowa, 2000b)

konceptualne grafe. Za opredeljevanje konceptualnih grafov samih je razvil ontologijo za uporabo konceptualnih grafov kot metameta jezika. Ontologijo je prenesel na druge načine zapisa, vključno z jezikom UML. Interno je zbirka znanja shranjena v konceptualnih grafih, zunanje pa je mogoče grafe prevajati v angleški ali francoski jezik. V tej izvedbi se je kot pomembna izkazala jezikovna neodvisnost konceptualnih grafov.

Slika 10 prikazuje tok informacij – pomena podatkov, od dokumentov k logiki in nadalje k dokumentom oziroma različnim računalniškimi predstavitev. Črtkana puščica med dokumenti in nadzorovanimi jeziki predstavlja potrebo po človeški podpori. Neprekinjene puščice predstavljajo popolnoma samodejno prevajanje, ki je lahko izvedeno v enem ali več sistemih. Jezik za poenotenje predstavitev pomena je logika, ki je lahko izvedena v različnih podnizih in izrazih različnih programskih orodij. Vsi podnizi vendarle uporabljajo isti slovar izrazov naravnega jezika, ki kartira isto ontologijo konceptov in razmerij. Razlike med programskimi orodji niso posledica razlik v jeziku ampak razlik v namenih programskih orodij.

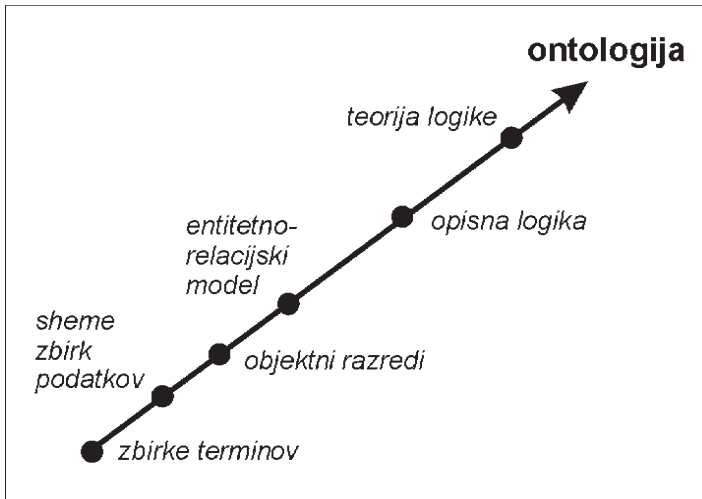
Pragmatika je tretji del semiotike, ki obravnava dejansko, izvrševalno uporabo jezika – namena in učinka izrekanja stavkov. Obstaja veliko stavkov, ki ne prenašajo informacij, ampak so enakovredni akcijam. V takšnih primerih izrekanje stavkov pomeni govorni nastop. Govorni nastopi se delijo v več kategorij, med katerimi je za prostorsko ontologijo najpomembnejša kategorija napovedovalnih, deklarativnih govornih nastopov, zaradi učinkov in posledic, ki se z njimi dosejajo.

### 4.3 ONTOLOGIJA IN INFORMACIJSKI SISTEM

Ontologije se uspešno uporabljajo v industriji, ne da bi jih uporabniki tako imenovali. Konec devetdesetih let dvajsetega stoletja se je uporaba ontologij začela pospešeno razvijati na področju informacijske tehnologije, še posebej za informacijsko poizvedovanje (McGuinness, 1998, Guarino et al., 1999). Na področjih upravljanja zbirk podatkov, standardizacije v informatiki in razvoja programske opreme se pojavljata potreba po presejanju razlik in potreba po skupnem razumevanju simbolov obravnavanih domen. Skupno razumevanje je mogoče z doseganjem soglasja o primernem načinu konceptualizacije obravnavane domene ter njenem nedvoumnom izražanju z izbranimi jeziki. Ontologija je rezultat soglasja o skupnem razumevanju domene in jo je zato mogoče uporabiti za različne namene v široki paleti miselnih povezav ter skupnih ciljev, ki vključujejo zamisli medopravnosti in ponovne uporabnosti.

Značilnosti ontologije navaja Uschold (1998) v obliki sledeče opredelitve: »Ontologija ima lahko različne oblike, vendar mora vsebovati slovar terminov ter nekatere opredelitve pomena navedenih terminov. Z opredelitvijo in navedbo povezav med pojmi se določi – vsili – sestava domene in omeji mogoče interpretacije simbolov na dovoljene« Ontologija se šteje za najpomembnejše sredstvo za reševanje semantične raznolikosti informacijskih sistemov. Gre za pristop reševanja problemov z eksplicitnim specificiranjem semantike terminov, ki se uporabljajo v informacijskih sistemih (Zhan et al., 2000), (slika 11).

Jasaper in Uschold (1999) sta razvila zasnovo za razumevanje in uvrščanje ontoloških aplikacij. Večji del zasnove se sestoji iz niza scenarijev uporabe ontologij. Scenarij je primer uporabe (use case) za predstavitev značilnosti med seboj podobnih aplikacij. Uporabni primer pojasnjuje določeno situacijo, v kateri se ontologija uporabi za določen namen. Poleg niza



Slika 11: Spekter ontologij (povzeto po Zhan et al., 2000)

scenarijev se zasnova sestoji iz naslednjih ključnih dimenzij: namena ali koristi, vloge ontologije, nabora akterjev za izvedbo scenarija, podporne tehnologije ter ravni zrelosti ontologije. V nekaterih scenarijih sta uporabljeni še dve dimenziji: način predstavitve pomena simbolov, formalni oziroma neformalni, ter porazdeljevanje oziroma izmenjava informacij.

V osnovi so ontologije uporabne za izboljšavo komunikacije, bodisi med ljudmi bodisi med računalniki. Za izboljšavo komunikacije med ljudmi zadošča neformalna, vendar nedvoumna ontologija. Medopravilnost med raznovrstnimi računalniškimi sistemi se doseže s prevajanjem in usklajevanjem pomena med različnimi metodami modeliranja, jezikov in programskih orodij. Pri tem se ontologija uporabi kot izmenjevalni format. Koristnost uporabe ontologije v sistemskem inženiringu se izraža v več značilnostih (Jasaper in Uschold, 1999). Ponovna uporabnost; ontologija predstavlja osnovo za formalno kodiranje pomembnih elementov, lastnosti, postopkov ter njihove medsebojne povezanosti v obravnavani domeni. Iskanje informacij; ontologijo je mogoče uporabiti kot meta podatke, ki služijo kot iskalni seznam v podatkovnem skladišču. Zanesljivost; formalna predstavitljivost ontologije omogoča samodejno preverjanje doslednosti, kar zagotavlja večjo zanesljivost izdelane programske opreme. Podrobni opis; ontologija lahko služi kot osnova v postopku ugotavljanja zahtev in opredeljevanja podrobnih opisov sistemov informacijske tehnologije. Vzdrževanje; za sisteme, ki so grajeni s pomočjo formalnih ontologij, je mogoče enostavno izboljšati programsko dokumentacijo, kar znižuje vzdrževalne stroške. Pridobivanje znanja; z uporabo obstoječih ontologij kot začetnih stanj in osnov za vodenje pridobivanja znanja, je mogoče izboljšati hitrost in zanesljivost pri izgradnji sistemov, temelječih na znanju.

#### 4.3.1 Vloge ontologije v uporabnih primerih

Scenarij uporabe ontologij je lahko nepreverjena ideja, predstavitev majhnega obsega potencialne aplikacije v raziskovalnem okolju ali pa zrela aplikacija v tržnem okolju. Scenarij mora pojasnjevati, katere ontologije so uporabljene v uporabnem primeru in na kakšen način. V ta namen so opredeljene tri vloge podatkov v različnih scenarijih. Vloga podatkov kot

operativnih podatkov, ki so uporabljeni in izdelani med izvrševanjem aplikacije; oznaka te vloge je L0. Informacije na ravni L0 so zapisane z uporabo simbolov, navedenih v slovarju, opredeljenem na ravni L1. Vloga podatkov kot ontologije, kjer informacije opredeljujejo izraze in opredelitive pomembnih konceptov določene domene; oznaka te vloge je L1; ontologije so uporabljene v razvojnem postopku izdelave aplikacije. Vloga podatkov kot jezika za predstavitev ontologije, kjer se informacije uporabijo za zapis ontologij na ravni L1; oznaka te vloge je L2; informacije ravni L2 uporabljajo avtorji ontologij in razvijalci aplikacij. Iste informacije lahko v različnih kontekstih igrajo različne vloge. Za porazdeljevanje ali izmenjavo podatkov na ravni Ln je potrebno sklicevanje na model Ln+1. Primer ravni L0 je družina standardov STEP, ki opredeljujejo izmenjavo primerov shem preko kodiranja tekstov (ISO 10303-21, 1994). Za izmenjavo shem je potreben zapis shem na ravni L1, v jeziku EXPRESS. Za porazdeljevanje ontologij (raven L1) je potrebno sklicevanje na raven L2 (Ontolingua, OIL, RDS).

### 4.3.2 Vloge igralcev za izvedbo scenarija uporabe ontologije

Vsak scenarij uporabe določene ontologije vključuje več igralcev. Igralec (oseba ali aplikacija) v scenariju privzema vlogo. Igralci so lahko sledeči: avtor ontologije je vloga, ki jo običajno igra oseba; avtor operativnih podatkov je vloga avtorja operativnih podatkov; razvijalec aplikacije je vloga razvijalca aplikacije; uporabnik aplikacije je vloga uporabnika aplikacije; uporabnik znanja je vloga, ki jo igra oseba, ki uporablja znanje. Posamezne osebe ali aplikacije lahko igrajo v določenem scenariju več vlog. Obstaja obsežen nabor tehnologij, ki podpirajo razvoj ontoloških aplikacij: predstavitveni jeziki ontologij (UML, Express, Ontolingua, OIL, XML), jeziki za izmenjavo znanj (KIF, PIF, CDIF), translacijska orodja (Ontolingua translator, CDIF orodja, Step Tools) in orodja in standardi za porazdeljevanje objektov (CORBA, COM).

Pomen oziroma znanje, ki je zajeto v določenih ontologijah, je spremenljivo z vidika količine v ontologiji predstavljenega pomena kakor tudi v stopnji formalnosti predstavitve pomena. Količina pomena je neposredno povezana z omejevanjem mogočih interpretacij, kar služi zmanjševanju dvoumnosti. Manj, ko je mogočih interpretacij simbola in s tem dvoumnosti, večja je količina pomena. Stopnja formalnosti sega od naravnega jezika do formalne logike. Za občevanje med ljudmi je bolj primerna neformalna opredelitev pomena. Nizko stopnjo dvoumnosti je mogoče doseči s formalnimi opredelitvami, ki so v podporni dokumentaciji podprte z neformalnimi opredelitvami. To je še posebej pomembno za udeležence razvoja informacijskih sistemov, ki niso specialisti informacijskih sistemov. Na krajši rok je v delujočih informacijskih sistemih lažje uporabiti ontologije z manjšo količino pomena – lahke ontologije. Primer večkrat uporabljene lahke ontologije je semantično strukturiranje podatkovnih skladišč za namene iskanja in priklicevanja podatkov. Vendar celo pri opravlilu iskanja in priklicevanja podatkov očitna korist nastopi kot posledica bogatih ontologij. Za namene medopravnosti in pri razvoju translacijskih orodij si z lahkimi ontologijami ni mogoče pomagati.

Ontologije so glede na uporabo na področjih industrije in raziskav uvrščene v naslednje glavne skupine (Jasaper in Uschold, 1999): ontologija kot nepristransko avtorstvo, ontologija kot podrobnejši opis, ontologija za skupni dostop do informacij ter iskanje informacij na osnovi ontologije. Ontologija kot nepristransko avtorstvo; informacijski avtorski izdelek je zapisan v enem jeziku in je preveden v različne oblike za uporabo v več ciljnih informacijskih sistemih; koristi takšnega pristopa so večkratna uporaba znanja, izboljšano vzdrževanje in ohranjanje znanja. Ontologija kot podrobnejši opis; ontologija določene domene se oblikuje in uporabi

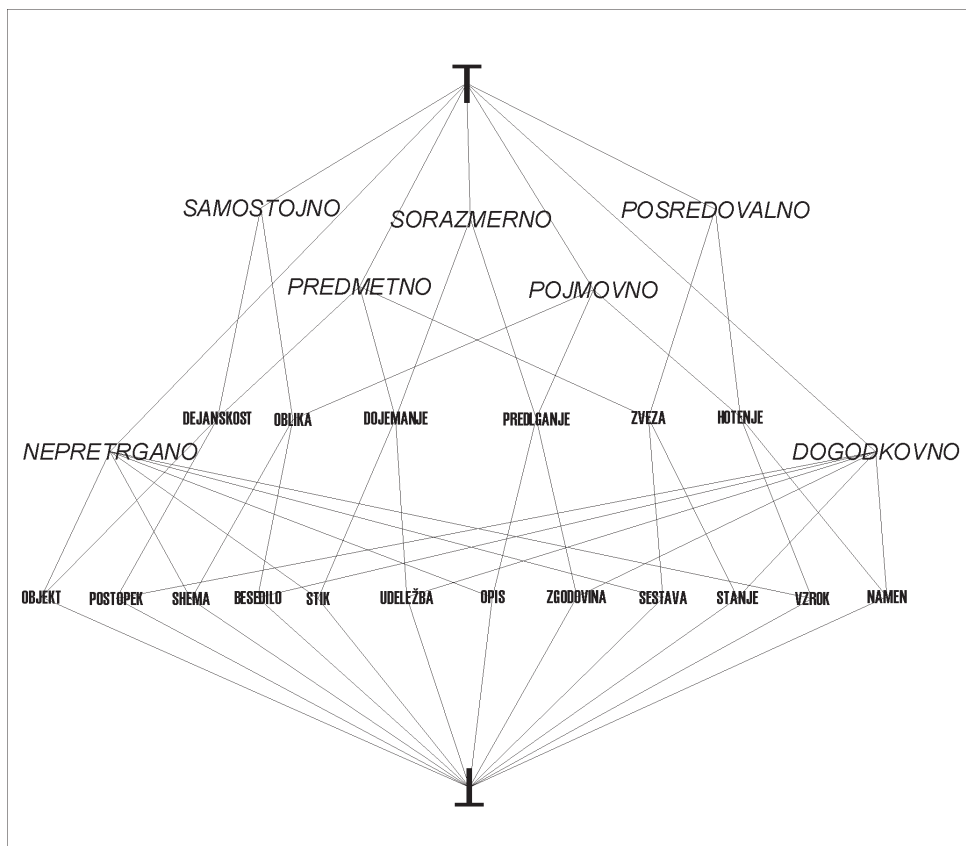
kot osnova za podroben opis in razvoj določene programske opreme; koristi tega pristopa vključujejo kakovostno dokumentacijo in vzdrževanje, zanesljivost aplikacije in večkratno uporabo znanja. Ontologija za skupni dostop do podatkov/informacij; informacije potrebuje več akterjev oziroma aplikacij z različnih problemskih področij, vendar so informacije izražene z nejasnimi izrazi nepoznanega slovarja; ontologija s porazdeljenim razumevanjem izrazov ali z jasnim kartiranjem med nizi izrazov pripomore k razumljivemu priklicu/predstavitvi informacij; koristi tega pristopa vključujejo medopravilnost in bolj učinkovito (večkratno) uporabo virov znanja. Iskanje informacij na osnovi ontologije; ontologija se uporabi pri iskanju informacij v podatkovnih skladiščih (dokumentih, spletnih straneh); izrazita korist je hitrejši pristop do pomembnih informacijskih virov, kar predstavlja doprinos k učinkovitosti večkratne uporabe virov znanja. Formalna ontologija (Stuckenschmidt, 2000) je opredeljena na formalno opredeljeni terminologiji. Ontologije imajo pomembno vlogo porazdeljevanja terminologije za postopke integracije informacij. Osnovni ontološki disciplini sta mereologija, teorija o celoti in njenih delih, ter topologija, teorija o stiku, združevanju in povezanosti. Uporabnost ontologije je odvisna od njene prilagojenosti splošni uporabi. To pomeni, da mora biti prilagojena ljudem glede razumevanja sestave, pregledovanja in iskanja obstoječih objektov v shemi razredov kakor tudi glede enostavnosti uvrščanja novih objektov v obstoječe razrede ontološke sheme. Težje je razviti uporabno ontologijo kot pa razviti natančno ontologijo. Razvoj splošne - univerzalne, idealne ontologije je že dolgo cilj ontologov. Splošna ontologija namreč poveže svojstvena in različna področja.

Splošno ontologijo je zasnoval Sowa (2000a) kot ontologijo predstavitev znanja. V taki ontologiji opredelitve povezujejo simbole z berljivim besedilom, ki opisuje, kaj simboli pomenijo oziroma opisuje pravila, ki omejujejo tolmačenje teh simbolov (slika 12). Splošno ontologijo opisuje z devetimi prvobitnimi kategorijami (supertipi): nepretrgano, samostojno, predmetno, sorazmerno, pojmovno, posredovalno, pojavno; s šestimi posredovalnimi kategorijami (podtipi): dejanskost, oblika, dojetanje, predložnost, zveza, hotenje ter z dvanajstimi kategorijami: objekt, proces, shema, besedilo, stik, udeležba, opis, zgodovina, sestava, stanje, vzrok, namen. Pojem ontologija povzroča težave pri obravnavi informacijskih sistemov (Gruber, 1993a). Koncept ontologije izvira iz filozofije, kjer se ontologija nanaša na sistematično obravnavo obstajanja, bitja, osnove, vzrokov in najsplošnejših lastnosti stvarnosti. Pogosto se ontologija zamenjuje z epistemologijo, ki je filozofska disciplina, ki obravnava izvor, strukturo, metodo spoznavanja in veljavnost spoznanja. Gruber (1993b) je kot pionir uporabe ontoloških metod v informacijski znanosti opredelil ontologijo kot eksplicitno specifikacijo konceptualizacije.

Z vidika porazdeljevanja informacij je ontologija opis konceptov za porazdeljevanje podatkov in razmerij med njimi, ki obstajajo za programskega agenta ali skupnost programskih agentov. Ta opredelitev se razlikuje od opredelitve ontologije v filozofskem smislu v tem, da filozofska opredelitev ne poudarja, komu je ontologija namenjena. Pri ustvarjanju ontologij za namene porazdeljevanja podatkov, informacij in znanj je pomembno opredeliti ontološko obveznost za niz programskih agentov. Ontološka obveznost je koncept na ravni znanja in je neodvisna od ravnih simbolov.

Ontološka obveznost je dogovor/pogodba z agenti o uporabi slovarja na način, ki je skladen s teorijo, ki je opisana z ontologijo. Mogoče je trditi, da agent spoštuje ontologijo, če so njegove aktivnosti skladne z opredelitvami v ontologiji. Z opazovanjem agentovih aktivnosti se agentu pripiše posedovanje znanja. Agent »ve«, če se vede, kot da bi imel informacijo o znanju. Aktivnosti agentov, vključno s strežniki zbirk znanj in strežniki zbirk podatkov, je mogoče opazovati s posredovanjem vmesnika »povej in vprašaj«, preko katerega odjemalec sodeluje z agentom, in sicer preko postavljanja logičnih trditev in postavljanja logičnih vprašanj (Levesque, 1984).





Slika 12: Splošna ontologija predstavitev znanja (povzeto po Sowa 2000a)

Zbirka znanja ni treba porazdeljevati med agente, ki si delijo porazdeljeni slovar, opredeljen z ontologijo. Vsak agent ima določena lastna znanja, za katera drugi agenti ne vedo. Agent, ki je zavezan ontologiji, ni obvezan odgovoriti na vsa vprašanja, ki jih je mogoče postaviti v okviru porazdeljenega slovarja. Predanost ontologiji, glede na povpraševanja in trditve, izvedene z uporabo slovarja, opredeljenega v ontologiji, je porok skladnosti, ne pa popolnosti. Ontologija je tesno povezana s teorijo logike.

#### 4.4 ONTOLOGIJA V GEOINFORMATIKI

Beseda ontologija se vse pogosteje pojavlja na strokovnem področju geoinformacijske znanosti. Ontologija se uveljavlja kot sredstvo za izboljšanje dostopa in porazdeljevanja obstoječih geografskih informacijskih virov. Neformalno se ontologije že dolgo uporabljajo v obliki objektnih oziroma atributnih katalogov. Vendar obstoječe ontologije geografskih informacijskih sistemov vsebujejo pogled na svet, ki se osredotoča na obstoječe vsebine zbirk podatkov in ne upošte-

vajo povezav s človeškimi aktivnostmi (Gallagher in Carnahan, 2000; Harvey et al., 1999). Guarino (1998) ločuje ontologijo v filozofiji ter ontologijo v inženirstvu znanja. V inženirstvu znanja se z ontologijo obvladuje jezikovno izraženo znanje. Ontološko raziskovanje presega konceptualno modeliranje. Do nedavnega ni bilo povezave med raziskavami ontologije oziroma epistemologije prostorskih objektov geografske razsežnosti ter med raziskavami predstavitev prostorskih pojavov (Winter, 2001). Z ontologijo kot znanstveno disciplino se na nepristranski način proučujejo različni tipi objektov in njihovih povezav v enotno in celotno stvarnost objektov, lastnosti, razmerij in procesov, ki sestavljajo svet različnih ravni in podrobnosti in katerih različni deli so predmet raziskav različnih ljudskih in znanstvenih disciplin. V člankih (Casati et al., 1998) so odkrivali ontologijo ljudske, zdravorazumske geoprostorske stvarnosti z uporabo formalnih logičnih metod.

Za GIS je potrebno razviti metode transformacije kvantitativnih prostorskih podatkov v vrsto kvalitativnih predstavitev prostorskih pojavov geografskih razsežnosti, ki so prilagodljive razumevanju GIS-izvedencev in laikov. Za osnovo tem metodam lahko služi le teorija ontologije zdravorazumskega pojmovanja geoprostorske stvarnosti (Smith in Mark, 2001). Rosch (1987) ločuje dve ravni kognitivne klasifikacije: osnovno in napredno. Vsaka družina zdravorazumskih kategorij je hierarhično organizirana v drevo, ki ima na svojem začetku zelo splošne kategorije, v nadaljevanju pa postajajo kategorije specializirane in organizirane po različnih načelih. Na področju človeških aktivnosti je vzročnost posledica želja in hotenj posameznikov (Dowty, 1998). S širjenjem GIS-aplikacij na spletu se povečuje potreba po prepoznavanju pristopov, s katerimi različne skupine uporabnikov konceptualizirajo domeno prostora geografskih razsežnosti. Potrebno je razumeti, kako uporabniki uspejo oziroma zakaj ne uspejo izmenjevati prostorskih podatkov, bodisi v osebni komunikaciji bodisi pri komunikaciji z računalniki. V bodočem razvoju geografskih informacijskih sistemov bo osrednji del predstavljala skladna in empirično utemeljena ontologija prostorskih pojavov geografskih razsežnosti (Smith in Mark, 2001). Podatki, ki predstavljajo iste pojave stvarnega sveta, so lahko shranjeni v neodvisnih zbirkah podatkov. Različne zbirke podatkov tako predstavljajo različne konceptualizacije prostora geografske razsežnosti. Medopravilni GIS bi moral biti neodvisen od razlik v geoinformacijski tehnologiji, poleg tega pa bi moral tudi nazorno komunicirati na visoki semantični ravni. Zato je potrebno izdelati ontologije geoinformacijske infrastrukture oziroma ontologije prostorskih informacijskih sistemov. Za povezovanje informacij je potrebno razviti metode, ki omogočajo razrešitev terminoloških in konceptualnih neskladij. Prvi poskusi razvoja primernih metod so temeljili na posameznih primerih zbirk podatkov. Pri tem je postalo očitno, da je potrebno razviti standardno ontologijo najvišje ravni.

Izgradnja ontologij o geografskih informacijah je zapleten postopek (Frank, 1999). Povezan je s porazdeljevanjem znanja o prostoru geografskih razsežnosti in večkratno uporabo znanja (Bucher, 2000). Pri porazdeljevanju znanja je smiselno združevanje ontologij in znanja o reševanju problemov (Gomez in Benjamins, 1999). Znanje o obstoju geografskih informacij je obsežno. Za integracijo tega znanja je potreben metapodatkovni sistem, ki opisuje dosegljive, a raznovrstne geografske informacijske vire (Bucher, 2000).

Konceptualna povpraševanja v zbirkah prostorskih podatkov so po naravi hierarhična. Z vključitvijo različnega števila sinonimov v povpraševanje je mogoče zožiti ali razširiti povpraševanje v okvir hierarhije konceptualnih vprašanj. Kakovost odkrivanja konceptov v zbirkah prostorskih podatkov je odvisna od kakovosti ontologije oziroma zbirke znanja v slovarjih ali besednjakih. Splošni slovarji obravnavajo koncepte splošnega sveta s celostnega vidika, na primer WordNet (Frank, 1999). Vsebujejo besedne sinonime in osnovna semantična razmerja med njimi. Problemi izražanja istih konceptov z različnimi naravnimi, nacionalnimi jeziki se

rešujejo s povezovanjem konceptov v okviru ontologije formaliziranih semantičnih različnosti. Banares (2000) predlaga uvrstitev različnih geoprostorskih ključnih besed iz različnih virov, kot so besednjaki, slovarji, leksikoni, enciklopedije in drugi ustrezni viri, v relacijsko tabelo. Sistem integracije zbirki podatkov vsebuje poleg temeljne ontologije še ontologijo tipoloških elementov geoinformacijske infrastrukture. Becam (2000) za izdelavo integracijskega sistema predlaga metapodatkovni model, katerega opis temelji na elementih več ontologij. Temeljna ontologija obsega vse ravni, od posamezne entitete do kategorije najvišje ravni domene. Prostorski referenčni sistemi omogočajo uporabnikom geografskih informacijskih sistemov integracijo prostorskih podatkov glede na njihovo geometrijo, ne pa glede na njihovo semantiko. Po analogiji je za semantično integracijo prostorskih podatkov mogoče opredeliti in izdelati semantični referenčni sistem. Tako sta mogoča elementa semantičnega referenčnega sistema semantični datum in semantično referenčno omrežje. Semantični datum opredeljuje pomen osnovnih izrazov zunaj sistema, semantično referenčno omrežje pa deluje kot koordinatni sistem, ki formalno opredeljuje pomen izrazov. S funkcionalnega vidika je potrebno opredeliti semantično referenciranje kot osnovni postopek vzpostavljanja razmerij med izrazi podatkovnega modela ter semantičnim referenčnim omrežjem. Semantična projekcija omogoča kartiranje iz bolj kompleksnih v bolj enostavne podatkovne modele. Semantična translacija pa je opredeljena kot matematični proces transformacije iz enega v drug semantični referenčni sistem (Kuhn, Raubal, 2003).

#### 4.4.1 Aktivnosti v prostoru geografske razsežnosti

GIS bi moral nuditi podporo izvajanju človeških aktivnosti. Namesto tega je GIS v večini primerov razvit kot pasivni model sveta s premajhnim upoštevanjem okolja, v katerem se uporablja. Napačno se predpostavlja, da postanejo geografski podatki uporabni sami po sebi, če so le dovolj verodostojna predstavitev stvarnosti. Za izboljšanje uporabnosti računalniško zasnovanih GIS-sistemov je potrebno razvoj osredotočiti na človeške aktivnosti (Laurel, 1993). Da bi postale geografske informacije bolj koristne in uporabne, bi bilo potrebno razviti metodo za gradnjo ontologije z osredotočenjem na človeške aktivnosti v prostoru (Kuhn, 2001). Metoda mora: temeljiti na človeških aktivnostih; uporabljati zapisana besedila, ki opisujejo človeške aktivnosti; povezati aktivnosti z objekti domene.

Obstoječe ontološke teorije geografskih informacijskih sistemov so statične. Poudarjajo objekte in attribute, njim podrejena pa so razmerja in operacije. Vzroki za pristranost v ontologijah in podatkovnih modelih geografskih in drugih prostorskih informacijskih sistemov so sledeči: izvor GIS-tehnologije je v statičnih, kartografskih modelih sveta; poudarek GIS-tehnologije je na objektih in atributih; neuporaba na logiki temelječih formalnih jezikov za upravljanje operacij in njihove semantike; v teorijah človeškega prostorskega zaznavanja je objektom vnaprej podeljena prioriteta; pomanjkljivo razumevanje o tem, kako so aktivnosti predstavljene v naravnih jezikih s prepoznano pristranostjo v prid samostalnikom (Fellbaum, 1999).

Potrebno je prepoznati, kako določeni agenti delujejo v svetu obravnave. Prepoznati je potrebno aktivnosti, ki se pojavljajo v svetu, neodvisno od načina modeliranja človeškega znanja o svetu v informacijskem sistemu. Teorije domen, ki nimajo povezave z aktivnostmi, nimajo pomena. Modeli aktivnosti brez znanja o domeni so neuporabni. Prostor geografskih razsežnosti je potrebno obravnavati sočasno kot sistem objektov in aktivnosti. Pri izdelavi GIS-ov za podporo aktivnostim v prostoru je potrebno – pri izbiranju in formalizaciji objektov – primarno upoštevati vidik aktivnosti v prostoru (Camara et al., 2000). Teorija aktivnosti (Kaptelinin et

al., 1999) vsebuje sledeča načela: interakcija s svetom oblikuje človeško konceptualizacijo sveta; aktivnosti so usmerjene k objektom (stvarem ali živim bitjem); aktivnosti so hierarhično strukturirane; aktivnosti pogojujejo odvisnost pomena informacij od konteksta; nova orodja in izdelki razvijajo nove človeške aktivnosti.

Hierarhija aktivnosti sega od visokonivojskih aktivnosti, kot sta hotenje in motivi, preko ciljno usmerjenih akcij do posameznih operacij, ki so sestavni del akcij. Vsaka od navedenih ravni lahko vsebuje vrsto slojev različnih stopenj kompleksnosti. Pri konceptualizaciji sveta in človeških aktivnosti v svetu ni potrebno ločevati med tehničnimi stvarmi in naravnimi objekti. Potrebno je enakovredno obravnavati interakcije vseh delov prostora geografskih razsežnosti.

Gibson (1986) je za uravnovešeno obravnavo objektov in aktivnosti vpeljal koncept omogočanja. Ljudje razlikujejo med stvarmi v svetu, ki omogočajo aktivnosti. Logični začetek razvoja ontologije prostora je analiza aktivnosti domene. Takšen pristop se ne uporablja pogosto zato, ker se koncepte a priori enači z objekti. V splošnem se ontologi sicer strinjajo, da obravnavajo objekte, razmerja, stanja, dogodke in postopke. V praksi pa navajajo zgolj razrede z atributi, ki predstavljajo določena stanja in razmerja. Dogodke, procese in aktivnosti prepuščajo poznejši, ločeni, največkrat nikoli izvedeni obravnavi. Obravnavo aktivnosti je ključni element modeliranja v kontekstu prostora.

Kuhn (2001) za razvijanje ontologij v domeni prostora geografskih razsežnosti predlaga analizo tekstov, zapisanih v naravnih jezikih, ki opisujejo človeške aktivnosti. Prepoznavanje konceptov, ki naj bi bili vključeni v ontologijo domene, je mogoče izvajati z analiziranjem tekstualnih opisov aktivnosti. Opisi zakonov so pogosto dober vir znanj o mnogih domenah, kjer se uporabljajo prostorske informacije. Zaradi pravnih in administrativnih pravil zakoni vsebujejo mnogo opisov prostorskih aktivnosti. Pri odkrivanju aktivnosti v besedilih se je potrebno osredotočiti na iskanje glagolov in glagolnikov ter spajanje sinonimov ter različnih slovničnih oblik v enotno formo. Sledi uvrščanje odkritih aktivnosti v hierarhično strukturo. Aktivnosti imajo samosvojo, naravno pripadajočo semantiko, ki omogoča urejanje, ki presega modeliranje nadobjektov in podobjektov. V naslednjem koraku izdelave ontologije se odkrivajo objekti, ki omogočajo aktivnosti v prostoru. Ko so objekti kategorizirani pri upoštevanju aktivnosti, se atributi objektov uporabijo za nadaljnje razvrščanje.

Ontologije so pogosto izvedene s formalizacijo besed naravnih jezikov. Posebej zapleteno stanje v informacijskih sistemih nastane, ko mora informacijski sistem služiti potrebam uporabnikov, ki govorijo različne naravne jezike. Mark et al. (2003) je analiziral geografske in prostorske izraze različnih slovarjev dveh naravnih jezikov. Ugotovil je, da besede osnovne ravni ne ustrezajo istim konceptom. Benslimane et al. (2000) je predlagal opredelitev omrežja večnivojske ontologije. Ta je sestavljena iz ontoloških ravni, ki so sestavljene iz generične funkcionalne strukture in več ontologij domen. Generična funkcionalna struktura vključuje splošne ontološke koncepte, ki so opisani kot abstraktni tipi.

Meta podatke je potrebno obravnavati kot podatke neodvisne od izvornih podatkovnih nizov. Vendar mora biti vsak podatkovni niz opremljen z meta podatkovnim nizom (ontologijo), ki opredeljuje pomen izvornih podatkov (Giger, Najar, 2003).

# 5

---

## ONTOLOGIJA PROSTORA GEOGRAFSKIH RAZSEŽNOSTI

Geografski prostor je sestavljen iz enot, ki niso zgolj fizični objekti, ampak v pomembnem obsegu ustvarjeni v človeškem umu (Smith in Mark, 1998). Predstavljena je ontologija prostora geografskih razsežnosti (OPGR), na osnovi katere je mogoče izvajati pomensko integracijo podatkov zbirk prostorskih podatkov. OPGR vsebuje semantični slovar prostorskih izrazov. Pri izdelavi modela konceptualnega omrežja so upoštevani izsledki raziskav ontološkega modeliranja. V nadaljevanju poglavja so izvedene opredelitve mer za ocenjevanje semantike zbirk prostorskih podatkov. Izvedena je analiza semantičnih razmerij aplikativnih konceptov v okviru OPGR. V sklepnem delu poglavja so nakazani metode in postopki semantičnega povezovanja zbirk prostorskih podatkov.

Razvita je bila metodologija posrednega povezovanja zbirk prostorskih podatkov. Metoda temelji na posredovalnem sistemu, ki omogoča prenašanje pomena med podatkovnimi zbirkami. Posredovalni sistem je sestavljen iz semantičnega referenčnega sistema ter semantičnih razmerij med semantičnim referenčnim sistemom in katalogi obravnavanih zbirk podatkov. Za razvoj semantičnega referenčnega sistema je bila razvita metodologija izdelave ontologije prostora geografske razsežnosti. Za ocenjevanje semantičnih razmerij pa je bila razvita metoda ocenjevanja semantike zbirk prostorskih podatkov.

Za razvoj metodologije posrednega povezovanja zbirk prostorskih podatkov so bile uporabljene nekatere predhodno razvite metode obravnavanja semantike zbirk podatkov oziroma modifikacije in njihova nadgradnja v novo razvito metodologijo. Razvita metodologija se sestoji iz štirih delov: metode izdelave ontologije prostora geografskih razsežnosti, metode ocenjevanja semantike zbirk prostorskih podatkov, metode ocenjevanja skladnosti med koncepti posameznih katalogov prostorskih podatkov ter metode izvedbe semantične povezave skladnih konceptov zbirk prostorskih podatkov.

### 5.1 PROSTORSKI OBJEKTI GEOGRAFSKEGA OBSEGA

Prostorski objekti geografskih razsežnosti so objekti, ki so večji od človeškega telesa in jih v celoti ni mogoče zaznati v okviru enega zaznavnega dejanja (Egenhofer in Mark, 1995). Geografski prostor je sestavljen iz enot, ki niso zgolj fizični objekti, ampak v pomembnem obsegu ustvarjeni v človeškem umu (Smith in Mark, 1998). V splošnem se ločuje med pred-

metnimi objekti in nepredmetnimi objekti geografskega obsega. Predmetni objekti obstajajo v fizični stvarnosti neodvisno od človeškega uma; primeri predmetnih objektov geografskega obsega so hiše, ceste, mostovi, utrdbe, gore, reke in morja. Nepredmetni objekti so ustvarjeni v človeškem umu; primeri nepredmetnih objektov so pravni objekti, na primer zemljiška posest, politične enote razdelitve prostora (Bittner in Frank, 1999).

Znanje o predmetnih objektih je pridobljeno z opazovanji in meritvami. S tehnikami opazovanj in meritev je mogoče pridobiti vedenje, znanje o približkih lastnosti in lokacij prostorskih objektov. Opazovanja in meritve lokacije predmetnih geografskih objektov podajajo znanje o približku lokacije v okviru lokalnih, regionalnih razdelitev, ki so bile ustvarjene ali izbrane s pomočjo opazovalnega ali merskega postopka (Bittner, 1999). Razlikovati je mogoče dve skupini predmetnih objektov geografskega obsega:

- predmetni objekti/pojavi z določenimi mejami,
- predmetni objekti/pojavi z nedoločenimi mejami.

Pojavi z določenimi mejami imajo ostre meje in lokacijo le-teh je mogoče natančno izmeriti, in sicer v okviru omejitev podrobnosti oziroma občutljivosti meritev. Primeri pojavov z določenimi mejami so hiše, mostovi, ceste. Poožaja pojavov z nerazločnimi, nedoločenimi mejami ni mogoče natančno izmeriti, kajti meje so nerazločne, nedoločene, se premikajo ali jih ni mogoče opazovati. Primeri pojavov z nerazločnimi mejami so reke, gore, otoki. Nepredmetni objekti so konstrukt človeškega uma. Ustvarjeni so z opredelitvami. Ker ljudje ustvarjajo in določajo pravne objekte z nedvoumnimi opredelitvami, je znanje o lastnostih in lokacijah pravnih objektov natančno in popolno. Primer pravnega objekta je zemljiška posest in administrativno-politične enote razdelitve prostora. Za pravne objekte je jasna predstavitev natančne lokacije možna in nujno potrebna, za fizične objekte je lahko poznana in predstavljena le približna lokacija.

### 5.1.1. Koncept lokacije v ontologiji prostora

V zvezi z ontološkim in epistemološkim ujemanjem je mogoče ločevati med formalizacijo in predstavnostjo natančnih in približnih lokacij. Pojem lokacije se nanaša na razmerja med posameznimi enotami in referenčnim okvirom. Pojem lociranja znotraj referenčnega okvira dovoljuje abstrakcijo, ki posamezne individuum nadomesti z ustreznimi razredi. Skupina individuumov, ki ima isto lokacijo znotraj izbranega referenčnega okvira, tvori ustrezní razred. Lokacijska abstrakcija pomeni obravnavo ustreznih razredov posameznih individuumov, ki si delijo isto lokacijo, ne pa obravnavo posameznih individuumov. Razlikovati je mogoče med natančno lokacijo in približno lokacijo.

V primeru natančne lokacije vsebujejo vsi ustrezní razredi iste lokacije samo posameznike, ki so natančno solocirani (Casati in Varazi, 1994; 1995). Dva individuum sta natančno solocirana, če zasedata isti del prostora. Predmetni objekti ne morejo biti solocirani. Dva različna predmetna objekta ne moreta imeti natančno iste lokacije v istem trenutku časa. Solocirani so lahko le objekti, ki so ontološko različni. Za veliko razredov geografskih objektov (gore, doline, področja onesnaženja, Alpe) ni mogoče poznati natančne lokacije, temveč je za takšne objekte mogoče poznati le približno lokacijo. Pojem približne lokacije temelji na referenčnih okvirjih, ki predstavljajo regionalne razdelke obravnavanega prostora. Približna lokacija je opredeljena v smislu razmerij med prostorskimi objekti in razdelki obravnavanega prostora.

## 5.2 PRAVILA ZA MODELIRANJE ONTOLOGIJE PROSTORA GEOGRAFSKIH RAZSEŽNOSTI

Pri razvoju formalne ontologije in njenih kategorij, razredov, je potrebno ohraniti čim manjše število razredov najvišje ravni. To zagotavlja jasnost razumevanja členjenja ontologije in enostavnost njene uporabe. Vsaka uspešna komunikacija zahteva jezik, ki je zgrajen nad jedrom skupnih konceptov. Podrobnejša opredelitev formalne ontologije je: ontologija je členjena, strukturirana, omejevalna zbirka nedvoumno opredeljenih konceptov. Obstajajo štiri temeljna pravila za modeliranje formalne ontologije (Kuhn, 1996): ontologija je zbirka skupnih konceptov; koncepti ontologije so nedvoumno opredeljeni, zbirka konceptov je omejevalna: koncepti, ki niso opredeljeni v ontologiji, niso uporabni; zbirka ima model členitve, kar pomeni, da ontologija vsebuje razmerja med koncepti.

Ontologije je mogoče graditi za tematske vsebine geografskih objektov kakor tudi za njihove razsežnostne koncepte. Ontologije razsežnostnih konceptov prostorskih objektov so geometrija, topologija ter simbologija predstavitve geografskih objektov (Rodriguez et al, 1996). Zbirka skupnih konceptov – ontologija domene – je torišče primerjav aplikacijskih ontologij, ki so izdelane za vsak geografski podatkovni niz. Primer področne zbirke skupnih zasnov je domena topografskega kartiranja, ki predstavlja prvo ontološko raven. Za tovrstno domeno so podane opredelitve topografskih objektov, kot so ceste, železnice, zgradbe in podobno (Ravi, 1995). Druga raven so aplikacijske ontologije stvarnih geografskih podatkovnih nizov. V teh ontologijah so uporabljene besedne oznake za snemane oziroma kartirane zasnove (cesta, zgradba in podobno), vendar njihov natančni pomen v splošnem ni enak pomenu zasnov iz področne ontologije, v kateri so uporabljene enake ali podobne besedne oznake. Zaradi teh razlik je za vsak podatkovni niz, ki ga obravnavamo v postopku integracije, potrebno izdelati aplikacijske ontologije. Vključevalna pravila predstavljajo kriterije za zajem podatkov o objektih, določajo, kateri objekti bodo posneti. Predstavitvena pravila opredeljujejo, kako bodo objekti predstavljeni. Pravila poenostavljanja določajo, kako bodo stvarni objekti poenostavljeni v predstavitvah objektov. Pravila združevanja opredeljujejo, kako bodo stvarni objekti združevani v predstavitvi objektov.

### 5.2.1 Opredelitev pravil semantične podobnosti

Problem semantične medopravnosti je odkrivanje semantično podobnih objektov, ki pripadajo različnim podatkovnim zbirkam (Rodriguez et al., 1999). Semantična razmerja med razredi ontologije domene in razredi ontologije aplikacije opredeljujejo semantiko obravnavanega področja obravnave. Tako obstajata dve osnovni semantični razmerji (Sheth in Kashyap, 1993): prvo je razmerje enakovrednih razredov ontologije domene in razredov ontologije aplikacije; drugo je razmerje povezave dveh ali več razredov ontologije domene, tako da tvorita oziroma tvorijo v ontologiji aplikacije en sestavljeni agregatni razred.

Ti dve semantični razmerji določata, da mora biti ontologija domene dovolj bogata z zasnovami, da lahko obvlada vse koncepte ontologije aplikacij. Na osnovi dveh predhodno opredeljenih razmerij so opredeljena semantična razmerja med objektnimi razredi različnih podatkovnih nizov (Sheth in Kashyap, 1993). Objektni razredi različnih aplikacijskih ontologij so semantično enakovredni, ekvivalentni, če ustrezajo istim razredom v ontologiji domene. Objektni razredi

različnih aplikacijskih ontologij so semantično sorodni, če ustrezajo razredom v ontologiji domene, ki so drug drugemu podrazred ali nadrazred. Objektna razreda A in B iz različnih aplikacijskih ontologij sta semantično ustrezna, če je razred A enakovreden z razredom v ontologiji domene, ki predstavlja en del sestavljenega razreda B v ontologiji aplikacije. Semantična vsebina zbirke podatkov je povečana, če se pri modeliranju zbirke prostorskih podatkov uporabljajo koncepti višjih ravni za modeliranje kompleksnih prostorskih objektov, ki so podobni konceptom, s katerimi ljudje obravnavajo geografsko stvarnost. Druga razsežnost povečane semantičnosti zbirke podatkov je modeliranje razmerij med geografskimi pojavi v okviru zbirke podatkov, s čimer se zagotavlja njena pravilnost (Kempainen, 1995).

### 5.3 OPREDELITEV ONTOLOGIJE PROSTORA GEOGRAFSKIH RAZSEŽNOSTI

Da bi postale geografske informacije bolj koristne in uporabne, je bilo potrebno razviti metodo za gradnjo ontologije z osredotočenjem na človeške aktivnosti v prostoru. Metoda temelji na človeških aktivnostih, povezuje aktivnosti z objekti domene. Oblikovana je hierarhija konceptov. Pri oblikovanju se uporablja zdravorazumsko mišljenje, uporabljajo se zapisana besedila, ki opisujejo človeške aktivnosti. Konceptualna povpraševanja v zbirkah prostorskih podatkov so po naravi hierarhična (Frank, 1999). Rosch (1987) ločuje dve ravni kognitivne klasifikacije: osnovno in napredno. Vsaka družina zdravorazumskih kategorij je hierarhično organizirana v drevo, ki ima na svojem začetku zelo splošne kategorije, ki v nadaljevanju postajajo specializirane in organizirane po različnih načelih. Na področju človeških aktivnosti je vzročnost posledica želja in hotenj posameznikov (Dowty, 1998). V metodologiji je predpostavljena hierarhična razdelitev v sledeče ravni:

- raven 0: izhodišče ontologije,
- raven 1: inducirana kategorija,
- raven 2: superrazred,
- raven 3: razred,
- raven 4: objekt/shema (pojav),
- raven 5: atribut (objekta/sheme pojava).

Določene razlike človeških kognitivnih sistemov so posplošene tako, da je mišljenje splošno veljavno. Splošno, zdravorazumsko mišljenje torej obstaja neodvisno od različnosti človeških kognitivnih aktivnosti (Smith, 1992). Casati et al. (1998) je odkrival ontologijo ljudske, zdravorazumske geoprostorske stvarnosti z uporabo formalnih logičnih metod. Za GIS je bilo potrebno razviti metode transformacije kvantitativnih prostorskih podatkov v vrsto kvalitativnih predstavitev prostorskih pojavov geografskih razsežnosti, ki so prilagodljive razumevanju GIS izvedencev in laikov. Za osnovo tem metodam lahko služi le teorija ontologije zdravorazumskega pojmovanja geoprostorske stvarnosti (Smith in Mark, 2001). Metodo izdelave ontologije prostora geografskih razsežnosti je zaradi potrebe po hierarhičnosti modela, ki povečuje enoličnost strukture in hitrost iskanja konceptov, razdeljena v tri faze: določitev izhodišča ontologije, izdelava inducirane ravni ontologije ter stvarnih ravni ontologije prostora.

Z metodo, opisano v nadaljevanju, je izdelano konceptualno omrežje domene prostora geografskih razsežnosti v naravnem jeziku kakor tudi v obliki formalno opredeljenih izrazov



na osnovi ontologije konceptov predstavitve znanja privzeto po Sowa (2000b). Najprej so predstavljeni temeljni koncepti predstavitev znanja, primerni za opredelitev znanja, simbolov prostora. Univerzalna ontologija predstavitve znanja o prostoru geografskih razsežnosti je zasnovana kot sistem človeških aktivnosti (namenov) in prostorskih pojav (pojavnih oblik - shem), ki so vzrok (pojav) ali posledica (objekt) aktivnosti.

### 5.3.1 Določitev izhodišča ontologije prostora

Za določitev izhodišča ontologije prostora je določena omejitev koncepta »prostor« s parametrom razsežnost. Z obravnavano ontologijo je za željeno obravnavati objekte geografskega prostora, zato je parameter razsežnosti omejen na geografsko razsežnosti, ki sta jo opredelila Egenhofer in Mark (1995): »Prostorski objekti geografskega obsega (razsežnosti) so objekti, ki so večji od človeškega telesa in jih ni mogoče v celoti zaznati v okviru enega zaznavnega dejanja.« Izhodišču ontologije je določena oznaka ravni 0.

### 5.3.2 Izdelava inducirane ravni ontologije prostora

Pristop induciranja konceptov je privzet po metodologiji konstruktivne indukcije, ki je metoda klasifikacije v induktivnem logičnem programiranju. Uporablja se pri obravnavi nekaterih vrst problemov, kjer se je pokazalo, da obstoječi atributi in relacije ne zadoščajo za razumljiv opis danega koncepta. Problemi takšne vrste se rešujejo s samodejnim ali ročnim dodajanjem novih, vmesnih konceptov (Robnik-Šikonja, 1997).

Z izdelavo inducirane ravni ontologije prostora so povezani deli splošne ontologije predstavitve znanja, ki ga je razvil Sowa (2000a). Tri izbrane kategorije so formalno zapisane v obliki trditve [1], [2] in [3]. Aktivnosti ljudi v prostoru geografskih razsežnosti so opredeljene s kategorijo »namen«. Namen je hotenje, ki pojasnjuje stanje, situacijo:

$$\exists \text{Namen} = \text{pojmovno} \wedge \text{posredovalno} \wedge \text{pojavno} \quad [1]$$

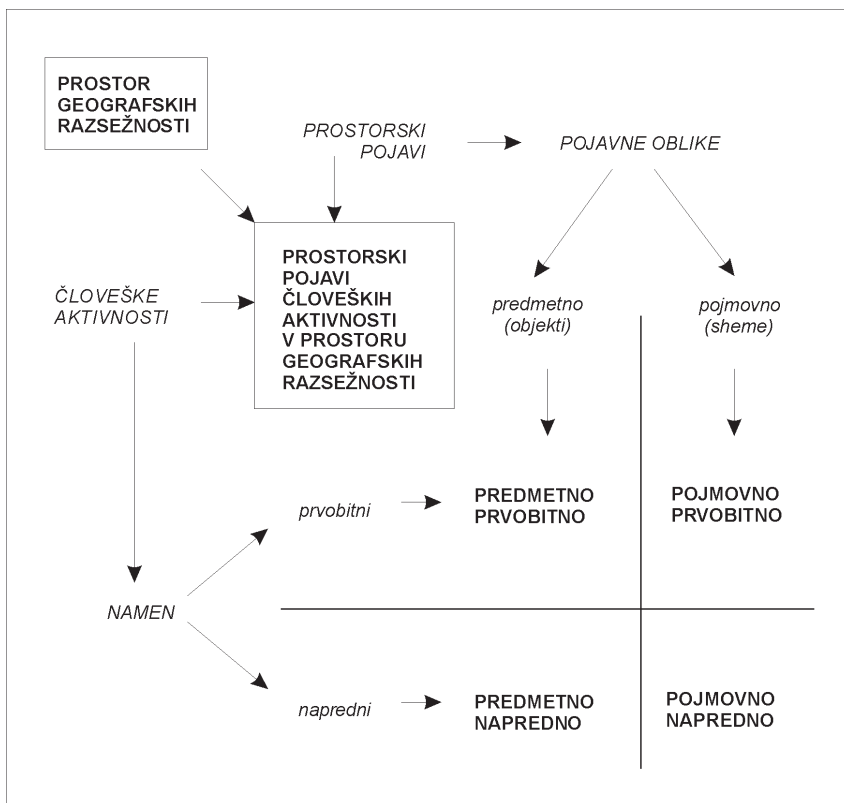
V ontologiji znanja (Sowa, 2000a) trditve povezujejo namen z aktivnostmi in povzročitelji, agenti. Trditve za namen in aktivnosti lahko vključujejo parametre, kot so časovno zaporedje, naključnost in uspeh ali neuspeh. Nameni človeških aktivnosti v prostoru geografskih razsežnosti se posredujejo v prostor geografskih razsežnosti preko dveh skupin dojetanja udeležbe agentov, in sicer kot prvobitne in napredne aktivnosti. Prvobitne človeške aktivnosti imajo svoj temeljni vzrok v namenu zadovoljevanja posameznikovih, individualnih hotenj in potreb. Napredne človeške aktivnosti imajo svoj vzrok v namenu zadovoljevanja kolektivnih hotenj skupnosti. Pojavne oblike v prostoru geografskih razsežnosti je mogoče opredeliti s kategorijama objekt in shema. Posamezna kategorija (objekt, shema) je konjunkcija enostavnejših monadičnih predikatov (Sowa, 2000a):

$$\exists \text{Objekt} = \text{predmetno} \wedge \text{samostojno} \wedge \text{nepretrgano} \dots\dots\dots [2]$$

$$\exists \text{Shema} = \text{pojmovno} \wedge \text{samostojno} \wedge \text{nepretrgano} \dots\dots\dots [3]$$

Trditve [1], [2] in [3] so izpeljane iz splošne ontologije predstavitve znanja (Sowa, 2000).

Nadalje so opredeljene kombinacije kategorij prvobitne in napredne aktivnosti ter kategorij pojavnih oblik objekt in shema. Kombinacije so izražene v matriki kategorij, razsežnosti 2 krat 2, ki predstavlja inducirano raven izdelane ontologije (slika 13 in preglednica 3). Kategorijam inducirane ravni (1) ontologije prostora so določene oznake vsebovanih kategorij. Formalno so zapisane kategorije inducirane ravni kot sledeče trditvene formule:



Slika 13: Konceptualno omrežje prostora geografskih razsežnosti: izhodiščna in inducirana raven

$\exists$  predmetno-prvobitno = predmetni objekti  $\wedge$  prvobitnih človeških aktivnosti  $\wedge$  geografske razsežnosti [4]

$\exists$  predmetno-napredno = predmetni objekti  $\wedge$  naprednih človeških aktivnosti  $\wedge$  geografske razsežnosti [5]

$\exists$  pojmovno-napredno = pojmovne sheme  $\wedge$  naprednih človeških aktivnosti  $\wedge$  geografske razsežnosti [6]

$\exists$  pojmovno-prvobitno = pojmovne sheme  $\wedge$  prvobitnih človeških aktivnosti  $\wedge$  geografske razsežnosti [7]

Tako so na formalni način ustvarjene kategorije inducirane ravni ontologije prostora geografskih razsežnosti.

### 5.3.3 Predpostavke modela

Konceptualno omrežje prostora temelji na človeških aktivnostih, namenih v geografskem prostoru. Kriterij za opredeljevanje aktivnosti je prostorska razsežnost – razpoložljivost prostora geografskih razsežnosti za izvajanje aktivnosti. Pojavna oblika človeške aktivnosti je prostorski objekt ali pojmovna shema prostorskih razsežnosti. Za aktivnosti modeliranja prostora na analogni ali digitalni medij z računalnikom ali drugimi sredstvi, na primer kartografskimi ali geodetskimi, prostor geografskih razsežnosti ni potreben. Za aktivnost modeliranja in napovedovanja je potreben miselni prostor v umu, na papirju ali v računalniškem pomnilniku in procesorju, ne pa prostor geografskih razsežnosti. Za aktivnosti opazovanja prostora je potreben prostor geografskih razsežnosti.

Kriterij za opredelitev novega temeljnega koncepta aktivnosti v ontologiji je razlika v namenu med človeškimi aktivnostmi, s katerih izvajanjem se ustvari prostorski predmet – objekt – oziroma pojmovna shema prostora. Kriterij za opredelitev nove podaktivnosti (razreda, podrazreda) je razlika v namenu izvedbe aktivnosti ali razlika v vrsti pojavnih oblike.

Zapletene objekte in sheme človeških aktivnosti v geografskem prostoru je mogoče predstaviti s kombiniranjem enostavnih kategorij objektov/shem – navedenih v OPRG. Nameni človeških aktivnosti in pojavnih oblike se prekrivajo v prostoru in času. Konceptualno omrežje prostora geografskih razsežnosti je zasnovano kot večplastna sestava domen raznovrstnih drugotnih prostorskih konceptov, zgrajenih nad osnovo homogene prostorske nepretrganosti nižje ravni kot domene prostorskih objektov in shem (Brentano, 1988).

Simboli, izrazi, ki se standardno uporabljajo za raznovrstne prostorske objekte, so atributi prostorskih namenov oziroma aktivnosti te ontologije (na primer hiša, stanovanje, lokal, parcela). Terminologija pod-objekt, nad-objekt je uporabljena le v smislu semantične primerjave med objekti. Ustvarjena ontologija ima hierarhično zasnovano. Zanj je značilna neodvisnost od podatkovnih objektov kakor tudi od stvarnih prostorskih objektov, ker je obravnavana ontologija opredeljena z nameni (prvobitni, napredni) oziroma aktivnostmi v prostoru ter značilnostmi njihovih pojavnih oblik, ki so lahko predmetne ali pojmovne. Primeri v hierarhiji modela so simboli, s katerimi se označujejo koncepti stvarnosti. Univerzalni obseg modela obravnava celotno domeno prostora geografskih razsežnosti. Mogoče je dinamično razširjanje modela z opredeljevanjem novih simbolov. Izražanje zapletenih konceptov se izvaja s kombiniranjem osnovnih konceptov. Pojavne oblike v prostoru geografskih razsežnosti je mogoče opredeliti s kategorijama objekt in shema. Pojavne oblike v prostoru geografskih razsežnosti so stanja v prostoru, ki jih akterji, agenti, dojemajo kot predmetne pojavnih oblike – objekte – in pojmovne pojavnih oblike – sheme. Obstajajo prostorski predmetni objekti kot pojavnih oblike v prostoru geografskih razsežnosti, za katere se štejejo le objekti človeškega izvora – ustvarjeni objekti. Obstajajo tudi prostorske pojmovne sheme kot pojavnih oblike v prostoru geografskih razsežnosti, za katere se štejejo dogovorno določene – fiat meje – in opazovane pojavnih oblike naravnih – bona-fide – in

Šifra	Prva raven
1	predmetni objekti prvobitnih človeških aktivnosti
2	predmetni objekti naprednih človeških aktivnosti
3	pojmovne sheme naprednih človeških aktivnosti
4	pojmovne sheme prvobitnih človeških aktivnosti

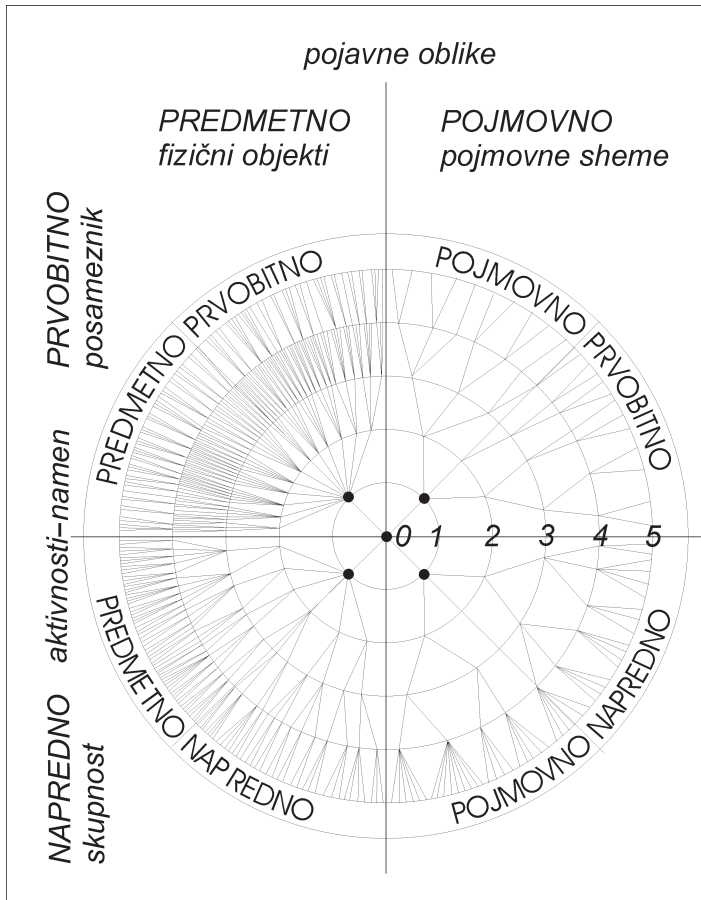
Preglednica 3: Ontologija prostora geografskih razsežnosti – prva raven

socioloških pojavov.

V preglednicah so navedeni koncept prve, inducirane ravni (preglednica 3) in druge ravni (preglednica 4) Ontologije prostora geografskih razsežnosti (OPGR).

### 5.3.4 Izdelava stvarnih ravni ontologije prostora

V tretji fazi izdelave ontologije so obravnavane stvarne ravni ontologije prostora. Vnaprej določeni štirje inducirani koncepti so osnova za ustvarjanje stvarnih konceptov na hierarhični ravni z oznako 2. Na tej ravni se izvede razvejanje ontologije. Da bi se ohranila preglednost, je razvejanost na drugi ravni omejena na 12 konceptov za vsako od štirih skupin, izhajajočih iz induciranih konceptov, tako je največji teoretični obseg druge ravni 48 konceptov. V razviti metodologiji je potrebno upoštevati načelo »od manj podrobnega k bolj podrobnemu«, ki bi ga lahko primerjali s strategijo enkapsulacije objektov v razrede v domeni objektnega modeliranja. Za oblikovanje novih konceptov je potrebno izvajati miselni postopek naravnega učenja, in sicer učenja z razumevanjem in vpogledom (Bratko et al., 1998), podrobneje s strategijo iskanja v



Slika 14: Semantično referenčno omrežje, ontologija prostora geografskih razsežnosti (OPGR)

## ONTOLOGIJA PROSTORA GEOGRAFSKIH RAZSEŽNOSTI

širino ter raziskovanja znanja kot strategijo iskanja v globino (Kononenko, 1997). Pri strategiji učenja iskanje v globino človek zaporedno predstavlja hipoteze in jih preskuša in spreminja, dokler je to možno, potem pa se vrne na izhodišče in postavi novo hipotezo. Tako učenje je za ljudi lahko, vendar ne vodi nujno k optimalni hipotezi in ponavadi zahteva veliko učnih primerov. S strategijo iskanja v širino in globino so bili na drugi ravni ustvarjeni splošni, stvarni in razumljivi koncepti, ki so nadalje razvejani na nižjih ravneh ontologije (slika 14). Osnovno vodilo za ustvar-

Šifra	Prva raven	Druga raven
	<b>predmetni objekti</b>	
1	<i>prvobitnih</i> človeških aktivnosti	
101		objekti aktivnosti BIVANJA
102		objekti aktivnosti PRIDOBIVANJA surovin in energije
103		objekti aktivnosti PROIZVODNJE izdelkov (industrijsko/obratno)
104		objekti aktivnosti PRENOSA (TRANSPORTA)
105		objekti aktivnosti IZMENJAVE proizvedenih izdelkov in urbanih storitev
106		objekti aktivnosti ODLAGANJA ostankov
107		objekti aktivnosti SPROSTITVE na prostem
108		objekti aktivnosti MANIPULACIJE na prostem
109		objekti aktivnosti PARKIRANJA prevoznih sredstev
110		objekti aktivnosti POSTAVLJANJA mejnih predmetov
111		objekti aktivnosti OSVETLJEVANJA površin in spomenikov
112		objekti aktivnosti VERSKIH OBREDOV
2	<i>naprednih</i> človeških aktivnosti	
201		objekti aktivnosti JAVNEGA UPRAVLJANJA (državno/lokalno) socialne storitve, združenja
202		objekti aktivnosti IZOBRAŽEVANJA
203		objekti aktivnosti ZDRAVSTVENE OSKRBE
204		objekti aktivnosti OBRAMBE države in ZAŠČITE ljudi in sredstev
205		objekti aktivnosti NADZORA PREHAJANJA DRŽAVNE MEJE
206		objekti aktivnosti SHRANJEVANJA narodnih rezerv
207		objekti aktivnosti IZMER državnih mrež
208		objekti aktivnosti REGULACIJE naravnih pojavov
209		objekti aktivnosti ODDAJANJA informacij skupnosti
	<b>pojmovne sheme</b>	
3	<i>naprednih</i> človeških aktivnosti	
301		sheme aktivnosti UPRAVLJANJA PRAVIC v prostoru
302		sheme aktivnosti VAROVANJA naravnih virov in spomenikov
303		sheme aktivnosti PROSTORSKEGA PLANIRANJA
304		sheme aktivnosti POPISOVANJA PREBIVALSTVA
305		sheme aktivnosti OBDAVČEVANJA
306		sheme aktivnosti RAZDELEJEVANJA DENARNIH PODPOR
307		sheme aktivnosti DOLOČANJA MEJA ADMINISTRATIVNIH STATUSOV V PROSTORU
4	<i>prvobitnih</i> človeških aktivnosti	
401		sheme aktivnosti OPAZOVANJA NARAVNIH PREDMETNIH POJAVOV geografskih razsežnosti
402		sheme aktivnosti VZPOSTAVLJANJA lastnih imen (shem) in geometričnih shem
403		sheme aktivnosti OPAZOVANJA POJAVOV V ČLOVEŠKI DRUŽBI pojavov geografskih razsežnosti

Preglednica 4: Ontologija prostora geografskih razsežnosti – prva in druga raven.

janje novih konceptov so aktivnosti ljudi v prostoru geografskih razsežnosti, katerih posledica je stvaritev predmetnih objektov ali pojmovnih shem geografske razsežnosti.

Koncepti druge ravni dedujejo lastnosti koncepta, ki predstavlja izbrano inducirano kategorijo (na ravni 1). Dodatno lastnost vsakemu konceptu druge ravni določa obravnavana aktivnost. Koncepti ravni 2 so formalno izraženi s formulo [8]:

$$\text{Koncept (superrazred)} = \text{pojavná oblika} \wedge \text{namen aktivnosti} \wedge \text{aktivnost} \wedge \wedge \text{razsežnost} \wedge \text{pomenska podrobnost} \quad [8]$$

Primer formalnega zapisa koncepta :

Koncept (POSTAVLJANJE mejnih predmetov) = Predmetni objekti  $\wedge$  prvobitne aktivnosti  $\wedge$  postavljanja mejnih predmetov  $\wedge$  geografske razsežnosti  $\wedge$  superazred

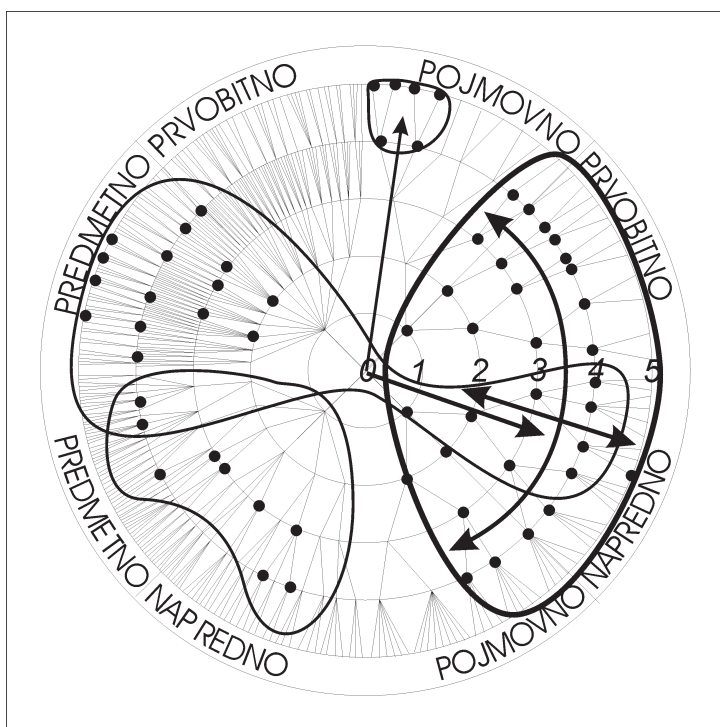
Tako izraženi koncept predstavlja odločevalno pravilo za uvrščanje konceptov na tretjo raven ontologije prostora. Sestavljanje govorečih šifer na ravni 2: enomesna šifra pripadajočega koncepta inducirane ravni (1) in k njej dodana dvomesna šifra koncepta druge ravni (10). Primer govoreče šifre za predhodno predstavljeni koncept (POSTAVLJANJE mejnih predmetov): 1\_10 = 110.

Simboli konceptov četrte in pete ravni so semantični slovar izrazov prostora geografske razsežnosti. Za koncepte, za katere obstajajo istopomenski izrazi, se dodaja povezava na njihove sinonimne simbole. Na enak način se izvaja tudi označevanje konceptov z izrazi v tujih jezikih. To omogoča uporabo izdelane ontologije za povezovanje konceptov zbirk podatkov, izraženih v različnih nacionalnih jezikih. Ontologija prostora geografskih razsežnosti obsega 588 konceptov na petih ravneh in je razširljiva. Konceptualno ogrodje je osnova za organizacijo semantičnega slovarja. Formalni zapis OPGR je bil izveden le za najvišje ravni ontologije, in sicer v jeziku OIL (glej priloga).

# 6

## OCENJEVANJE SEMANTIKE PROSTORSKIH PODATKOV

Kakovost predstavitve semantike zbirke prostorskih podatkov je mera za ocenjevanje stopnje razvoja določenega informacijskega sistema oziroma aplikacije le-tega v semantičnem smislu (Čeh, 2001). Za namene ocenjevanja semantike zbirk prostorskih podatkov je opredeljen nabor mer za ocenjevanje semantičnih značilnosti zbirk podatkov kot zaključenih celot in njihovih delov (slika 15). Semantična skladnost zbirk prostorskih podatkov je utemeljena kot stopnja podobnosti zbirk prostorskih podatkov glede na dva kriterija: semantično globino zbirke podatkov in semantično razpršenost zbirke podatkov z ozirom na ontologijo domene prostora. Večja ko je stopnja podobnosti med semantičnimi parametri primerjanih zbirk, večja je semantična skladnost primerjanih zbirk podatkov in s tem semantična skladnost posameznih konceptov.



Slika 15: Grafični prikaz semantičnih parametrov zbirk prostorskih podatkov

Podatek o razpršenosti zbirke podatkov je za poizvedovanje v namene integracije zbirk podatkov v okolju porazdeljenih geografskih informacijskih sistemov uporaben meta podatek. Semantična ustreznost ontologije je ocena stopnje primernosti določene ontologije za izražanje konceptov zbirk podatkov. To je relativna mera, s katero se izraža kakovost ontologije glede na stvarne zbirke podatkov. Zato se za izračun uporabijo rezultati izražanja konceptov zbirk podatkov s koncepti ontologije. Vsote ekvivalentnih in agregatnih razredov se povprečijo in izrazijo v odstotkih. Semantična pristranost koncepta je posledica podrejanja metodam zajemanja podatkov pri opredeljevanju konceptov prostora. Še posebej je značilna za koncepte v katalogih prostorskih podatkov, ki so zajeti z metodami daljinskega zaznavanja (satelitskimi in letalskimi). V opredelitvah pristranskih konceptov je značilna uporaba izrazov za opisovanje snovi (pesek, kamen, beton, asfalt, vegetacija), za katere je enostavno interpretirati valovno dolžino zaznanega valovanja. Druge v tem delu opredeljene mere za ocenjevanje semantike zbirk prostorskih podatkov so še:

- semantična globina prostorskega koncepta,
- semantična globina zbirke prostorskih podatkov,
- semantična razdalja dveh konceptov,
- semantična razpršenost zbirke prostorskih podatkov,
- kakovost predstavitve semantike zbirke prostorskih podatkov.

V prihodnosti bodo eksplicitno izražena razmerja med shranjenimi podatki v zbirki podatkov (semantičnih podatkov) postala sestavni del vsakega informacijskega sistema. Zbirke prostorskih podatkov se bodo kakovostno razlikovale tudi po opisu vsebovanih semantičnih podatkov, ker bo semantiko zbirke prostorskih podatkov mogoče predstaviti z različnimi semantičnimi merami. Če bo zbirka prostorskih podatkov med svojimi meta podatki vključevala mere, kot so absolutna semantična globina zbirke prostorskih podatkov, relativna prostranost zbirke podatkov, relativno razsipanje zbirke podatkov, semantična razpršenost zbirke podatkov, bo kakovost semantične predstavitve ocenjena z drugo stopnjo. Če bo zbirka prostorskih podatkov poleg naštetih mer med meta podatki vključevala tudi prevod kataloga v semantično referenčno omrežje, ogrodje (OPGR), bo kakovost semantične predstavitve ocenjena s prvo stopnjo.

V metodi posrednega povezovanja zbirk prostorskih podatkov je treba določiti semantična razmerja med semantičnim referenčnim sistemom, ki služi kot posredovalni sistem, ter katalogi obravnavanih zbirk podatkov. Izdelana je metoda ocenjevanja semantike zbirk prostorskih podatkov. Metoda temelji na semantičnem referenčnem omrežju – ontologiji prostora geografskih razsežnosti (OPGR). Metodologija ocenjevanja semantike zbirk prostorskih podatkov je sestavljena iz zaporednih faz, ki so podrobneje opisane v nadaljevanju:

1. prevedba konceptov kataloga zbirke podatkov v semantično referenčno omrežje,
2. določitev semantičnih razmerij med koncepti zbirke in koncepti omrežja,
3. analiza semantične ustreznosti OPGR za povezovanje zbirk,
4. določitev semantične globine zbirke prostorskih podatkov,
5. določitev semantične razpršenosti zbirke prostorskih podatkov,
6. ocena kakovosti predstavitve semantike zbirke prostorskih podatkov.

S postopkom prevedbe konceptov iz kataloga zbirke podatkov v koncepte semantičnega referenčnega sistema se istočasno izvajata procesa uvrščanja konceptov kakor tudi pomensko bogatenje obravnavane zbirke podatkov.

Po opisanem postopku v preglednici kataloga ustvarimo atribut »pomen v OPGR«. S tem



pridobljenim atributom je vsak od izvirnih konceptov kataloga zbirke podatkov pomensko obogaten. Pravimo, da je koncept pomensko obogaten, ker je z oznako (etiketo) ustreznega koncepta/konceptov iz OPGR pridobil vse pomenske attribute, ki so v OPGR formalno opredeljeni v okviru semantičnega referenčnega sistema, ki smo ga metodološko izvedli v razdelku o metodologiji izdelave ontologije prostora geografskih razsežnosti.

Za določitev semantičnih razmerij med koncepti zbirke in koncepti semantičnega referenčnega sistema smo privzeli in nadgradili opredelitev semantičnih razmerij med razredi ontologije domene in razredi ontologije aplikacije (Sheth in Kashyap, 1993). Obstajata dve osnovni semantični razmerji:

- opredelitev razmerja enakovrednih razredov ontologije domene in razredov ontologije aplikacije:  
se\_nanaša\_na\_enaokvredni\_razred [RazredDomene, RazredAplikacije];
- opredelitev razmerja povezave (dveh ali več) razredov ontologije domene tako, da tvorijo en sestavljen (agregatni) razred v ontologiji aplikacije:  
se\_nanaša\_na\_sestavljene\_razred [RazredDomene, RazredAplikacije]
- Ker je potrebna tudi pomenska analiza konceptov, za katere v ontologiji domene (OPGR) ni mogoče odkriti ustreznega koncepta, smo navedenima razmerjema dodali dodatno razmerje [11]:
- neobstoječe razmerje med razredi ontologije domene in razredi ontologije aplikacije  
se\_nanaša\_na\_nezdružljiv\_razred [RazredDomene, RazredAplikacije]

Za operativno uporabo smo navedena razmerja izrazili kot formule:

ekvivalent = se\_nanaša\_na\_enaokvredni\_razred [RazredDomene, RazredAplikacije]; [9]

agregat = se\_nanaša\_na\_sestavljene\_razred [RazredDomene, RazredAplikacije]; [10]

nezdružljiv = se\_nanaša\_na\_nezdružljiv\_razred [RazredDomene, RazredAplikacije]; [11]

V postopku določitve semantičnih razmerij med koncepti zbirke podatkov in koncepti semantičnega referenčnega sistema uporabimo predhodno v OPGR prevedeni in bogateni katalog obravnavane zbirke podatkov. V katalog uvedemo tri dodatne attribute (ekvivalent, agregat, nezdružljiv), s katerimi označujemo razmerja med koncepti zbirke podatkov in semantičnim referenčnim sistemom.

## 6.1. ANALIZA SEMANTIČNE USTREZNOSTI OPGR ZA POVEZOVANJE

Ker želimo poznati stopnje primernosti razvite ontologije za izražanje konceptov zbirk podatkov, ki jih obravnavamo, smo razvili metodo za izvedbo ocene semantične ustreznosti ontologije. Oceno semantične ustreznosti ontologije smo izvedli kot razmerje med številom uspešno izraženih konceptov zbirk podatkov in številom vseh konceptov zbirke podatkov (normalizacija):

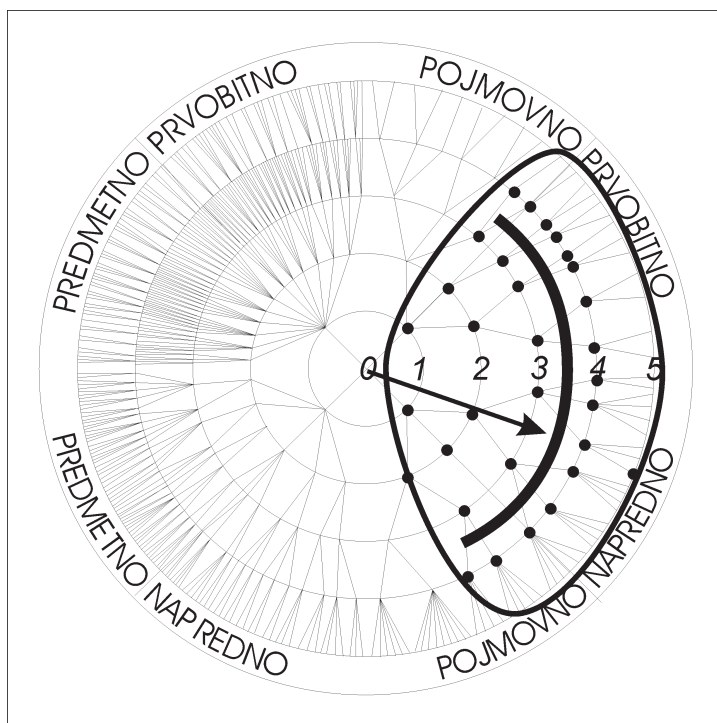
$$\text{semantična ustreznost ontologije} = \frac{\text{število uspešno izraženih konceptov/}}{\text{število vseh konceptov}} \quad [12]$$

Za izvedbo postopka določitve semantične ustreznosti ontologije uporabimo rezultate druge faze metode, in sicer prevedene kataloge obravnavanih zbirk podatkov z določenimi semantičnimi razmerji med koncepti zbirke podatkov in koncepti semantičnega referenčnega sistema (OPGR).

## 6.2 SEMANTIČNA GLOBINA

Za vsako zbirko prostorskih podatkov, ki bi jo radi povezali z drugimi zbirkami prostorskih podatkov, želimo poznati stopnjo tematske, konceptualne podrobnosti oziroma splošnosti, specializacije/generalizacije. Znanje o stopnji tematske podrobnosti zbirk podatkov omogoča primerjavo zbirk podatkov na meta podatkovni ravni, zato smo razvili metodo opisano v nadaljevanju.

### 6.2.1 Določitev absolutne semantične globine



Slika 16: Grafični prikaz semantične globine zbirke prostorskih podatkov

Absolutna semantična globina (slika 16) zbirke prostorskih podatkov (asgpz) je mera, opredeljena kot povprečna globina vseh konceptov zbirke podatkov glede na izhodiščno raven oziroma osrednji koncept ontologije prostora geografskih razsežnosti. Utemeljena je na osnovi izračunov semantičnih globin konceptov (sgk) zbirke podatkov glede na izhodiščno raven 0, ontologije domene prostora geografskih razsežnosti; (slika 17), ( $n$  - število konceptov zbirke podatkov).

$$asgpz = (\sum sgk)/n \quad [13]$$

Za ročno odčitavanje-izračun semantičnih globin konceptov na osnovi govorečih šifer konceptov si pomagamo s pripomočkom - preglednico 5.

Oznake	Izhodišče ontologije	Inducirana raven	Superrazred	Razred	Objekt/shema	Atribut
Oznaka ravni	0	1	2	3	4	5
Razsežnost šifre	-	1	2	2	2	2
Globina	0	1	2	3	4	5

Preglednica 5: Pripomoček za izračun semantičnih globin konceptov na osnovi govorečih šifer.

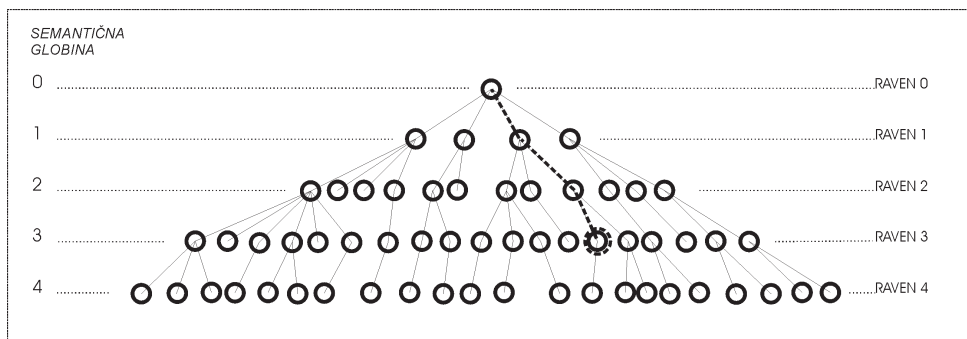
Primeri določanja semantične globine koncepta s pripomočkom (preglednica 5), so navedeni v preglednici 6:

Koncept	Šifra					Globina
hmeljišče	1	02	03	03		4
zemljišča/objekti za pridobivanje rastlinskih surovin	1	02	03			3
ruševje	4	01	07	01	09	5
Globina	1	2	3	4	5	

Preglednica 6: Določanja semantične globine koncepta na osnovi govorečih šifer.

## 6.2.2 Absolutna semantična globina posameznega prostorskega koncepta

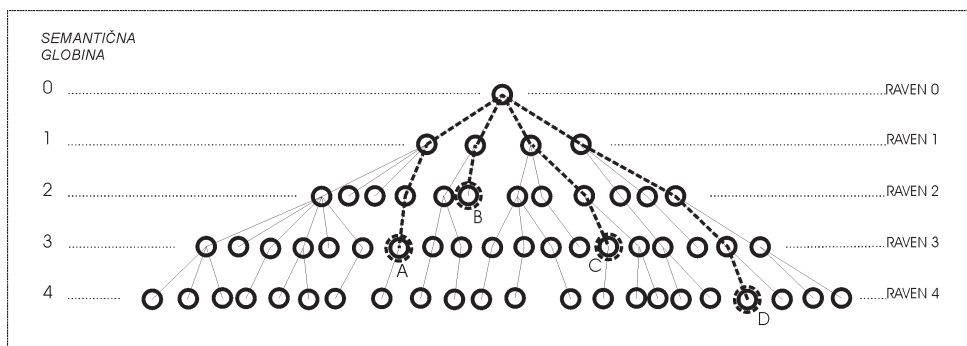
Semantična globina prostorskega koncepta (sgk) je mera, opredeljena glede na izhodiščno raven ontologije prostora geografskih razsežnosti (slika 17). Utemeljena je na osnovi izračuna semantične razdalje med izhodiščno ravno ontologije domene prostora ter obravnavanim prostorskim konceptom. Za semantične razdalje med posameznimi ravni ontologije je normirana vrednost 1 (ena enota).



Slika 17: Prikaz semantične globine koncepta (sgk) v konceptualnem drevesu (grafu). Semantična globina izbranega koncepta je 3

### 6.2.3 Absolutna semantična globina zbirke prostorskih podatkov

Absolutna semantična globina zbirke prostorskih podatkov (asgzp) je mera, opredeljena kot povprečna globina vseh konceptov zbirke podatkov glede na izhodiščno raven oziroma osrednji koncept ontologije prostora geografskih razsežnosti (slika 18). Utemeljena je na osnovi izračuna povprečne semantične globine konceptov (sgk) zbirke podatkov glede na izhodiščno raven ontologije domene prostora.



Slika 18: Prikaz absolutne semantične globine (asgzp) zbirke prostorskih podatkov P {A, B, C, D} v konceptualnem drevesu (grafu). Semantična globina zbirke podatkov je  $(3 + 2 + 3 + 4)/4 = 3$

Koncept, izražen z OPGR		Semantična globina koncepta	Semantična globina zbirke
1020303	hmeljišče	4	3
101	bivanje	2	
10405	tirni	3	
		9	

Preglednica 7: Primer izračuna absolutne semantične globine (asgzp) hipotetične zbirke prostorskih podatkov

Primer izračuna absolutne semantične globine zbirke prostorskih podatkov (preglednica 7): za hipotetično zbirko podatkov je izbran podniz treh konceptov iz kataloga zbirke podatkov rabe kmetijskih zemljišč P1 {hmeljišče, bivanje, tirni}. Za to zbirko podatkov treh elementov je izračunana semantična globina 3.

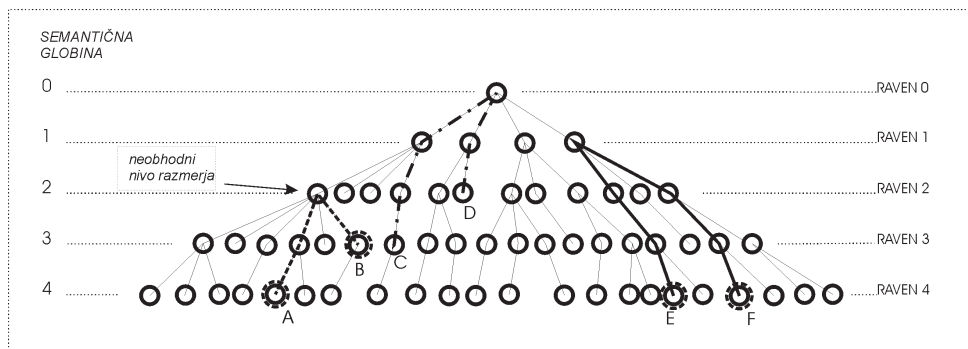
### 6.3 SEMANTIČNE RAZDALJE

Za vsako zbirko prostorskih podatkov, ki bi jo radi povezali z drugimi zbirkami prostorskih podatkov, želimo poznati stopnjo tematske, konceptualne osredotočenosti. Znanje o stopnji tematske osredotočenosti zbirk podatkov omogoča primerjavo zbirk podatkov na meta podatkovni ravni, zato smo razvili metodo, opisano v nadaljevanju. Semantična razpršenost zbirke prostorskih podatkov je mera splošnosti oziroma osredotočenosti, ki jo v razviti metodologiji izračunamo na osnovi semantičnih razdalj med vsemi koncepti obravnavane zbirke podatkov in uteži. Semantične razdalje se morajo izračunati za vse permutacije razmerij med koncepti.

#### Semantična razdalja razmerja

Semantična razdalja semantičnega razmerja med konceptoma je vsota odsekov poti v semantičnem referenčnem sistemu, ki so potrebni za vzpostavitev razmerja med konceptoma. V ta namen je potrebno poznati neobhodno raven v shemi, ki jo je za vzpostavitev razmerja med konceptoma treba doseči v semantičnem referenčnem sistemu.

Neobhodno raven je smiselno izraziti s semantično globino. Semantične razdalje razmerij med koncepti ( $d(A, B) = 3$ ,  $d(C, D) = 5$ ,  $d(E, F) = 6$ ); (slika 19).



Slika 19: Semantične razdalje razmerij med koncepti ( $d(A,B) = 3$ ,  $d(C, D) = 5$ ,  $d(E, F) = 6$ )

Preglednica 8: Določitev globin neobhodnih ravni (gnr) semantičnih razmerij (za vsak par primerjanih konceptov hipotetične zbirke podatkov P1 {hmeljišče, bivanje, tirni})

Globina neobhodne ravni razmerja (gnr)		hmeljišče	bivanje	tirni
		1020303	101	10405
hmeljišče	1020303	x		
bivanje	101	1	x	
tirni	10405	1	1	x

Določitev globin neobhodnih ravni semantičnih razmerij (za vsak par primerjanih konceptov hipotetične zbirke podatkov P1 {hmeljišče, bivanje, tirni}) je izveden v preglednici 8. Vsota semantičnih razdalj zbirke podatkov P1 {hmeljišče, bivanje, tirni} znaša 12 enot (preglednica 9).

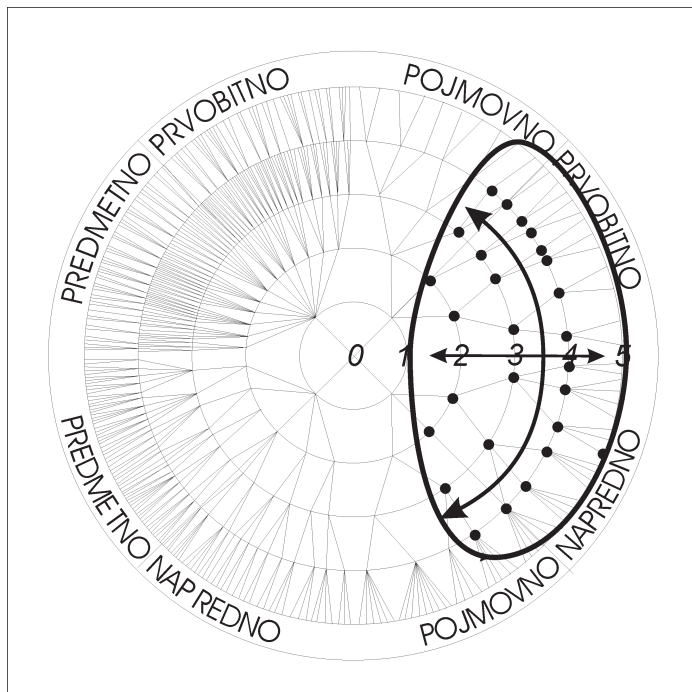
Semantična razdalja (r)		hmeljišče	bivanje	tirni
		1020303	101	10405
hmeljišče	1020303	0		
bivanje	101	1 + 3 = 4	0	
tirni	10405	2 + 3 = 5	2 + 1 = 3	0

Preglednica 9: Izračun semantičnih razdalj (r) semantičnih razmerij (za vsak par primerjanih konceptov hipotetične zbirke podatkov P1 {hmeljišče, bivanje, tirni})

### 6.3.1 Uteži semantičnih razdalj

Vsako izračunano semantično razdaljo med dvema konceptoma je potrebno razviti po metodi utežiti z dvema raznovrstnima utežema.

Utež neobhodne ravni (unr); globlja, ko je neobhodna raven za vzpostavitev povezave med dvema konceptoma, manjša je razsutost zbirke in manjša naj bo utež. Utež naj bo obratno sorazmerna z neobhodno globino, povečano za vrednost 1 (da se izognemo deljenju z 0 na ničelni ravni). Utež izraža prostranost zbirke v okviru ontologije domene (slika 20):



Slika 20: Relativna prostranost in relativna globina zbirke podatkov

$$unr = 1/(gnr + 1) \quad [14]$$

Absolutna prostranost zbirke podatkov (apzp) je vsota semantičnih razdalj konceptov (r) uteženih z utežjo neobhodne ravni (unr):

$$apzp = \sum((r) \cdot (unr)) \quad [15]$$

Relativna prostranost zbirke podatkov (rpzp) je razmerje med absolutno prostranostjo zbirke podatkov [15] in absolutno prostranostjo semantičnega referenčnega sistema - ogrodja (apso):

$$rpzp = apzp/apso \quad [16]$$

Utež relativne globine (urg); manjša kot je povprečna relativna globina (globinska razlika) od obeh konceptov do neobhodne ravni za vzpostavitev povezave med konceptoma, manjša naj bo utež. Utež odraža stopnjo relativnega razsipanja zbirke podatkov po semantični globini. Za vsako izračunano semantično razdaljo med primerjanima konceptoma je treba po razviti metodi na osnovi razlik semantičnih globin konceptov (sgk) in neobhodnih ravni razmerij konceptov, ki se izražajo kot globina neobhodne ravni razmerja (gnr), izračunati utež relativne globine (urg):

$$urg = sgk - gnr \quad [17]$$

Primer izračuna absolutne semantične prostranosti zbirke podatkov

Absolutna prostranost zbirke podatkov (apzp) [15] je vsota semantičnih razdalj zbirk podatkov, uteženih z utežmi neobhodne ravni (unr). Zato najprej izračunamo uteži neobhodne ravni (preglednica 10).

Utež neobhodne ravni (unr) $1/(gnr + 1)$		hmeljišče	bivanje	tirni
		1020303	101,00	10405
hmeljišče	1020303	x		
bivanje	101	0,50	x	
tirni	10405	0,50	0,50	x

Preglednica 10: Izračun uteži neobhodne ravni (unr) semantičnih razmerij (za vsak par primerjanih konceptov hipotetične zbirke podatkov P1 {hmeljišče, bivanje, tirni})

Utežene semantične razdalje $(r) \cdot (unr)$		hmeljišče	bivanje	tirni
		1020303	101	10405
hmeljišče	1020303	x		
bivanje	101	$4 \cdot 0,5 = 2$	x	
tirni	10405	$5 \cdot 0,5 = 2,5$	$3 \times 0,5 = 1,5$	x

Preglednica 11: Izračun uteženih semantičnih razdalj  $(r) \times (unr)$

Utež neobhodne ravni ( $unr$ ); ta utež odraža razsipanje zbirke podatkov po semantični širini (prostranosti). Za izračun absolutne semantične prostranosti zbirke podatkov ( $apzp$ ) semantične razdalje razmerij ( $r$ ) pomnožimo z utežmi neobhodne ravni ( $unr$ ) in tako utežene razdalje seštejemo.

Izračun vsote uteženih semantičnih razdalj oziroma izračun absolutne prostranosti zbirke podatkov P1 {hmeljišče, bivanje, tirni}:  $(r) \cdot (unr) = 6$  (preglednica 11).

## 6.4 SEMANTIČNA RAZPRŠENOST

Semantična razpršenost zbirke prostorskih podatkov je mera splošnosti oziroma osredotočenosti, generalizacije oziroma specializacije zbirke prostorskih podatkov. Izračun temelji na uteženih semantičnih razdaljah razmerij konceptov zbirke podatkov. V razviti metodi ločujemo dva pristopa k izračunu semantične razpršenosti zbirke podatkov.

Izračun semantične razpršenosti zbirke prostorskih podatkov – osnovni pristop

Pri osnovnem pristopu izračuna semantične razpršenosti zbirke podatkov izračunamo vsoto relativne prostranosti zbirke podatkov ( $rpzp$ ), [16], in relativne razsutosti zbirke podatkov po semantični globini ( $rrzpg$ ), [19].

Absolutna razsutost zbirke podatkov po semantični globini ( $arzpg$ ) je vsota semantičnih razdalj ( $r$ ), uteženih z utežmi relativne globine ( $urg$ ), [17]:

$$arzpg = \sum r \cdot urg \quad [18]$$

Relativna razsutost zbirke podatkov po semantični globini ( $rrzpg$ ) je razmerje med absolutno razsutostjo zbirke podatkov po semantični globini ( $arzpg$ ), [20], in absolutno razsutostjo semantičnega referenčnega sistema – ogrodja ( $arso$ ):

$$rrzpg = arzpg/arso \quad [19]$$

Semantična razpršenost zbirke podatkov ( $srzp$ ) je vsota relativne prostranosti zbirke podatkov ( $rpzp$ ), [16], in relativne razsutosti zbirke podatkov po semantični globini ( $rrzpg$ ), [19]:

$$(srzp) = (rpzp) + (rrzpg) \quad [20]$$

### 6.4.1 Izračun semantične razpršenosti zbirke prostorskih podatkov – poenostavljeni pristop

Pri poenostavljenem pristopu izračuna semantične razpršenosti zbirke podatkov izračunamo vsoto (dvakratno) uteženih semantičnih razdalj. Semantične razdalje se morajo izračunati za vse permutacije razmerij med koncepti. Vsako izračunano semantično razdaljo med dvema konceptoma istočasno utežimo z dvema raznovrstnima utežema:



- utež neobhodne ravni; utež izraža prostranost zbirke v okviru ontologije domene,
- utež relativne globine; utež odraža stopnjo relativnega razsipanja zbirke podatkov po semantični globini.

Semantična razpršenost zbirke podatkov se izračuna kot vsota uteženih semantičnih razdalj razmerij konceptov zbirke podatkov.

Primer izračuna uteži relativne globine

Za razmerja konceptov obravnavane zbirke podatkov na osnovi razlik semantičnih globin konceptov in neobhodnih ravni razmerij konceptov je najprej treba izračunati uteži relativne globine ( $urg$ ), [17] (preglednica 12).

Utež relativne globine ( $urg$ )		hmeljšče	bivanje	tirni	Semantična globina koncepta
		1020303	101	10405	
hmeljšče	1020303	x			4
bivanje	101	$(2-1, 4-1) 0,5 = 2$	x		2
tirni	10405	$(3-1, 4-1) 0,5 = 2,5$	$(3-1, 2-1) 0,5 = 1,5$	x	3
Semantična globina koncepta		4	2	3	

Preglednica 12: Izračun uteži relativne globine ( $urg$ ) semantičnih razmerij (za vsak par primerjanih konceptov hipotetične zbirke podatkov P1 {hmeljšče, bivanje, tirni})

Utež relativne globine odraža relativno razsipanje zbirke podatkov po semantični globini.

Primer izračuna semantične razpršenosti zbirke podatkov s poenostavljenim pristopom  
Izračun semantične razpršenosti hipotetične zbirke podatkov P1 znaša 54 enot (preglednica 13).

Uteži semantične razpršenosti $usr = (unr) \cdot (urg)$		hmeljšče	bivanje	tirni
		1020303	10100	10405
hmeljšče	1020303	x		
bivanje	101	4	x	
tirni	10405	6,25	2,25	x

Preglednica 13: Izračun uteži semantične razpršenosti ( $usr$ ) zbirke podatkov kot produkt uteži neobhodne ravni ( $unr$ ) in uteži relativne globine ( $urg$ )

Dobljeno vrednost (preglednica 14) je potrebno normalizirati, postaviti v razmerje do absolutne vrednosti (vsote uteženih razdalj vseh možnih razmerij med koncepti v ontologiji domene).

Utežene semantične razdalje		hmeljišče	bivanje	tirni
		1020303	10100	10405
1020303	hmeljišče	x		
101	bivanje	$4 \cdot 4 = 16$	x	
10405	tirni	$5 \cdot 6,25 = 31,25$	$3 \cdot 2,25 = 6,75$	x

Preglednica 14: Izračun uteženih semantičnih razdalj. Semantična razpršenost hipotetične zbirke prostorskih podatkov P1 (hmeljišče, bivanje, tirni),  $srzpp\ 1 = (16 + 13,25 + 6,75) = 54$

## 6.5 ANALIZA SEMANTIČNIH RAZMERIJ APLIKATIVNIH KONCEPTOV V OKVIRU SEMANTIČNEGA REFERENČNEGA OMREŽJA

V tem razdelku je izvedena analiza semantičnih razmerij aplikativnih konceptov v okviru univerzalne ontologije prostora (OPGR).

### 6.5.1 Analiza semantične ustreznosti OPGR

Izvedena je bila analiza povezljivosti posameznih katalogov prostorskih podatkov s semantičnim referenčnim omrežjem, ontologijo prostora geografskih razsežnosti. Določitev razmerij med simboli katalogov, ontologij aplikacij in koncepti ontologije domene OPGR je izvedena v ustreznih preglednicah (preglednica 19). Razmerja so ocenjena z dvema osnovnima semantičnima razmerjema, predikatoma in z negacijo razmerja:

- razredi ontologije aplikacije;
- razmerje povezave razredov (dveh ali več konceptov) ontologije domene, tako da v ontologiji aplikacije tvorijo en sestavljeni razred (agregat);

Katalog konceptov	Semantična skladnost				e + a	Semantična ustreznost OPGR		
	ekvivalent	agregat	nezdružljiv	skupaj		e + a/s (%)	n (%)	S (%)
Raba zemljišč – kataster	31	5	6	42	36	86	14	100
Kmetijske rabe zemljišč	34	15	9	58	49	84	16	100
TNN VM 500	155	56	3	214	211	99	1	100
Topografske baze VN	10	1	5	16	11	69	31	100
CORINE Land Cover	27	27	5	59	54	92	8	100
Prostorske oblike	95	60	7	162	155	96	4	100
Skupaj	352	164	35	551	516	<b>94</b>	6	100
	64 %	30 %	6 %	100 %				

Preglednica 15: Izračun semantične ustreznosti OPGR.

- če v OPGR ni mogoče najti ustreznega koncepta oziroma ustreznega agregata konceptov, je ocenjeno, da razmerje ne obstaja (nezdružljiv).

Povprečno je mogoče štiriindevetdeset odstotkov (94 %) od skupno 551 konceptov iz obravnavanih katalogov prostorskih podatkov izraziti s katalogom ontologije prostora geografskih razsežnosti (preglednica 15). Rezultat analize potrjuje visoko stopnjo semantične ustreznosti izdelane ontologije, še posebej, ker ontologija ni bila ustvarjena kot spoj obravnavanih zbirk podatkov – lokalnih ontologij, ampak kot neodvisna zasnova. Dokazano visoka stopnja semantične ustreznosti izdelane ontologije v razmerju do aplikativnih ontologij potrjuje visoko vrednost neodvisne zasnove OPGR in s tem njene univerzalnosti za domeno prostora geografskih razsežnosti. Zaradi visoke stopnje semantične ustreznosti je OPGR primerno orodje za ugotavljanje semantične skladnosti konceptov različnih zbirk prostorskih podatkov.

### 6.5.2 Analiza izrazov semantično nezdružljivih konceptov

Določenih konceptov posameznih katalogov zbirk prostorskih podatkov ni mogoče izraziti s simboli konceptov OPGR. Zato je bila izvedena analiza izrazov v nezdružljivih konceptih obravnavanih zbirk podatkov.

Rezultati analize razkrivajo, da se v nevrščenih konceptih najpogosteje pojavlja prostorski izraz zemljišče (12-krat). Sledijo mu prostorski izrazi stavba (7-krat), površina (4-krat), zgradba (3-krat) in objekt (3-krat). Torej sta najbolj dominantna izraza zemljišče in stavba. Izraza ostalo in drugo ne bosta predmet obravnave, ker se ne nanašata na prostor (preglednica 16 in preglednica 19).

Katalog	Izraz					
	stavba	zgradba	objekt	zemljišče	površina	ostalo/drugo
Katalog rabe zemljišč zemljiškega katastra	2	1	1	2	1	
Katalog kmetijske rabe zemljišč (projekt Zajem)				7		5
Katalog konceptov temeljnih topografskih načrtov velikih meril (TNN VM 500)	1	1				1
Katalog konceptov topografske baze večje natančnosti		1	1	1		1
Katalog konceptov kataloga CORINE land cover					3	
Katalog konceptov prostorskih oblik (GURS)	4		1	2		
Skupaj	7	3	3	12	4	7

Preglednica 16: Analiza izrazov semantično nezdružljivih konceptov

Izraz zemljišče se tako pomensko izkazuje kot sinonim izraza prostor, ki v kreirani ontologiji OPGR predstavlja izhodišče. Kot izhodišče ontologije je izraz prostor izjemno splošen in zato neprimeren za praktično označevanje konceptov. Očitno je pri oblikovanju kataloga kmetijske rabe zemljišč prevladala uporaba besede zemljišče kot sinonima za izraz prostor. V tem katalogu se je izraz zemljišče pojavil kar 7-krat. Tudi za uporabo izraza površina bi bilo mogoče sklepati, da je uporabljen za izražanje splošnega koncepta prostora, četudi preko svojega enakozvočnega, homonimnega pomena za koncept ploskev. Tako je mogoče izpeljati enakovredno razmerje med koncepti prostor – ploskev – zemljišče oziroma

združeno prostorska ploskev – zemljišče.

Analizo nadaljujemo z drugim najpogosteje zastopanim izrazom v nevrstljivih konceptih, to je z izrazom stavba (7), ki se mu sopomensko, kot sinonima, pridružujeta s skoraj enakim skupnim deležem v nevrstljivih konceptih izraz zgradba (3) in izraz objekt (3). Tudi za to skupino izrazov je mogoče izpeljati enakovredno razmerje med koncepti stavba – zgradba – objekt, ki jih ti izrazi predstavljajo. Izraz stavba ima v tem razmerju glede na numerične rezultate analize rahlo dominantno vlogo.

Če analizirane izraze in koncepte strnemo z vidika kognitivnega zaznavanja je mogoče pritrčiti, da je koncept prostora na najbolj splošni ravni sestavljen iz dveh elementov, ki sta:

- ploskev – zemljišče; element, ki ga zaznavamo kot odprt prostor in
- stavba – zgradba – objekt; element, ki ga v človeški percepciji zaznavamo kot tridi menzionalne kvadre, kubuse bolj ali manj pravilnih oblik in zato samodejno sklepamo na njihov umetni izvor.

Oba elementa sta predstavitev koncepta prostora, ki ga je mogoče predstaviti kot unijo:

$$\text{prostor} = \text{ploskev} - \text{zemljišče} \sqcup \text{stavba} - \text{zgradba} - \text{objekt} \quad [21]$$

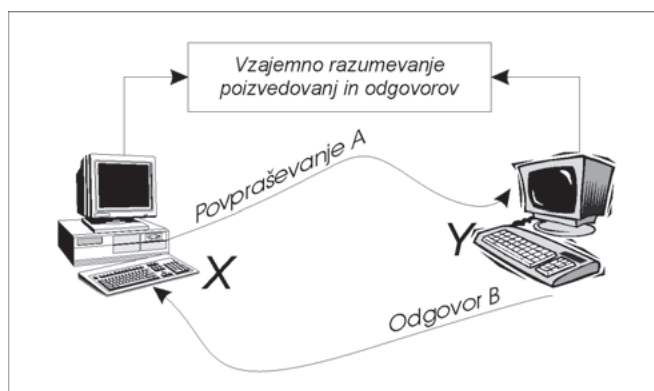
Pri opredeljevanju konceptov prostora za praktično uporabo (na primer v zbirkah podatkov) je treba presegati splošnost navedenih izrazov. Če se njihovi uporabi ni mogoče izogniti, jim je treba bogatiti pomen z dodajanjem simbolov/izrazov, ki imajo manj splošno semantiko oziroma večjo semantično globino. Takšni izrazi so povezani s koncepti človeških aktivnosti v prostoru.

# 7

## MEDOPRAVILNA INTEGRACIJA ZBIRK PROSTORSKIH PODATKOV V PORAZDELJENIH INFORMACIJSKIH SISTEMIH

Medopravilnost, medizvedljivost, interoperabilnost, je sposobnost sistema oziroma njegovih komponent za zagotavljanje porazdeljevanja informacij in nadzor postopkov medaplikativnega sodelovanja (Bishr, 1997). Medopravilnost je pogoj, ki se izpolnjuje med sistemi za elektronsko komuniciranje ali napravami za elektronsko komuniciranje, in to takrat, ko se podatki ali storitve med elektronskimi sistemi in/ali njihovimi uporabniki lahko izmenjujejo navidezno neposredno in zanesljivo. Izpolnjevanje pogoja medopravilnosti omogoča, podpira ter pospešuje delitev obdelav ter porazdeljevanje podatkov med različnimi sistemi. Medopravilnost je zmožnost komuniciranja, izvajanja programov in prenosa podatkov med različnimi funkcionalnimi enotami na način, ki od uporabnika ne zahteva posebnega poznavanja tehničnih značilnosti takšnih naprav.

Stopnja medopravilnosti je opredeljena v odvisnosti od posameznega primera komunikacije. Medopravilnost ima več vidikov, ki se nanašajo na naslednje značilnosti informacijskih sistemov: sposobnost pridobitve informacij in zagona obdelovalnih orodij, ne glede na to, kje se nahajajo; sposobnost razumevanja in uporabe pridobljenih informacij in orodij, ne glede na to, kakšne vrste je osnovna računalniška oprema, in ne glede na to, ali je delovanje lokalno ali daljinsko; sposobnost razvoja računalniškega okolja s tržnimi računalniškimi izdelki, brez pogojene navezave na ponudbe posameznega dobavitelja; sposobnost gradnje na osnovi informacijske in računalniške infrastrukture različnih dobaviteljev, za podporo tržnih niš, brez tveganja razpada sistema zaradi kasnejšega razvoja različne podporne infrastrukture; sposobnost povečanja h kakovosti trga, kjer dobrine in storitve predstavljajo odgovor na potrebe porabnikov in kjer so tokovi proizvodov odprti, tako da omogočajo rast trga za učinkovito podporo.



Slika 21: Medopravilnost informacijskih sistemov (povzeto po Bishr, 1997)

Stvarna medopravilnost – posebnost dveh sistemov, da izmenjujeta podatke in uporabljata izmenjane podatke – je odvisna od več dejavnikov in je časovno pogojena. Dva sistema X in Y lahko medopravilno sodelujeta, če lahko sistem X pošlje poizvedovanje po storitvi A sistemu Y, ob vzajemnem razumevanju storitve A s strani sistemov X in Y. Sistem Y lahko vrne sistemu X odgovor B, ob vzajemnem razumevanju B kot odgovora na poizvedovanje A s strani sistemov X in Y (slika 21).

Obstaja več ravni medopravilnosti. Komunikacijska medopravilnost temelji na poveztivosti tehnologij CORBA, Java, XML, GML in WAP. Na sistemski ravni medopravilnosti Frank et al. (2000) navaja opazovanje razvoja tehnologije za podporo sistemske medopravilnosti, ki ga je organizirala Evropska komisija v okviru Združenega raziskovalnega centra (Joint Research Centre – JRC) kot program opazovanja tehnologije za izmenjavo podatkov ter splošno, medopravilno interpretacijo pomena podatkov – semantično medopravilnost. Tretja raven medopravilnosti se nanaša na zakone, pravilnike, pogodbe in podobno, kar zagotavlja pravice uporabnikov za uporabo sistemov in izmenjavo podatkov; institucionalna medopravilnost je večkrat omejena z določenimi prepovedmi izmenjave in uporabe podatkov. Ravni izvedbe medopravilnosti od najnižje proti najvišji so sledeče (Bishr, 1997): medopravilnost mrežnih protokolov, datotečnih sistemov, postopkov daljinskih klicev, medopravilnost upravljanja zbirke podatkov in iskanja po njej, medopravilnost geografskih informacijskih sistemov in medopravilnost aplikacij. Medsebojno razumevanje povpraševanj in odgovorov med sistemi ima različen pomen, odvisno od ravni izvedbe medopravilnosti.

## 7.1 PORAZDELJENI INFORMACIJSKI SISTEMI IN SEMANTIČNA MEDOPRAVILNOST

Porazdeljeni sistem je zbirka neodvisnih, avtonomnih računalnikov, povezanih z računalniškim omrežjem in opremljenih s programsko opremo, ki omogoča izvajanje integriranih računalniških postopkov (Coulouris et al., 2001). Porazdeljeni informacijski sistem je sestavljen iz več neodvisnih procesnih elementov, ne nujno homogenih, ki so medsebojno povezani z računalniško mrežo in sodelujejo pri izvajanju določenih nalog (Ozsu, Valduriez, 1999). Glavne značilnosti porazdeljenih sistemov so: porazdeljevanje virov, odprtost, konkurenčnost, razširljivost, strpnost do napak v sistemu in odkritost.

V porazdeljenih – distribuiranih – informacijskih sistemih se porazdeljujejo procesna logika, funkcije, podatki in nadzor. Kot osnova za razvoj večine porazdeljenih informacijskih sistemov služi enostavna arhitektura strežnik-odjemalec (Coulouris et al., 2001). Strežnik nudi dostop odjemalcem do različnih porazdeljenih sistemskih naprav, storitev in virov. Drugi pristop k razvoju porazdeljenih informacijskih sistemov se imenuje objektno zasnovani sistem, v katerem se vsako porazdeljeno sredstvo obravnava kot objekt. Vsi objekti se obravnavajo na enoličen način, kar ne velja za arhitekturo strežnik-odjemalec.

Arhitektura objektno zasnovanega porazdeljenega informacijskega sistema je arhitektura, imenovana »posrednik objektnih zahtev/povpraševanj« (Object Request Broker – ORB), ki je opredeljena v specifikaciji, imenovani CORBA (Common Object Request Broker Architecture). Posrednik objektnih zahtev ORB vzdržuje zvezo in občevanje med objekti in njihovimi odjemalci in prikriva vidike porazdeljenosti in raznovrstnosti. Poleg objektov so za to arhitekturo značilni pripomočki, kot so vodoravne in navpične storitve. V specifikaciji je posebej opredeljen model oddaljenega objekta. Svetovni splet (World Wide Web – WWW) je porazdeljeni sistem dokumentov, ki temelji na arhitekturi odjemalec-strežnik. S povezovanjem dokumentov se ustvarja

nov, integrirani dokument. Pomemben element sistema porazdeljenih dokumentov je sredstvo za sklicevanje, povezava – hiperlink – na dokumente, označene z enotnim lokatorjem virov (Uniform Resource Locator – URL). Za izbiranje dokumentov in njihov pravilen prikaz skrbi brkljalnik. Operacijski sistemi za porazdeljene informacijske sisteme so zelo podobni tradicionalnim operacijskim sistemom po njihovi vlogi upravljanja virov – strojne opreme, računalniškega omrežja ter podatkov (preglednica 17). Vendar je najpomembnejša lastnost porazdeljenih sistemov prikrivanje porazdeljenosti sestavnih delov sistema ter zagotavljanje navideznega pogleda, v njegovem okviru je izvrševanje aplikacij nemoteno (Tanenbaum in Steen, 2002). Operacijski sistemi za porazdeljene računalnike se delijo na dve skupini: tesno povezani sistemi oziroma porazdeljeni operacijski sistem – POS (Distributed Operating Systems – DOS) in ohlapno povezani sistemi oziroma mrežni operacijski sistemi – MOS (Network Operating Systems – NOS). Porazdeljeni operacijski sistemi so namenjeni upravljanju večprocesorskih naprav ali več homogenih računalnikov. Mrežni operacijski sistemi so namenjeni upravljanju več nehomogenih računalnikov. Z nastavitvijo MOS postanejo lokalne storitve in naprave dosegljive oddaljenim uporabnikom v mreži. Najpogosteje napačno interpretirani kot porazdeljeni sistemi so sistem porazdeljevanja računalniškega časa, ohlapno ali čvrsto spojen večprocesorski sistem in podatkovni sistem, ki obstaja na

	Operacijski sistemi za porazdeljene informacijske sisteme			
	Porazdeljeni OS		mrežni OS	vmesni OS
	večprocesorski računalnik	več homogenih računalnikov		
<b>stopnja prikrivanja</b>	zelo velika	velika	nizka	velika
<b>enak OS na vseh vozliščih?</b>	da	da	ne	ne
<b>število kopij OS</b>	1	n	n	n
<b>osnova komunikacije</b>	deljene spomin-ske enote	sporočila	datoteke	glede na model
<b>upravljanje virov</b>	globalno, centralno	globalno, porazdeljeno	na vozlišču	na vozlišču
<b>razširljivost</b>	ne	zmerno	da	spremenljivo
<b>odprtost</b>	zaprt	zaprt	odprt	odprt

Preglednica 17: Primerjava med operacijskimi sistemi POS, MOS in VOS (privzeto po Tanenbaum, Steen, 2002)

enem od vozlišč računalniškega omrežja – centralizirana zbirka podatkov na mrežnem vozlišču (Ozsu in Valduriez, 1999).

Po prodoru mrežnih operacijskih sistemov so v mnogih organizacijah spoznali, da imajo veliko mrežno porazdeljenih aplikacij, ki jih ni enostavno integrirati v enoten sistem. Za doseganje integracije porazdeljenega sistema je treba porazdeljeni operacijski sistem izboljšati z vidika upravljanja neodvisnih in nehomogenih računalnikov. Z vidika prikrivanja porazdeljenosti oziroma zagotavljanja enotnega systemskega pogleda je potrebno izboljšati mrežni operacijski sistem, torej razviti sistem, ki ima prikrito porazdeljenost, transparentnost porazdeljenih operacijskih sistemov in zagotavlja razširljivost in odprtost mrežnih operacijskih sistemov. To vodi v posredovalni, vmesni sistem – middleware – za porazdeljeni informacijski sistem. To je

mogoče doseči z uvajanjem dodatne vmesne programske ravni med ravnijo aplikacij (CORBA, HTTP/XML, ActiveX, Java, ODBC) in ravnijo mrežnega operacijskega sistema, ki jo imenujemo trinivojska arhitektura porazdeljenih informacijskih sistemov.

Da bi bila razvoj in integracija porazdeljenih aplikacij čimbolj enostavna, naj bosta osnovana na enakem modelu porazdeljevanja in komunikacije. Modeli porazdeljevanja in komunikacije za vmesni sistem so naslednji: porazdeljeni datotečni sistemi, klicanje oddaljenih postopkov (Remote Procedure Call – RPC), porazdeljeni objekti in porazdeljeni dokumenti. Mnogo posredovalnih sistemov omogoča storitve, kot so prikrit dostop do oddaljenih podatkov in povezovanje aplikacij na porazdeljene datotečne sisteme ali porazdeljene zbirke podatkov. Primer prikritega enosmernega prenašanja podatkov/dokumentov z oddaljenih lokacij je splet. Storitve porazdeljenih transakcij omogoča, da se večkratne operacije zapisovanja/čitanja podatkov izvajajo atomično – posamezna operacija uspe v celoti, sicer propade. Storitve varovanja za vmesno raven niso osnovane neposredno na lokalnih operacijskih sistemih, ampak so delno nadgrajene. Aplikacije, zgrajene za določen porazdeljeni sistem so neodvisne od operacijskega sistema, vendar so zelo odvisne od vmesnega porazdeljenega sistema. Vmesni sistem je odprt manj, kot trdijo proizvajalci, kar predstavlja problem za doseganje medopravnosti (Ozsu in Valduriez, 1999).

### 7.1.1 Programski agent v porazdeljenih informacijskih sistemih

Programski agent je inteligentni posrednik med uporabnikom in programi. Stacionarni programski agenti so pasivni, čakajo na zunanjo iniciativo in imajo vlogo programskega orodja. Nasprotno so mobilni programski agenti proaktivni. Lastnosti mobilnega programskega agenta so: neodvisnost, usmerjeno delovanje, samoiniciativnost, selektivno zaznavanje in odzivanje, časovna neprekinjenost identitete in stanja, sodelovanje, delovanje v imenu drugih, sposobnost učenja, selitev – mobilnost – po svoji volji. Vse, kar lahko naredijo mobilni programski agenti, lahko naredijo tudi stacionarni agenti, vendar na drugačen način. Namesto povpraševalne zahteve oddaljenemu viru in lokalne obdelave se mobilni agent preseli na oddaljen računalnik, kar omogoča zmanjšan pretok po mreži, neodvisno in sočasno delovanje v računalniškem omrežju. Posledica mobilnosti agentov je zmanjšanje nadzora in napovedljivosti informacijskih sistemov, poveča se nevarnost pred zlonamernimi agenti (Pogačnik in Privošnik, 1998).

### 7.1.2 Porazdeljevanje kopij porazdeljenih podatkovnih skladišč

Ozsu in Valdarez (1999) ločujeta med porazdeljeno zbirko podatkov, porazdeljenim sistemom za upravljanje zbirke podatkov in porazdeljenim podatkovnim sistemom. Porazdeljena zbirka podatkov – PZP (Distributed Database – DDB) je zbirka večkratnih, logično povezanih zbirke podatkov porazdeljenih preko računalniškega omrežja. Porazdeljeni sistem za upravljanje zbirke podatkov (Distributed Database Management System – D-DBMS) je računalniški program, ki upravlja PZP in zagotavlja dostopni mehanizem, ki za uporabnike prikrije porazdeljenost. Porazdeljeni podatkovni sistem – PPS (Distributed Database System – DDBS) je vsota porazdeljene zbirke podatkov in porazdeljenega sistema za upravljanje zbirke podatkov; PPS = DDB + D-DBMS. Poznani so trije različni tipi reprodukcij, replikacij porazdeljenih podatkovnih skladišč: stalne kopije (zrcala spletnih strani, porazdeljene zbirke podatkov), dinamična reprodukcija na pobudo strežnika in dinamična reprodukcija na pobudo odjemalca. Zbirke podatkov so lahko porazdeljene in reproducirane med številne strežnike, ki skupaj tvorijo skupino delovnih postaj (Cluster of



Workstations – COW), kar se imenuje tudi arhitektura porazdeljenega niča, s čimer se poudarja, da se med procesorje delovnih postaj ne porazdeljujejo niti pomnilniške enote niti hitri pomnilnik. Vendar je med številne geografsko porazdeljene podatkovne strežnike porazdeljena zbirka podatkov, kar je poznano kot zvezno – federativno – porazdeljena zbirka podatkov.

### 7.1.3 Razširjanje obnovljenih podatkovnih nizov v porazdeljenih sistemih

Operacije obnavljanja podatkov se v okviru porazdeljenih sistemov izvajajo na strani odjemalcev. Razširjanje obnovljenih podatkovnih nizov v porazdeljenih in reproduciranih podatkovnih skladiščih poteka od odjemalca v eno od kopij, od tam pa na vse ostale kopije. V osnovi se razširjanje izvaja na tri načine: razširjanje sporočila o obnovitvi podatkov; prenos podatkov iz prve v drugo kopijo; razširitev operacije obnavljanja na preostale kopije.

Če so operacije obnovitve podatkov razširjene na pobudo strežnikov, se imenujejo protokol s pristopom poriva. Protokol s pristopom potega se uporablja, kadar strežnik ali odjemalec zahteva od drugega strežnika, naj mu pošlje obnovljene podatke. Slabosti in prednosti obeh protokolov so povzročile razvoj hibridnega razširjanja, ki temelji na zakupu. V zakupnem protokolu se strežnik zaveže, da bo določen zakupni čas »porival« obnovitve podatkov proti odjemalcu. Po preteku zakupnega časa mora odjemalec sam »potegniti« obnovitve podatkov od strežnika ali pa s strežnikom skleniti novo zakupno razmerje. Prilagodljivi zakup omogoča dinamično prilagajanje trajanja zakupa glede na različne zakupne kriterije.

Prednosti sestavljeno, kompozitno, porazdeljenih informacijskih sistemov pred centraliziranimi informacijskimi sistemi so pogosto opisane kot: povečana zanesljivost in dosegljivost, krajši čas dostopanja do podatkov, zmanjšano komunikacijsko prekrivanje, boljša varnost podatkov (Oszu in Valduriez, 1999). Z vidika uporabnika je kompozitno porazdeljen IS enoličen logični sistem, opisan z globalno shemo, ki daje uporabniku iluzijo centraliziranega informacijskega sistema. Ker vse sestavine sistema uporabljajo isti podatkovni model in isti univerzalni komunikacijski jezik, je kompozitno porazdeljeni informacijski sistem homogen. Globalna shema v takšnem sistemu opredeljuje razmerja med lokalnimi shemami sodelujočih sestavin. Slabost kompozitno porazdeljenega informacijskega sistema je ponavljanje postopka konstruiranja globalne sheme vsakič, ko je zaželeno sprememba lokalne sheme posameznega vozlišča, kar je stroškovno in časovno zahteven postopek.

Z razvojem interneta je postal takšen sistem preokoren. Zato je s strani uporabnikov zaželeno ohlapnejše povezovanje vozlišč. To je mogoče doseči s povezovanjem neodvisnih, avtonomnih, IS, ki izmenjujejo podatke brez osrednjega nadzora, brez centralno upravljane globalne sheme. Takšna povezava se imenuje zveza – federacija. Posamezni neodvisni IS se v zvezni povezavi imenuje vozlišče, ki neodvisno odloča o pogojih za izmenjavo in porazdeljevanje podatkov z ostalimi vozlišči. Cilja vzpostavitve zveznih informacijskih sistemov sta sodelovanje – kooperacija – in medsystemska izvedljivost – medopravilnost – med vozlišči.

Za neodvisne informacijske sisteme so značilne različne oblike avtonomnosti (Veijalainen, 1988). Oblikovna avtonomnost pomeni, da vsako vozlišče porazdeljenega informacijskega sistema neodvisno izbira lastno obliko (podatkovni model, poizvedovalni jezik, predstavitev podatkov, omejitve pristopa in podobno). Komunikacijska avtonomnost pomeni neodvisnost vozlišča pri izbiranju komunikacijskih značilnosti z drugimi vozlišči – avtonomno odločanje o tem kdaj in kako odgovarjati na povpraševanja drugih vozlišč. Izvedbena avtonomnost pomeni, da so lokalne operacije znotraj vozlišča neodvisne od operacij, ki potekajo v kombinaciji z ostalimi vozlišči. Avtonomnost neodvisnih informacijskih sistemov povzroča z vidika povezovanja tovrstnih

sistemov več tipov raznovrstnosti. Nekatere raznovrstnosti so posledica oblikovnih razlik v DBMS, druge pa razlik v pomenu podatkov – semantiki. Pomenska raznovrstnost podatkov – semantična heterogenost – se pojavlja, kadar obstaja neskladje med osnovno razlago, interpretacijo podatkov in med nameravano uporabo podatkov, lahko pa tudi, kadar je isti pojav modeliran v dveh shemah na različen način, četudi sta shemi oblikovani na osnovi istega podatkovnega modela. Pri prehajanju iz obravnave računalniških sistemov (formalni, dobro definirani računski vzorci za upravljanje znakov) na obravnavo pomenskih, semantičnih, vprašanj (vprašanje pomena podatkov in uporabe podatkov v različnih aplikacijah) je potrebno obravnavati in uravnati protislovja, nedoslednosti, nepopolnost informacij, različne razlage istovrstnih podatkov s strani različnih uporabnikov, ki delujejo na različnih strokovnih področjih. Pomen podatkov je v manjši meri ponazorjen v okviru konceptualne sheme, v večji meri pa v okviru programske, aplikativne sheme. Pomen podatkov ni ponazorjen v programski kodi, ampak v predpostavkah o podatkih, ki jih aplikacija notranje vsebuje. Za obvladovanje pomenske raznovrstnosti podatkov v okolju zvezno porazdeljenih informacijskih sistemov je treba nedvoumno odkriti semantične informacije, predpostavke skrite v aplikacijah, in jih eksplicitno prikazati s pojmovnimi, konceptualnimi shemami.

## 7.2 STANDARDI ZA POVEZOVANJE ZBIRK PROSTORSKIH PODATKOV

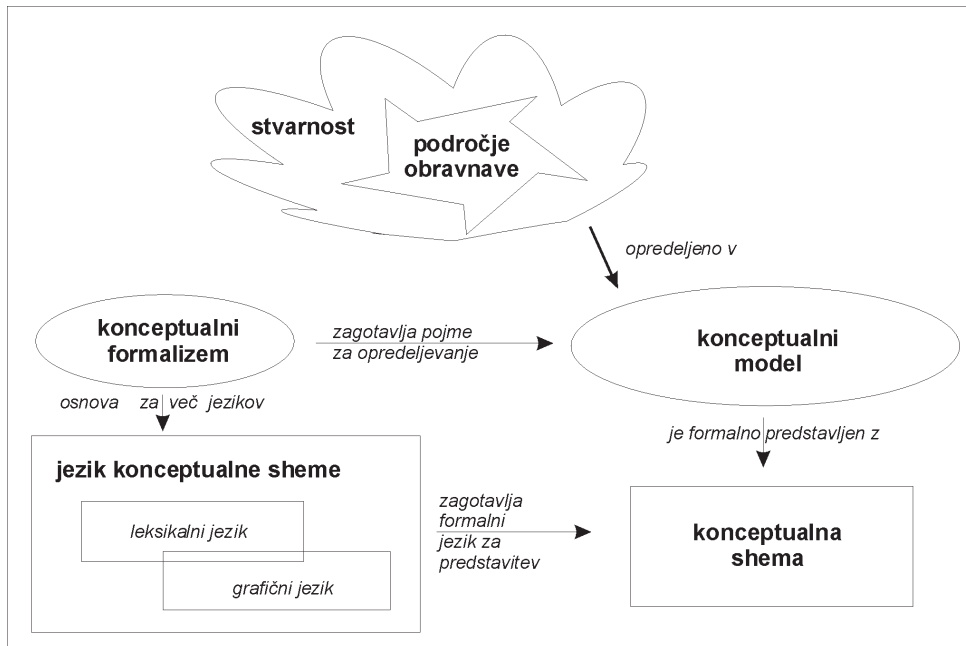
Popolnoma medopravilni GIS je predmet številnih raziskav (Vckovski, 1998). Med uporabniki in proizvajalci informacijske tehnologije se povečuje zavedanje o tem, da je pomemben vidik organiziranja in uporabe digitalnih podatkov dosežen z vpisovanjem, registracijo in indeksiranjem podatkov v odvisnosti od položaja v geografskem prostoru (Albrecht, 1999). Ker se veliko zbirk podatkov že navezuje na položaj v prostoru, je narasla potreba po standardizaciji tovrstnih podatkov. Razvita je bila skupina mednarodnih standardov za področje prostorskih geografskih podatkov in geomatike (ISO/TC 211). Najpomembnejši cilj razvoja navedene skupine standardov je pospeševanje integracije in porazdeljevanje geografskih podatkov oziroma tako imenovana medopravilnost geografskih informacijskih sistemov v porazdeljenih informacijskih okoljih oziroma sistemih.

Standardizacijski postopek za izdelavo skupine mednarodnih standardov ISO 19100 je osnovan na povezavi konceptov geografskih informacij in konceptov informacijske tehnologije. Zato so pri izdelavi referenčnih modelov za standarde geografskih informacij uporabljeni koncepti izpeljani iz standardov za okolja odprtih sistemov (Open Systems Environment – OSE) opisanih v standardu ISO/IEC TR 14252 in drugih standardih. Standard za referenčni model (ISO 19101) je organiziran v naslednje glavne sklope: konceptualno modeliranje, referenčni model domene, referenčni model arhitekture in profili.

Konceptualno modeliranje je pri opredelitvi skupine standardov ISO 19100 pomembno z vidika podatkov in informacij kakor tudi z vidika računalništva. Konceptualno modeliranje je uporabljeno za natančen opis geografskih informacij, za opredelitev storitev transformiranja in izmenjave geografskih informacij. Konceptualno modeliranje za izdelavo skupine standardov ISO 19100 temelji na referenčnem modelu za odprte porazdeljene obdelave ISO/IEC 10746 (Open Distributed Processing – ODP) in načelih, ki so del pripomočkov za modeliranje konceptualnih shem ISO/IEC 14481 (Conceptual Schema Modelling Facilities – CSMF).

Zasnove, ki so podlaga konceptualnemu modeliranju skupine mednarodnih standardov

19100, so: določitev jezika konceptualnih shem, pristop in načela konceptualnega modeliranja, ki so opredeljena v pripomočkih za modeliranje konceptualnih shem (SCMF) in koncept integracije modelov.



Slika 22: Grafična predstavitev vloge konceptualnega modeliranja za opredeljevanje prostorskih podatkov določenega področja obravnave stvarnega sveta (privzeto po ISO/DIS 19101)

Domena geografskih informacij je predstavljena z referenčnim modelom domene. Referenčni model domene geografskih informacij ISO 19101 opredeljuje glavne koncepte, ki so uporabljeni za predstavitev, organiziranje, shranjevanje, izmenjavo in analize geografskih informacij na računalniški način, v skupini mednarodnih standardov ISO 19100 (slika 21). Glavni koncepti v referenčnem modelu domene geografskih informacij so: prostorski objekti, opisi položajev prostorskih objektov, sloj, aplikacijska shema, metapodatkovni niz, geoinformacijska storitev.

Konceptualni model je abstraktni opis delov stvarnega sveta geografskih razsežnosti (na primer vodni tokovi, otoki, gore) in/ali niz pripadajočih konceptov. Niz pripadajočih konceptov je na primer niz geometričnih sestavnih delov kot so točke, linije in ploskve, s katerimi je mogoče predstaviti dele stvarnega sveta geografskih razsežnosti. Konceptualni model, ki je opisan z jezikom konceptualnih shem se imenuje konceptualna shema. Jezik konceptualnih shem zagotavlja sintaktične in semantične elemente, s katerimi je mogoče natančno opisati konceptualni model in zanesljivo izražati pomen. Uporaba jezika konceptualnih shem je temeljni pogoj za razvoj standardnih opredelitev geografskih informacij in geografskih informacijskih storitev. Standardne sheme v skupini standardov ISO 19100 služijo razvijanju skladnih shem za občevanje med geografskimi informacijskimi storitvami in programskimi sistemi. Za specifikacijo predpisanih delov skupine standardov ISO 19100 je kot temeljni jezik konceptualnih shem izbran jezik UML, ob podpori jezika ISO IDL (Interface Definiti-

on Language) ter jezika OCL (Object Constraint Language). Ta izbira omogoča izdelavo ogrodja za izdelavo standardov, ki podpirajo sintaktično medopravilnost in semantično medopravilnost. V jeziku UML je mogoče opredeliti konceptualne sheme, ki so namenjene izvedbi v določeni zbirki podatkov.

Konceptualna modeliranje (ISO/DIS 19101) je opis področja obravnave – nominalne osnove s konceptualnim modelom. Model lahko vsebuje poleg pojavov tudi njihove lastnosti, funkcije in razmerja, ki obstajajo med njimi. Konceptualni model je izveden s konceptualno shemo, pri katere zgradnji se uporablja jezik konceptualne sheme. Ta jezik upošteva konceptualni formalizem, ki vsebuje pravila, omejitve, mehanizme dedovanja, dogodke, funkcije, postopke in druge elemente, ki so sestavni deli jezika. Konceptualni formalizem zagotavlja formalne opredelitve vseh znanj, ki so smotrna pri izgradnji aplikacije informacijske tehnologije. Isti konceptualni formalizem je mogoče uporabljati v več jezikih konceptualnih shem, leksikalnih ali grafičnih.

Načela konceptualnega modeliranja za razvoj konceptualnih shem v skupini standardov ISO 19100 so opisana v standardu ISO/IEC 14481. Načelo stotih odstotkov pomeni, da mora konceptualna shema opredeljevati izbrano področje obravnave v popolnosti. Načelo konceptualizacije predpisuje, da lahko konceptualna shema vsebuje le tiste strukture in vedenjske vidike konceptov, ki so relevantni za področje obravnave, ne pa vidikov notranje ali zunanje predstavitve podatkov; to pomeni, da mora biti konceptualna shema razvita neodvisno od fizičnih vidikov računalniške tehnologije. Helsinško načelo predpisuje, da mora vsaka pomembna izmenjava govornih ali napisanih trditev temeljiti na dogovorjenih nizih sintaktičnih in semantičnih pravil; načelo uporabe formalno določene sintakse predpisuje uporabo formalno določene sintakse v jeziku konceptualne sheme. Načelo samoopredelitve predpisuje, da mora biti možen samoopis predpisanih gradnikov mednarodnih standardov, na primer ISO 14046. Koncept integracije modelov omogoča medopravilnost storitev geografskih informacij in geografskih informacijskih sistemov v porazdeljenih informacijskih okoljih. Konzorcij odprtega GIS (OGC) združuje več kot sto petdeset organizacij članic, akademskih, uporabniških in dobaviteljskih. V okviru tega industrijskega združenja skušajo razviti specifikacije, standarde, ki bi omogočali medopravilnost za prostorske podatke in postopke obdelave v geografskih informacijskih sistemih (OpenGIS, 1998).

Koncept "katalog" v specifikaciji OpenGIS opisuje niz storitvenih vmesnikov, ki podpirajo organizacijo in odkrivanje prostorskih informacij. Katalog vsebuje meta podatke, ki opisujejo vsebino in uporabnost podatkov v različnih virih geografskih informacij. Standard za pridobivanje znanja iz shranjenih meta podatkov predvideva orodja za upravljanje semantičnih razmerij med koordinatami in nadzorovanimi ključnimi besedami. Orodja temeljijo na tehnikah pridobivanja znanja z umetno inteligenco. Tovrstne tehnike omogočajo poleg horizontalne integracije, to je združevanje podatkov istega tipa za različna področja, tudi vertikalno integracijo različnih tipov podatkov za isto področje, kar je velika potreba pri integraciji porazdeljenih zbirk podatkov o geografskem prostoru (Banares, 2000).

Osnutek mednarodnega standarda za metodologijo objektnih katalogov geografskih informacij (ISO/DIS 19110). Ta standard zagotavlja standardni okvir za organiziranje in poročanje o klasifikaciji pojavov stvarnega sveta v nizih geografskih podatkov. Katalog objektnih tipov ne more v celoti zajeti zapletenosti in raznolikosti geografske stvarnosti. Vendar mora biti uporabnikom zbirk podatkov jasno predstavljen in dosegljiv niz uporabljenih abstrakcij, ki so opredeljene s katalogom objektnih tipov. Uporaba standarda ne zagotavlja harmonizacije in medopravilnosti aplikacij. Standard ne predpisuje kriterijev za smemalna pravila – za zajemanje podatkov posameznih pojavov stvarnosti. Smemalna pra-

vila morajo biti posebej, za vsak podatkovni niz, opredeljena v specifikaciji zbirke podatkov (Uitermark 2001). Standard opredeljuje metodologija za katalogiziranje objektnih tipov. Z njim je mogoče ustvarjati nove kataloge ali pregledovati in popravljati obstoječe kataloge. Ta standard je mogoče uporabiti kot osnovo za opredeljevanje področja obravnave, ki se ga modelira v aplikacijah (objektni atribut, objektni katalog, objektna operacija, objektno razmerje, objektni tip, funkcionalni jezik).

Jezik UML (Unified Modelling Language) je kot industrijski standard razvil konzorcij OMG (Object Management Group). UML je objektno usmerjen, umetni, formalni jezik, ki je neodvisen od področja in okolja uporabe. Glavni cilj jezika UML je poenotenje sodobnih objektno usmerjenih jezikov za modeliranje sistemov (Šumrada, 1999). Jezik UML sestavljajo tri skupine elementov: osnovni bloki, pravila za uporabo blokov ter povezovalna načela s pripadajočimi mehanizmi. Osnovni bloki so treh tipov: predmetni, relacijski in diagramski. Običajne vidike informacijskega sistema je mogoče predstaviti z diagrami, ki so sestavni del jezika UML. Običajni vidiki na informacijski sistem so: povezave med razredi in hierarhije razredov, sestava posameznih objektov, procesno obnašanje posameznih objektov – dinamika objekta, stanja sistema – dinamika sistema in uporabniški pogledi na sistem – vrste uporabe sistema. Za izmenjavo modelov izdelanih z jezikom UML je bila izdelana specifikacija XMI (XML Metadata Interchange). Specifikacija temelji na uporabi jezika XML. Jezik XML je namenjen poenostavitvi uporabe jezika SGML (Standard Generalized Markup Language), ki je opredeljen z mednarodnim standardom ISO 8897. XML je zgled za kodni jezik GML pri prenosu geografskih podatkov v sklopu skupine standardov ISO 19118 (Encoding). Specifikacija XML bo v prihodnosti omogočila vzpostavitev medopravnosti med programskimi orodji za računalniško podprto inženirstvo programskih sistemov (CASE) oziroma (OOAD).

GML (Geography Markup Language) je različica XML za kodiranje geografskih objektov. GML omogoča sestavo nevtralnega formata za prostorske, geografske podatke, s katerim lahko kodiramo večino obstoječih oblik prostorskih podatkov in informacij. GML razvija industrijsko združenje OGC (Open GIS Consortium), kjer sodelujejo tudi vodilni proizvajalci GIS orodij. Z GML je mogoče prenašati geografske informacije kot ločene objekte in uravnavati njihov prikaz. Uporabniki lahko geografske informacije prikazujejo z običajnimi pregledovalniki internetnih vsebin. Pomembnejše koristi uporabe GML so sledeče: povečana kakovost prikaza geografskih informacij, uporaba obstoječih pregledovalnikov brez nadgradnje z odjemalskimi programi, uporabniško prilagajanje prikaza informacij, vstavljanje povezav med objekt in spletnimi naslovi, neposredno povpraševanje po lastnostih objekta, neposredno določanje vsebine prikaza, animacija prikaza objektov, ki se s časom spreminjajo, univerzalnost formata za različne naprave kot so miniaturni računalniki in brezžične komunikacijske naprave – mobilni telefoni, ki omogočajo uporabo XML formata in verižno dodajanje storitev za izdelavo informacij (Lake, 2002).

OWL (Ontology Web Language) je spletni jezik za izražanje ontologij, ki je predlog standarda. Zagotavlja standardiziran način izražanja semantike.

ODMG 3.0 je specifikacija za shranjevanje objektov (The Standard for Storing Objects), (ODMG, 2000). ODMG je presek treh obstoječih domen standardov: standarda zbirke podatkov (SQL), objektnih standardov (OMG) in objektnih programskih jezikov (C++, Smalltalk in Java). ISO SQL3 je standard objektno nadgradnje SQL2, ki se imenuje tudi objektno poizvedovalni jezik (Object Query Language – OQL). Je deklarativni jezik za učinkovito povpraševanje po objektnih zbirkah podatkov vključno z zelo prvinskimi objektnimi nizi in sestavami struktur (Cattell in Douglas, 2000).

SQL Multimedia (SQL/MM) je mednarodni standard ISO (ISO/IEC 13249, 2000), sestavljen

iz več delov, ki so delno neodvisni drug od drugega. Povezuje jih prvi del, imenovan Ogradje, v katerem so opredeljeni skupni pojmi in sestava ostalih delov: polno besedilo SQL/MM, prostorski SQL/MM, negibne podobe SQL/MM, rudarjenje podatkov SQL/MM.

SQL/MM prostorski standard je namenjen standardizaciji uporabe podatkov (avtomatizirana kartografija, GIS in podobno), in sicer s prostorskih vidikov, kot so geometrija, položaj v prostoru, lokacija in topologija. SQL/MM prostorski standard opredeljuje več vrst hierarhij, med katerimi je najbolj splošni tip ST\_Geometry (maksimalni supertip). Za opis prostorskih referenčnih sistemov je opredeljeni tip ST\_SpatialRefSys. Prostorski standard SQL/MM je tesno povezan in usklajen z drugimi prostorskimi standardi, kot je skupina mednarodnih standardov ISO 19100.

Algoritem prekrivanja poligonov podpira opredelitev GIS sistema kot integracijskega medija različnih zbirk podatkov. Tako Crown (1990) opredeljuje GIS kot sistem za podporo odločanja, ki vključuje integracijo zbirk prostorskih podatkov v okolju za upravljanje problemov. Pri tej opredelitvi je poudarek na integraciji, še posebej je mišljena funkcionalnost prekrivanja poligonov, ki je temeljni pristop za sintezo različnih podatkovnih virov prostorskih podatkov. Opredelitev žal ne upošteva vidikov integracije prostorskih podatkov na osnovi semantike. Žalik in Podgorelc (2001) obravnavata GIS kot sistem integracije različnih podatkovnih tipov podatkov, znanj, uporabnikov in socioloških dejavnikov.

Podatkovni model specifikacije OpenGIS vsebuje semantično neusklajenost konceptov pojava z geometrijo (feature with geometry) in sloja (coverage). Uporaba industrijsko ustvarjene terminologije "coverage" je vir nerazumevanja. OpenGIS specifikacija bi morala opisovati obravnavane koncepte v bolj abstraktni in formalni terminologiji (Camara et al., 1999). Nekateri semantični modeli imajo bogatejšo vsebino kot model OpenGIS (na primer "category-feature" v MGE-Integraph, in "geo-field" v SPRING sistemu). Dokler ne bodo v konzorciju izvedli semantičnega modeliranja na višji ravni, se zaradi navedenih pomanjkljivosti odsvetuje prenos in povezovanje podatkov z upoštevanjem te specifikacije (Camara et al., 1999). Na področju integracije zbirk podatkov se za standardizacijo strukture shem uporablja postopek prevajanja shem.

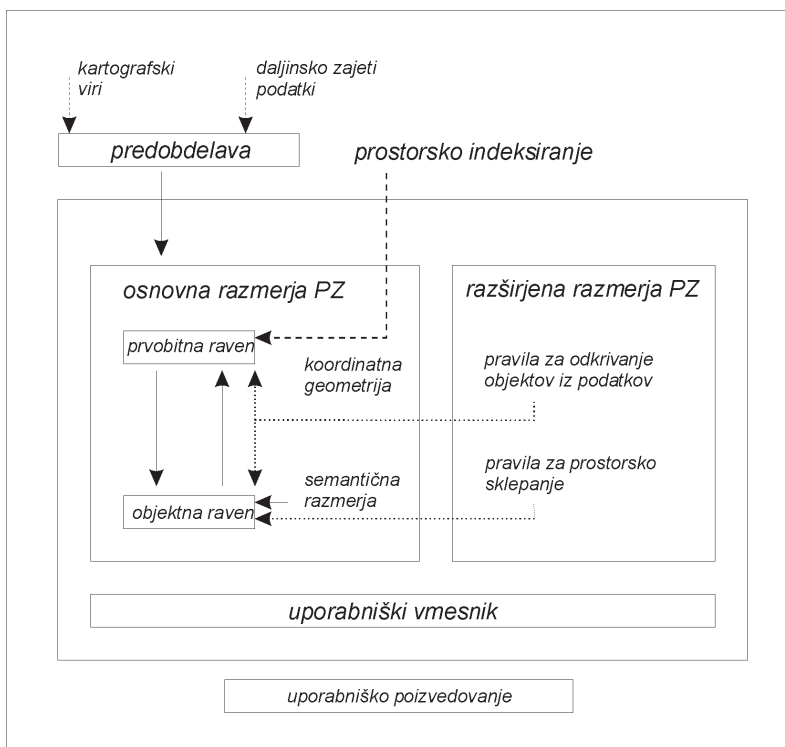
Semantični modeli obstoječih informacijskih sistemov predstavljajo jasne ovire za doseganje medopravnosti geografskih informacijskih sistemov. V večini primerov so namreč semantični modeli izvedeni na osnovi praktičnih vidikov v zvezi z uporabljenimi podatkovnimi strukturami za predstavitev prostorskih podatkov na računalniških sistemih (Camara et al., 1999). Za modeliranje programskih sistemov na tehnološkem področju geografskih informacijskih sistemov je značilna uporaba nestandardnih metodoloških pristopov. Razlogi za to so v kratkosti razpoložljivega obdobja za razvoj sistema do operativne faze ter nezadostna upravljalvska in finančna podpora (Šumrada, 1999).

### 7.3 DEDUKTIVNI OBJEKTNO USMERJEN PRISTOP ZA MEDOPRAVILNO UPRAVLJANJE ZBIRK PROSTORSKIH PODATKOV

V objektno usmerjenem geografskem podatkovnem modelu (OO GDB) je za izražanje značilnosti prostorske strukture zapletenih prostorskih razmerij potrebno uporabiti deduktivni pristop (Abdelmoty et al., 1993). Dedukcija je sklepanje s splošnega na posamezno. Prostorsko sklepanje je metoda sklepanja, s katero se tvorijo informacije, ki niso bile posebej shranjene v zbirki podatkov. Deduktivne sposobnosti sistema za upravljanje zbirk prostorskih podatkov predstavljajo pomemben mehanizem za izražanje povpraševanj in omejitev pri priklicevanju podatkov iz zbirke podatkov.

Povpraševanja se izražajo z uporabo jezika logike prvega reda (Abdelmoty et al., 1993). Prostorsko-podatkovna zbirka se obravnava kot sestava dveh ločenih nizov razmerij v zbirki podatkov, in sicer (slika 23): osnovna razmerja zbirke podatkov – ekstenzije – ter poglobljena razmerja zbirke podatkov – intenzije. Osnovni niz razmerij zbirke podatkov se obravnava na dveh ravneh. Na prvobitni ravni so prostorski objekti predstavljeni s primernim geometričnim podatkovnim modelom vektorske ali mozaične vrste (tesalacije oziroma rastri). Na tej ravni obstajajo tudi algoritmi za prostorske operacije koordinatne geometrije – COGO – ter strukture prostorskega indeksiranja. Na objektni ravni so pojavi stvarnega sveta predstavljeni kot razredi objektov in semantičnih razmerij med njimi, kar omogoča obravnavo zapletenih prostorskih pojavov. Med obema ravnema obstajajo povezave. Razširjeni niz razmerij zbirke podatkov je niz pravil za sklepanje, veljavnih za objekte osnovnega niza, ki se uporabljajo za prostorsko sklepanje v prostoru geografskih razsežnosti, pri odkrivanju objektov in prepoznavanju njihovih značilnosti iz podatkov, ki se nahajajo na prvobitni ravni.

Zapleteni prostorski objekt se opredeli z znaki, značilnimi za druge, enostavnejše prostorske objekte, ter z razmerji zapletenega objekta do drugih prostorskih objektov. Opredelitve razredov zapletenih prostorskih objektov se izvedejo s pravili prostorskega sklepanja v okviru razširjenega niza razmerij. Pravila za opredelitev določenega razreda objektov so lahko različna glede na lokacijo v prostoru, kar izvira iz oblikovnih, kulturnih, upravnih in drugih različnosti obravnavanih območij. Zaradi upravne različnosti so na primer pravila, ki opredeljujejo gradnjo hiš, cest in podobno, različna.



Slika 23: Sestavni deli deduktivne objektno usmerjene zbirke prostorskih podatkov (povzeto po Abdelmoty et al., 1993)

Prostorska razmerja se izražajo na ravni objektnih razredov. Nekatera so splošna in veljajo za vse primere objektnega razreda. Posebna prostorska razmerja veljajo le med določenimi primeri objektov enega ali več objektnih razredov. Posebna razmerja med določenimi primeri objektnih razredov so zelo pogosta v aplikacijah GIS. V strukturah za indeksiranje, kot so drevesne strukture, je ta razmerja mogoče shraniti vnaprej (na primer kd-drevesa, r-drevesa). Drug možen pristop je dinamično ocenjevanje prostorskih razmerij s pomočjo prostorskega sklepanja. Za izvedbo deduktivnega prostorskega sklepanja so potrebne identifikacija, klasifikacija ter formalna opredelitev prostorskih razmerij. Egenhofer (1991) je razvil niz temeljnih prostorskih razmerij (razdružen, se\_dotika, enak, prekrit\_z, vsebuje, pokriva, prekriva), s katerimi je mogoče izvesti štiriinšestdeset binarnih razmerij.

## 7.4 STRATEGIJA RAZVOJA MEDOPRAVILNIH DRŽAVNIH ZBIRK PROSTORSKIH PODATKOV

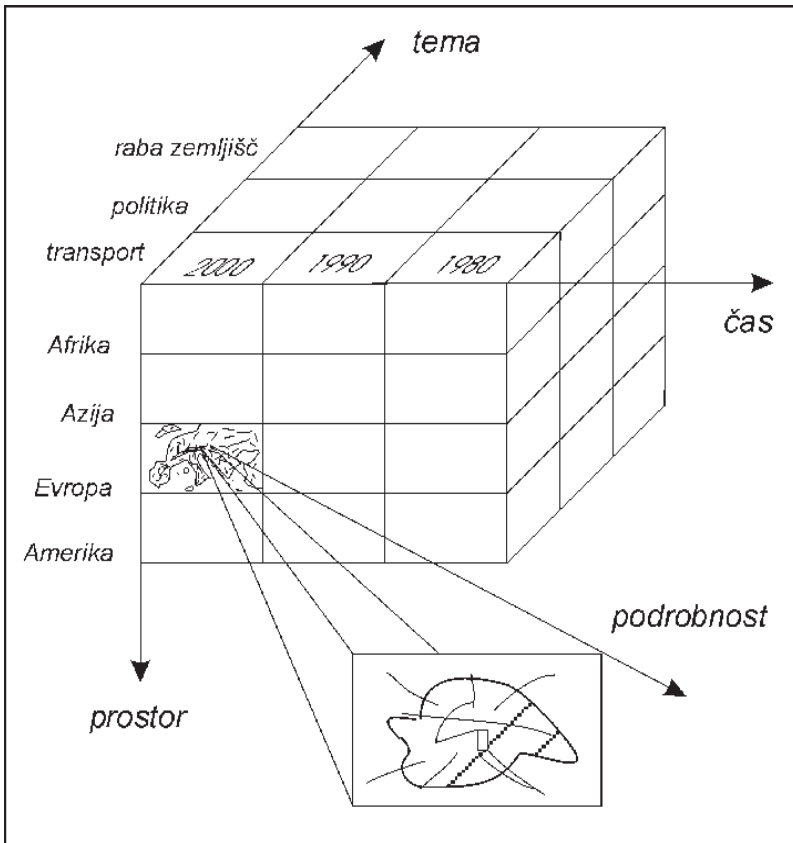
Oblikovanje državne informacijske strategije za prostorske podatke (National Geographic Information Strategy – NGIS) v prvi vrsti zadeva državno upravo. Ta je skoraj v vseh državah primarni pokrovitelj državnega geodetskega sestava in osnovnega sestava topografskih podatkov (slika 24). Poleg tega je država sponzor vzpostavitve drugih državnih zbirk podatkov, kot so zbirke podatkov geologije, pedologije, meteorologije, onesnaženja, demografije, lastništva nepremičnin, obdavitve nepremičnin, statističnih zbirk podatkov in drugih (Rhind, 1995).

Največji korak pri upravljanju prostorskih podatkov od iznajdbe kart je bil storjen z zasnovano geografske informacijske tehnologije (GIT, GIS). Tehnologija GIS je neobhoden, vendar nezadosten element za učinkovito delovanje sistema državne infrastrukture prostorskih podatkov. Ta obstaja v večini razvitih in razvijajočih se dežel, v večji ali manjši meri. S pojmom »državna informacijska strategija za prostorske podatke« je izražena potreba po aktivnosti državne uprave za izboljšanje obstoječe infrastrukture, zaradi povečanja odzivnosti na nova povpraševanja trga, ki jih povzroča nezadržen napredek informacijske in geomske tehnologije.

Primerjalna analiza delovnih postopkov v lokalnih upravah devetih evropskih držav je pokazala tesno povezavo med dosegljivostjo digitalnih prostorskih podatkov in razširjenostjo uporabe tehnologije GIS. Iz analize je očitna potreba po razvoju strategije vladnih organizacij, ki podpira integracijo zbirk podatkov državne informacijske infrastrukture. Na oblikovanje strategije vpliva več posrednih činiteljev, kot so zakoni, odloki, navodila, razvojne in podatkovne cenovne politike (Masser et al., 1996, Masser, 1998). V več državah so bili ustanovljeni koordinacijski odbori za oblikovanje in izvedbo državne informacijske strategije za geografske podatke NGIS.

Prvi koordinacijski odbor na evropski celine je bil leta 1984 ustanovljen nizozemski odbor za nepremičninske informacije (RAVI). Izdelan je bil načrt strategije nacionalne geografske informacijske infrastrukture za spodbujanje ekonomije, za podporo prostorskemu planiranju in investiranju ter za pospeševanje uravnoveženega razvoja (Ravi, 1995a, b). Na ministrski konferenci G7 leta 1995 v Bruslju je bila predlagana evropska geografska informacijska infrastruktura, s katero se predvidevajo sledeče bodoče koristi: povečanje učinkovitosti združenih evropskih tržišč; zmanjšanje problemov pri izvedbi prekomejnih in vseevropskih projektov; nove poslovne priložnosti za evropsko geografsko informacijsko industrijo (GII); sposobnost oblikovanja medopravilnih tehničnih rešitev za nadaljnji razvoj; povečanje uporabe evropskih veččin v GII in izboljšanje njenega tržnega položaja; izboljšanje možnosti za vseevropsko planiranje in odločanje.





Slika 24: Različne razsežnosti geografskih podatkov v državni informacijski strategiji za medopravilne prostorske podatke NGIS (privzeto po Abdelmoty et al. 1993)

Uspesne državne informacijske strategije za prostorske podatke imajo lahko pomembne učinke na družbo kot celoto, in sicer pozitivne in negativne. Dva možna scenarija razvoja in učinkov strategije navajata Wegner in Masser (1996). Prvi je tržno usmerjen scenarij: prostorski podatki se tržijo kot vsi ostali izdelki. V tem scenariju lahko državne uprave izgubijo nadzor nad podatki in se izgubijo v mednarodnem informacijskem konglomeratu. Drugi je scenarij »velikega brata«: »evropske združene države« uporabljajo prostorske podatke za obrambo držav in državljanov pred kriminalom in uničevalnimi aktivnostmi ilegalnih emigrantov in davčnih prestopnikov.

Prostorske podatke je potrebno obravnavati kot premoženje, ki ga je treba upravljati v nacionalnem interesu (Ordinance Survey, 1996). V smislu zajemanja prostorskih podatkov je mogoče opredeliti tri ločene sestavine nacionalnega interesa: javni interes, koristi in splošna uporabnost. Javni interes se izkazuje pri kartiranju območij države, kjer sicer podatki ne bi bili zajeti. Koristi, ki izhajajo iz doslednih, vsebinsko, ažurno, stilno in metodološko skladnih podatkov, kar predpisujejo potrebe, so drugačne od potreb trga podatkov. Primer koristi so prostorski podatki za potrebe nacionalne obrambe in splošne koristi uporabe temeljnih zbirk prostorskih podatkov kot podlag, ki omogočajo prostorsko združevanje zbirk podatkov.

Splošna uporabnost podporne infrastrukture, kot je državna geodetska mreža, državne GPS-mreže, sistem državne razdelitve na liste, obsega zajem interoperabilnih podatkov o prostoru,

ki jih uporabljajo vse državne ustanove in javnost brez plačevanja prispevka. Zbirke prostorskih podatkov kot državno premoženje so pod vplivom dveh ravni upravljanja. Prva raven je skrbništvo nad zbirko podatkov, ki daje odgovornost za njeno kreiranje in vzdrževanje zbirke podatkov. Druga raven pa je nadzor nad skrbnikom zbirke podatkov, ki zagotavlja, da so aktivnosti skrbnikov v skladu z nacionalnim interesom.

Skrbnik v imenu države upravlja zbirko podatkov, katere lastnik je država. Zaradi razširljivosti in možne integracije različnih zbirk prostorskih podatkov je mogoče opredeliti tri ključne sestavine državne informacijske strategije za prostorske podatke: razvoj metapodatkovnih storitev za podporo medopravilnosti; koordinacija med skrbniki zbirk podatkov v smislu razvoja navodil in standardov; opredelitev in razlikovanje referenčnih in izpeljanih zbirk podatkov prostorskih podatkov. To omogoča medopravilno kombiniranje porazdeljenih zbirk podatkov za doseganje največje možne učinkovitosti vzdrževanja in integrirane uporabe zbirk prostorskih podatkov. Vsakdo, ki izvaja določene aktivnosti v zvezi z informacijami geografskega prostora, se sklicuje na posebne podatke in zbirke podatkov, ki se imenujejo referenčni podatki oziroma referenčne zbirke podatkov. Referenčni podatki in referenčne zbirke podatkov nudijo prostorski vidik oziroma prostorsko strukturo za izvajanje aktivnosti v geografskem prostoru kakor tudi mehanizem za povezovanje oziroma integracijo zbirk podatkov.

Zaradi njihovega strateškega pomena referenčne zbirke prostorskih podatkov predstavljajo, poleg metapodatkovnih storitev in koordinacije med skrbniki zbirk podatkov, tretjo temeljno sestavino državne informacijske strategije za prostorske podatke. Vnaprej je treba identificirati ključne, temeljne zbirke prostorskih podatkov državne infrastrukture in izpeljane zbirke prostorskih podatkov državne infrastrukture. Referenčne zbirke prostorskih podatkov so temelj za izvedbo izpeljanih zbirk podatkov državne infrastrukture kakor tudi temelj za različne medopravilne aplikacije. Globalna prostorskopodatkovna infrastruktura (Global Spatial Data Infrastructure - GSDI) je opredeljena kot koordinirana aktivnost držav in organizacij, ki spodbujajo uporabo mehanizmov in skupnih standardov za razvoj in medopravilnost digitalnih geografskih podatkov in tehnologij za podporo večnamenskemu odločanju (Stensen, 2001).

## 7.5 OPREDELITEV METOD MEDOPRAVILNEGA POVEZOVANJA ZBIRK PODATKOV IN GIS

Cilj povezovanja oziroma integracije zbirk prostorskih podatkov v okolju neodvisno porazdeljenih informacijskih sistemov sta sodelovanje – kooperacija – in medsistemska izvedljivost – medopravilnost – med informacijskimi vozlišči. Za izvedbo sodelovanja med neodvisnimi informacijskimi vozlišči je pomembna informacija o stopnji povezljivosti izbranih zbirk podatkov oziroma smiselnosti povezovanja konceptov. Zato smo razvili metodo ocenjevanja skladnosti med koncepti posameznih katalogov prostorskih podatkov, ki temelji na predhodno razviti ontologiji prostora geografskih razsežnosti ter metodah ocenjevanja semantike zbirk prostorskih podatkov.

Metoda ocenjevanja skladnosti med koncepti posameznih katalogov prostorskih podatkov Posredovalno, semantično referenčno omrežje, ki omogoča prenašanje pomena med podatkovnimi zbirkami in je bilo razvito v prejšnjih razdelkih, predstavlja neobhodno osnovo za ocenjevanje skladnosti konceptov posameznih katalogov prostorskih podatkov. V ta namen razvita metoda je sestavljena iz dveh zaporednih faz:

- ocenjevanje skladnosti semantičnih meta podatkov primerjanih zbirk podatkov,
- ocenjevanje skladnosti konceptov primerjanih zbirk podatkov.

### 7.5.1 Ocenjevanje skladnosti semantičnih meta podatkov primerjanih zbirk podatkov

Za ocenjevanje skladnosti semantičnih meta podatkov primerjanih zbirk podatkov uporabimo podatke o absolutni semantični globini, relativni prostranosti, relativnem razsipanju in semantični razpršenosti obeh zbirk podatkov, katerih opredelitve smo izračunali v metodi ocenjevanja semantike zbirk prostorskih podatkov. Če so podatki relativno skladni, nadaljujemo delo z drugo fazo, in sicer z ocenjevanjem skladnosti konceptov primerjanih zbirk podatkov. Če podatki ne izkazujejo relativne skladnosti, opustimo izvedbo druge faze, razen če iščemo razmerje med podatkovnima zbirkama za posebej izbran koncept. V tem primeru obstaja možnost, da takšno razmerje odkrijemo, ne glede na relativno semantično neskladnost primerjanih zbirk podatkov.

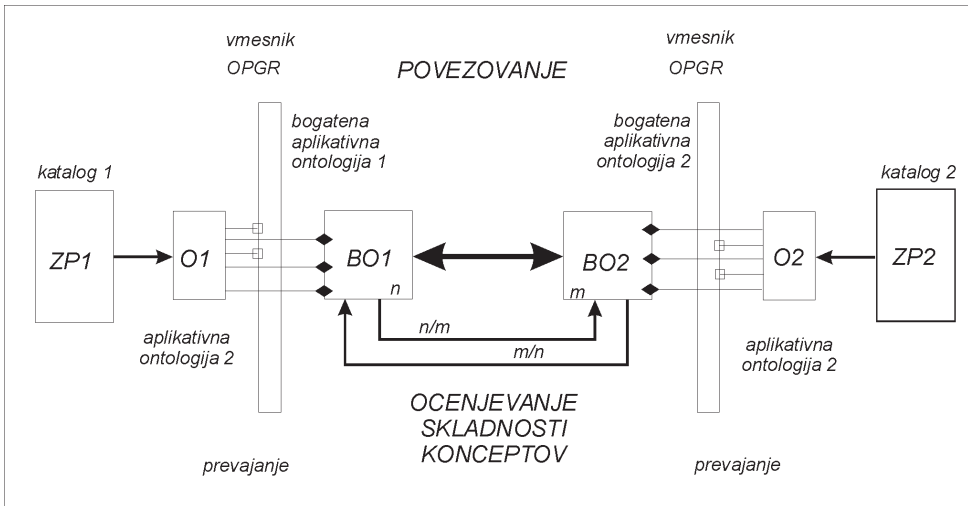
### 7.5.2 Ocenjevanje skladnosti konceptov primerjanih zbirk podatkov

Za metodo ocenjevanja skladnosti konceptov primerjanih zbirk podatkov smo razvili postopek primerjanja konceptov zbirk podatkov ter postopek izvedbe ocene povezljivosti dveh zbirk podatkov. V metodo smo privzeli opredelitve stopenj semantičnih razmerij »enakovredno«, »sorodno«, »ustrezno« ter opredelitev »ni razmerja« po Shethu in Kashyapu (1993) in temu dodali še stopnjo »ni izražen z OPGR« (preglednica 18).

Stopnje	Semantična razmerja	Angleški izraz	Opis
1	enakovredno	equivalent	Objektni razredi različnih aplikacijskih ontologij so semantično enakovredni (ekvivalentni), če ustrezajo istim razredom v ontologiji domene.
2	sorodno	related	Objektni razredi različnih aplikacijskih ontologij so semantično sorodni (related), če ustrezajo razredom v ontologiji domene, ki so drug drugemu podrazred ali nadrazred.
3	ustrezno	relevant	Objektna razreda A in B iz različnih aplikacijskih ontologij sta semantično ustrezna (relevant), če je razred A enakovreden z razredom v ontologiji domene, ki predstavlja en del sestavljenega razreda B v ontologiji aplikacije.
4	ni razmerja	irrelevant	
5	ni izražen z OPGR		

Preglednica 18: Lestvica možnih semantičnih razmerij za ocenjevanje skladnosti konceptov različnih zbirk podatkov

Postopek primerjanja konceptov zbirk podatkov (ZP1 in ZP2) in postopek izvedbe ocene povezljivosti dveh zbirk podatkov je grafično prikazan na sliki 25. Iz zbirk podatkov sta izvedeni aplikativni ontologiji O1 in O2. Koncepti aplikativnih ontologij so prevedeni in kartirani na vmesniku OPGR in s tem tudi semantično obogateni. Nekateri koncepti obeh ontologij z vmesnikom niso skladni in so zato iz povezovanja izključeni. Prevedeni koncepti so vključeni v postopek ocenjevanja skladnosti



Slika 25: Ocenjevanje skladnosti konceptov in povezovanje katalogov dveh zbirk prostorskih podatkov

obeh obogatenih aplikativnih ontologij in ustrezno povezani med seboj.

### 7.5.3 Metoda izvedbe semantične povezave skladnih konceptov zbirk prostorskih podatkov

Metoda izvedbe semantične povezave skladnih konceptov zbirk prostorskih podatkov je končni postopek za vzpostavitev stvarne povezave med koncepti obravnavanih zbirk prostorskih podatkov. Metoda temelji na predhodno izvedenih metodah metodologije posrednega povezovanja zbirk prostorskih podatkov.

Stopnje skladnosti konceptov, določene v drugi fazi metode ocenjevanja skladnosti med koncepti posameznih katalogov prostorskih podatkov, so osnova za odločanje o vzpostavitvi povezave med koncepti obravnavanih zbirke podatkov. Najprimernejša stopnja skladnosti za povezovanje konceptov je stopnja »enakovredno« (1), sledita ji stopnja »sorodno« (2) ter stopnja »ustrezno« (3). Za samodejno usklajevanje podatkovnih objektov in njihovo samodejno povezovanje je primerna le stopnja skladnosti »enakovredno« (1).

Za izvedbo semantične povezave skladnih konceptov zbirk prostorskih podatkov smo razvili dva postopka. Osnovni postopek je izvajanje metode semantične povezave, ki po zaključku omogoča vključitev mehanizmov za samodejno povezovanje pomensko enakovrednih podatkov. Vloge ontologov, predvidene v postopku, so v primeru združevanja zbirke podatkov, ki so v pristojnosti istega skrbnika podatkovnih zbirke, lahko združene v eno samo vlogo.

1. Navedi koncept iz obravnavane zbirke podatkov.
2. Določi simbole za iskani koncept.
3. Določi sinonime simbolov v domačem jeziku.
4. Določi sinonime simbolov v tujem jeziku (v primeru mednarodnega poizvedovanja).
5. Izberi administrativno prostorsko enoto (državo, regijo, lokalno skupnost).
6. Izberi želeno stopnjo povezljivosti (kriterija semantične globine in semantične razpršenosti).
7. Izberi vrsto zbirke podatkov (atributna tabela, GIS-tematski sloj).

8. Sproži iskanje ustreznih zbirk podatkov.
9. Analiziraj semantiko konceptov odkritih zbirk podatkov – navedene meta podatke o semantiki konceptov.
10. Začni pogajalski postopek za izvedbo povezave.
11. Vzpostavi povezavo.
12. Vzdržuj povezavo.

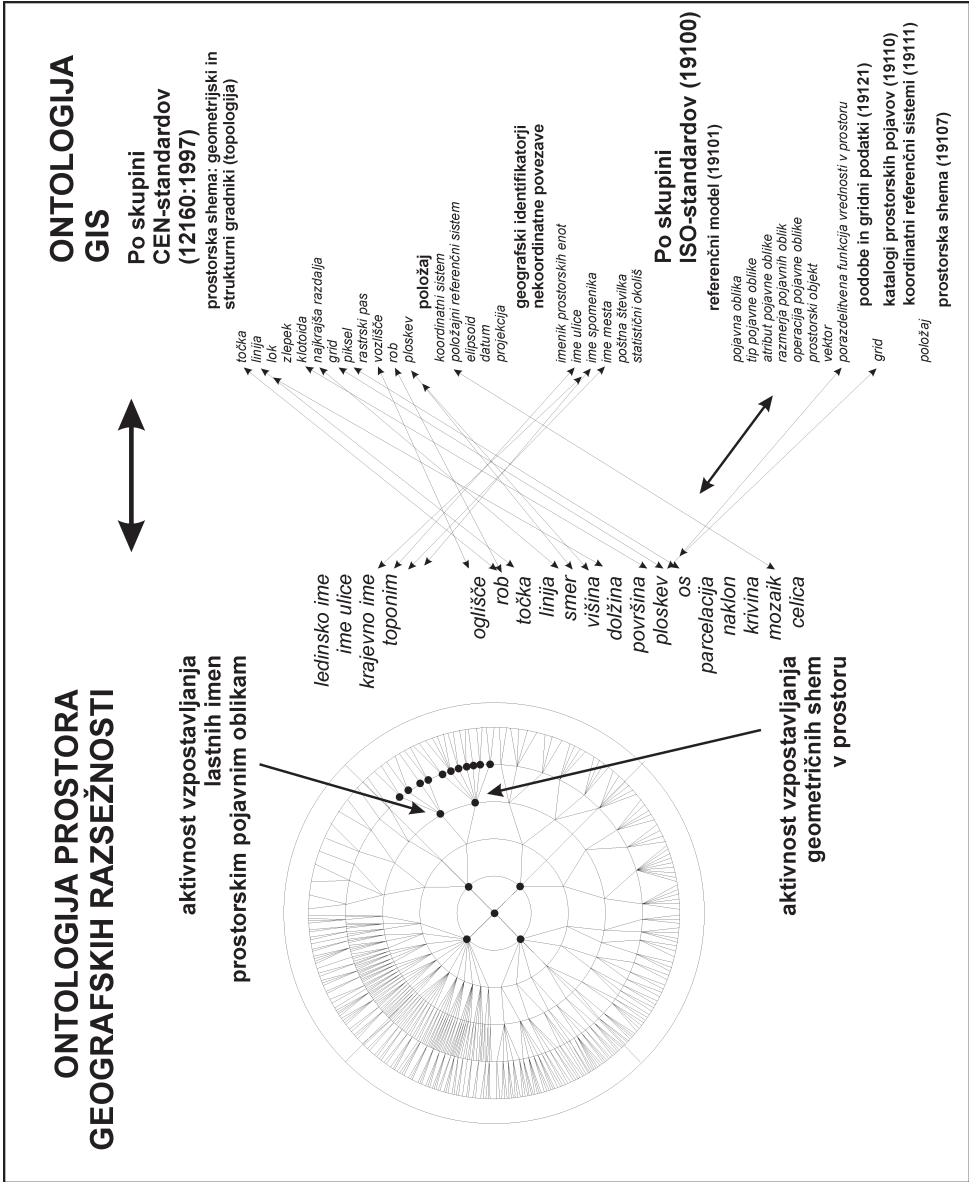
Izvedba semantične povezave v okolju globalne prostorsko-podatkovne infrastrukture je predvideni postopek uporabe metode za izvedbo semantičnega povezovanja zbirk prostorskih podatkov za GIS-uporabnika z vidika združevanja vseevropske prostorskopodatkovne infrastrukture v zvezno porazdeljenih informacijskih sistemih.

#### **7.5.4 Povezovanje konceptov geografskega prostora in konceptov tehnologije GIS**

Za uspešno obravnavo medopravnosti geografskih informacijskih sistemov ter zbirk prostorskih podatkov je potreben celostni pristop, ki obravnava tako geografski prostor kakor tudi geoinformacijsko tehnologijo.

Predlagamo model povezovanja, ki ločeno izraža semantično referenčno omrežje prostora ter koncepte geoinformacijske tehnologije z dvema neodvisnima in delno povezljivima ontologijama. Semantično referenčno omrežje geografskega prostora je predstavljeno z ontologijo prostora geografskih razsežnosti – OPGR. Geoinformacijska tehnologija je predstavljena z ontologijo GIS – OGIS, katere koncepte privzemamo iz skupin standardov, in sicer CEN 12160:1997 (prostorska shema, geometrijski in strukturni gradniki – topologija, položaj, geografski identifikatorji – nekoordinatne povezave) in skupine standardov ISO referenčni model (19101), podobe in gridni podatki (19121), katalog prostorskih podatkov (19110), koordinatni referenčni sistemi (19111) ter prostorska shema (19107).

Za povezovanje tipoloških prostorskih konceptov podatkovnih modelov OGIS s koncepti OPGR je odločilen koncept iz OPGR, ki ga imenujemo »aktivnosti vzpostavljanja geometričnih shem v prostoru«. Skupina pripadajočih podrazredov – konceptov točka, linija, območje, celica mozaika (raster), predstavlja temeljno povezavo splošnih konceptov prostora s tehnološkimi koncepti informacijske tehnologije geografskih informacijskih sistemov oziroma ontologije GIS (slika 26). Za opisno podajanje prostorske lokacije v nekoordinatnih prostorskih referenčnih shemah je pomemben koncept iz OPGR z imenom »vzpostavljanje lastnih imen prostorskim pojavnim oblikam«. Semantično referenčno omrežje geografskega prostora oziroma ontologija prostora geografskih razsežnosti, zasnovana na pojmih namena človeških aktivnosti v prostoru, je večstransko uporabna zasnova. V prvi vrsti jo odlikuje neodvisnost od posameznih prostorskih objektov in tako predstavlja parameter za izračun relativnih semantičnih razdalj oziroma ocenjevanje relativne semantične skladnosti zbirk prostorskih podatkov. Uporabna je tudi kot ogrodje za povezovanje podatkovnih objektov različnih zbirk podatkov, v najboljšem primeru za samodejno usklajevanje podatkovnih objektov in njihovo medopravilno povezovanje.



Slika 26: Povezovanje konceptov geografskega prostora in konceptov tehnologije GIS

## ZAKLJUČEK

Upravljanje zbirk prostorskih podatkov v neodvisno porazdeljenih informacijskih sistemih je stvarnost. Učinkovito samodejno povezovanje tovrstnih zbirk podatkov je prihodnost. Strukturna medopravilnost sistemov za upravljanje prostorskih podatkov je predmet tehnološkega razvoja, ki že ponuja rešitve v obliki standardov za geografske informacijske sisteme. Semantična medopravilnost znanja o prostoru pa je predmet intenzivnih raziskav. Semantični vidik inteligentne integracije informacij vključuje izdelavo sistema znanja za integracijo podatkov. Gre za ohranjanje pomena informacijskih entitet v različnih kontekstih, kar je mogoče obravnavati kot klasifikacijsko nalogo. S širjenjem GIS-aplikacij na svetovnem spletu se povečuje potreba po prepoznavanju pristopov, s katerimi različne skupine uporabnikov konceptualizirajo domeno prostora geografskih razsežnosti. Potrebno je razumeti, kako uporabniki uspejo oziroma zakaj ne uspejo izmenjavati prostorske informacije, bodisi v osebni komunikaciji bodisi pri komunikaciji med računalniki. Namen integracije geografskih podatkovnih nizov je neovirano porazdeljevanje informacij med različnimi viri geografskih informacij. Še posebej je pomembno preučevanje integracije geografskih podatkovnih nizov v smislu porazdeljevanja obnovljenih podatkovnih nizov, to je večkratne uporabe obnovljenih podatkov v več različnih zbirkah podatkov.

GIS bi moral nuditi podporo izvajanju človeških aktivnosti. Namesto tega je GIS v večini primerov razvit kot pasivni model sveta s premajhnim upoštevanjem okolja, v katerem se uporablja. Predpostavlja se, da postanejo geografske informacije uporabne same po sebi, če so le dovolj verodostojna predstavitev stvarnosti. Da bi postale bolj koristne in uporabne, bi bilo potrebno razviti metodo za gradnjo ontologije z osredotočenjem na človeške aktivnosti v prostoru. Potrebno je prepoznati, kako določeni agenti delujejo v svetu obravnave, in aktivnosti, ki se pojavljajo v svetu, neodvisno od načina modeliranja človeškega znanja o svetu v informacijskem sistemu.

Za teorijo prostorskih aktivnosti sta najpomembnejši sledeči pravili: (1) semantika aktivnosti in objektov je neločljiva, (2) aktivnosti in objekti so hierarhično strukturirani. Prepoznavanje konceptov, ki naj bi bili vključeni v ontologijo domene, je mogoče izvajati z analiziranjem tekstualnih opisov aktivnosti. Prostorska integracija podatkov temelji na zanesljivem geometričnem referenčnem sistemu in na razumni ravni ujemanja podrobnosti različnih podatkovnih nizov. Semantična integracija geografskih podatkovnih nizov – integracija kart – je postopek vzpostavljanja razmerij med odgovarjajočimi primeri objektov za različne, neodvisno izdelane geografske podatkovne nize določenega območja.

Semantični vidik geografskih podatkovnih nizov je odvisen od namena aplikacije, za katero so bili podatkovni nizi izdelani. Isti geografski pojav/objekt je v različnih izvedbenih področjih lahko različno poimenovan. Še posebej je pomembno preučevanje integracije geografskih podatkovnih

nizov v smislu neoviranega porazdeljevanja obnovljenih podatkovnih nizov, to je večkratne uporabe v več različnih zbirkah podatkov. Organizacijska narava kognitivnega kartiranja je v iskanju pomena stvari. Za organiziranje podatkov o izkušnjah v prostoru je potreben čitljiv okvir, ki mora biti nepretrgan in sestavljen iz razpoznavnih, a povezanih delov. Del organizacijskih povezav izkušenj v prostoru je odvisen od svojstvenih lastnosti prostora kot samoorganizacije. Preostali del okvira je prežet z osebnim pomenom. Kognitivne predstavitve prostora tako postanejo polnopomenske. Kognitivno kartiranje je miselna obdelava prostorskih informacij, miselna sposobnost pojmovnega zbiranja, organiziranja, priklicevanja in upravljanja informacij o prostoru. Teorija kognitivnega kartiranja temelji na hierarhiji predstavitev prostorskega znanja, ki se imenuje prostorska semantična hierarhija. Vsaka raven hierarhije ima svojo lastno ontologijo (niz objektov in relacij za opisovanje sveta) in svoj niz metod reševanja problemov in sklepanj. Za izražanje znanja o prostoru je mogoče izbirati med dvema sredstvoma, in sicer med podobami in besedami. Za zunanje predstavitve miselnega prostorskega znanja s slikovnim sredstvom izražanja se uporablja niz simbolov, ki niso enaki nizu kognitivnih simbolov v interni predstavitvi v človeškem umu. Tudi za besedno izražanje obstaja omejen niz simbolov, ki omejuje izražanje internih kognitivnih simbolov navzven. Če oseba izvede miselni prevod iz enega v drugo izrazno sredstvo, rezultat verjetno predstavlja spretnost te osebe za izvedbo prevoda, ne pa le načina predstavitve znanja o prostoru. Problem enoličnega prevajanja postane enostavnejši, če se obravnava interno prostorsko znanje, informacije kot slovar – seznam besed, interne simbole pa kot slovnična pravila.

Znanje o prostoru in njegovo klasifikacijo je mogoče obravnavati z več klasifikacijskimi razsežnostmi, kot so: vrsta prostorskega znanja (deklarativna, postopkovna, konfigurativna), viri prostorskih informacij za kognitivno kartiranje, človeška interakcija s svetom in prirejena uporaba jezikov. V zgodovini metafizike obstaja večno nasprotje med stvarmi, ki so najdene oziroma odkrite, in med narejenimi oziroma ustvarjenimi stvarmi. Za izdelavo ontologije geografskega prostora je mogoče izbirati med več doktrinami: med skrajnim idealizmom, zmernim realizmom in skrajnim realizmom. Doktrina mereološkega aktualizma razlikuje dve osnovni vrsti prostorskih objektov: naravne (bona-fide) ter umetne (fiat) prostorske objekte. Prostorski fiat objekti so posledica človeških kognitivnih aktivnosti in delovanja. Prostorski objekti geografskih razsežnosti so objekti, ki so večji od človeškega telesa in jih v celoti ni mogoče zaznati v okviru enega zaznavnega dejanja. Nepredmetni objekti so ustvarjeni v človeškem umu. Predmetni objekti obstajajo v fizični stvarnosti neodvisno od človeškega uma.

Razlikovati je mogoče dve skupini predmetnih objektov/pojavov geografskega obsega: predmetni objekti/pojavi z določenimi mejami, predmetni objekti/pojavi z nedoločenimi mejami. Lokacij pojavov z nerazločnimi ali nedoločenimi mejami ni mogoče natančno izmeriti, kajti meje so nerazločne, nedoločene, se premikajo oziroma jih ni mogoče opazovati. V zvezi z ontološkim in epistemološkim ujemanjem je mogoče ločevati med formalizacijo in predstavnostjo natančnih in približnih lokacij. Pojem lokacije se nanaša na razmerja med posameznimi enotami, individuumi in referenčnim okvirom. V primeru natančne lokacije vsebujejo vsi ustrezni razredi iste lokacije samo posameznike, ki so natančno solocirani. Solocirani so lahko le objekti, ki so ontološko različni.

Sestavine formalnega pristopa izgradnje integracijskega sistema so ontologije topografskega kartiranja – ontologija domene posebne ontologije za vsak podatkovni niz – aplikacijske ontologije, ter kriteriji za zajemanje podatkov oziroma snemalna pravila. Snemalna pravila pojasnjujejo načela transformacije objektov stvarnega sveta v objekte geografskih podatkovnih nizov. Podobnosti v snemalnih pravilih kažejo na različna skladnostna razmerja na ravni konceptov, torej med ontologijo domene in aplikacijskimi domenami. S prepoznavanjem teh razlik v pravilih je mogoče na ravni objektnih primerov določiti semantično podobnost. V ta namen se uporabljajo metode računske geometrije, na primer operacija prekrivanja.



Porazdeljevanje in večkratna uporaba obnovljenih podatkov v različnih podatkovnih nizih predstavlja komunikacijski problem. Izdelava sistema, ki bi med odgovarjajočimi, semantično podobnimi primeri geografskih objektov zmanjšal operaterjeva posredovanja pri operacijah vzdrževanja zbirke podatkov, je zapletena. Temeljni predpogoj za izdelavo takšnega sistema je pomenska podobnost različnih predstavitev objektov.

Izvedena je bila opredelitev ontologije prostora geografskih razsežnosti (OPGR), na osnovi katere je mogoče izvajati pomensko integracijo podatkov zbirke prostorskih podatkov. V navedenem konceptualnem omrežju je pomen besed, s katerimi so označeni prostorski objekti v stvarnosti, pojasnjen z uvrščanjem izrazov o objektihi v konceptualno omrežje prostora geografskih razsežnosti. OPGR vsebuje tudi semantični slovar. Pri izdelavi modela konceptualnega omrežja smo upoštevali izsledke raziskav ontološkega modeliranja. V nadaljevanju smo izvedeli opredelitve mer za ocenjevanje semantike zbirke prostorskih podatkov, in sicer semantične globine prostorskega koncepta, semantične globine zbirke prostorskih podatkov, semantične razdalje dveh konceptov, semantične razpršenosti zbirke prostorskih podatkov, semantične skladnosti zbirke prostorskih podatkov, kakovosti predstavitve semantike zbirke prostorskih podatkov, semantične ustreznosti ontologije in semantične pristranosti koncepta. Izvedena je analiza semantičnih razmerij aplikativnih konceptov v okviru OPGR in izračunana stopnja ustreznosti OPGR za obravnavane preizkusne zbirke podatkov. Izdelan je bil nov analitični sistem za primerjanje zbirke prostorskih podatkov – metodologija posredne primerjave konceptualnih shem zbirke prostorskih podatkov. Gre za primerjalno metodo analiziranja zbirke podatkov, ki je oblikovana v naslednjem postopku:

- izdelava univerzalnega konceptualnega omrežja prostorskega znanja (univerzalne ontologije prostora geografskih razsežnosti);
- izdelava aplikativnih ontologij na osnovi shem podatkovnih modelov (objektnih katalogov);
- prevajanje – translacija – konceptov aplikativnih ontologij v prevedeno aplikativno ontologijo, izražanje konceptov aplikativne ontologije s koncepti univerzalne ontologije;
- določitev stopnje povezljivosti aplikativnih ontologij z univerzalno ontologijo (enakovredno, agregatno, nepovezljivo);
- določitev stopnje semantične povezljivosti aplikativnih, primerjanih ontologij;
- bogatitev slabo povezljivih aplikativnih konceptov in ponovna analiza;
- neposredna povezava dobro povezljivih aplikativnih konceptov.

Za integracijo konceptualnih shem zbirke podatkov se v splošnem izdeluje globalna konceptualna shema na osnovi preučevanja več lokalnih shem. Od takšnega pristopa se razvita metoda razlikuje po vnaprejšnji izdelavi globalne konceptualne sheme, ki temelji na neodvisni in podrobni proučitvi domene obravnavanega področja, in sicer domeni prostora geografskih razsežnosti ter izdelavi podrobnega univerzalnega omrežja konceptov prostora. Ker je metoda izvedena na izviren način s kombinacijo štirih osnovnih konceptov, te koncepte na kratko povzemam:

- prostor geografskih razsežnosti,
- človeške aktivnosti,
- nameni človeških aktivnosti (prvobitni, napredni),
- pojavne oblike (predmetni – fizični – objekt oziroma pojmovna shema), ki jih v prostoru ustvarijo oziroma vzpostavijo človeške aktivnosti.

Ugotovili smo, da je razviti model OPGR zelo zanesljiv in dovolj obsežen za namene semantičnega izražanja prostorskih konceptov. Rezultati poskusov so pokazali, da je štiriindevetdeset odstotkov

(94 %) vseh konceptov obravnavanih katalogov prostorskih podatkov mogoče izraziti s koncepti ontologije prostora geografskih razsežnosti. Potrjena je vrednost neodvisne zasnove OPGR in s tem njene univerzalnosti za domeno prostora geografskih razsežnosti, še posebej, ker ontologija ni bila izdelana kot spoj obravnavanih zbirk podatkov (lokalnih ontologij), ampak kot neodvisna zasnova. Zaradi visoke stopnje semantične ustreznosti je OPGR primerno orodje za ugotavljanje semantične skladnosti konceptov raznovrstnih zbirk prostorskih podatkov. V analizi smo ugotovili semantično pristranost nekaterih konceptov obstoječih zbirk podatkov, ki se nanaša na metodo zajema prostorskih podatkov. Razvita metoda posredne primerjave konceptualnih shem zbirk prostorskih podatkov je bila izvedena za omejeno število aplikativnih katalogov obstoječih zbirk prostorskih podatkov.

V raziskavah samodejne integracije informacij z vidika semantike podatkov je cilj ohranitev pomena informacijskih entitet v različnih informacijskih okoljih. Prvobitni vir pomena sta stvarni svet in znanje zastopnikov (agentov), ki uporabljajo znake za predstavitev entitet stvarnega sveta in njihovih namenov. Medopravilnost na podatkovni ravni pomeni skladno interpretacijo pomena podatkov pri povezovanju zbirk podatkov – semantično medopravilnost. Kriteriji za klasifikacijo podatkovnih virov za namene integracije podatkov so neodvisnost, raznovrstnost, porazdeljevanje in odkritost podatkovnih modelov. Pri praktičnem izvajanju konceptualnega modeliranja obstaja problem, da naravni jeziki ne zagotavljajo ustreznih besed za vse koncepte na želeni ravni abstrakcije.

Struktura jezika, razmerja med stvarnostjo in jezikom ter dejanska uporaba jezika so še posebno pomembni v smislu komunikacije, zato so jezikoslovne teorije pomemben temelj konceptualnega modeliranja. Semiotika je veda o znakih in je del teorije informacij, ki obravnava komunikacijo med različnimi vrstami agentov (ljudmi, živalmi, računalniki) preko jezika in drugih simbolnih sistemov (alfanumeričnih, grafičnih). Za učinkovito komunikacijo pomena simbolov se uporabljajo opredelitve. Opredelitev je postopek (govorjen ali izveden kako drugače), v katerem se posamezniki med seboj obvestijo o pomenu obravnavanega simbola. Opredelitev navaja, na kateri koncept se nanaša simbol. Vedno opredeljuje simbol, nikoli koncepta. Sredstvo za poenotenje predstavitev pomena je logika, ki je lahko izvedena v različnih podnizih različnih programskih orodij. Semantična dvoumnost simbolov izvira iz dovoljenih večpomenskih (polisemičnih) izrazov v zbirki znanja, kjer je en izraz povezan z več pomeni. Večpomenski izrazi se pojavijo celo, če se klasifikacija uporablja dosledno. Reševanje dvopomenskosti izrazov je mogoče razrešiti z doslednim preverjanjem konteksta sheme, ontologije.

Semantika se v nasprotju s sintakso, ki obravnava le obliko izražanja v jeziku, nanaša na pomen izrazov v jeziku ali logiki. Pet semantičnih prvin in različne mehanizme za opredeljevanje drugih operatorjev logike prvega reda je mogoče prirediti veliko različnim načinom zapisovanja, vključno z naravnimi jeziki, podatkovnim standardom XML in orodjem za opisavanje virov podatkov RDF, XML in RDF omogočata jasno ločevanje med formatiranjem in pomenom besedila. Oba za predstavitev pomenov uporabljata oznake oziroma etikete v besedilu.

Konceptualne sheme so strukture, ki predstavljajo odnose med objekti v stvarnem svetu oziroma nominalni osnovi. Raziskovanje integracije shem se sooča s problemi pridobivanja in predstavitev semantičnega znanja. Pridobivanje semantičnega znanja je povezano z odkrivanjem semantike konceptov, odkrivanjem semantične podobnosti in protislovij, izdelavo arhitekture za orodja, ki so namenjena integraciji shem, in izdelavo semantičnih slovarjev.

Semantična relativistična teorija predpostavlja, da je mogoče z uporabo pripadajočih semantičnih razmerij identificirati koncept in druge koncepte, ki so s tem v odnosu. Semantična sorodnost je združljivost, pri kateri se dva ali več konceptov različnih shem nanašata na enake ali podobne objekte v stvarnosti. Semantična sorodnost je pri integraciji shem zbirk podatkov opredeljena s klasifikacijo semantičnih razmerij: istovetno (natančno enak), enakovredno (uporaba neenakih, a

enakovrednih sestavin ter uporaba istih spoznanj), skladno (neenakovredne vendar uporabljene sestave in nenasprotujoča si spoznanja), neskladno (vsi elementi so si nasprotujoči). Odkrita nesoglasja se poskušajo urediti z določenimi integracijskimi orodji, če so na razpolago. Neskladja, ki jih ni mogoče urediti z integracijskimi orodji, se predstavijo uporabnikom, da ti sami sprejmejo odločitve o njihovem reševanju.

Za samodejno integracijo konceptualnih shem zbirk podatkov je potrebno zgraditi semantični slovar in opredeliti kriterije za presojo semantične podobnosti. Primerjanje shem zadeva tolmačenje modelov stvarnega sveta. Modele ustvarijo različni ljudje z različnimi nameni. Zato bi bila popolna avtomatizacija postopkov primerjave shem neutemeljena. Metode merjenja semantične podobnosti določajo semantične razdalje med primerjanimi koncepti. Za izdelavo semantičnega slovarja je potrebna osnova, ki se imenuje semantično omrežje. To je sestavljena predstavitev semantičnih razmerij med objekti ali pojavi v stvarnem svetu. Konceptualni graf je struktura konceptov in konceptualnih razmerij, kjer vsaka povezava povezuje konceptualno vozlišče in vozlišče konceptualnega razmerja in pogosto služi kot semantično omrežje. Za opisovanje niza konceptov in razmerij med koncepti je prikladna uporaba hierarhije. Vsem metodam primerjanja shem je skupno, da informacije, ki so vsebovane v lokalnih shemah, niso zadovoljive za izdelavo zanesljivega niza povezovalnih trditvev. Zato se v vse novejšie tehnike primerjanja shem uvajajo metode semantičnega bogatenja lokalnih shem.

Pri semantični bogatitvi shem za primerjavo se jasno poudarijo različne vrste skritih informacij, ki zadevajo semantiko modeliranih objektov/pojavov stvarnega sveta. Metode za semantično bogatitev shem vključujejo uporabo dodatnih virov znanja, poleg tega, ki ga vsebujejo sestavine konceptualnih shem. Če so sheme obdelane z enako tehniko oziroma metodo obogatitve, potem semantična obogatitev navidezno vsili globalni pogled lokalnim shemam. Obogatene lokalne sheme so namreč pogosto povezane s splošnim modelom oziroma shemo, povezani skupaj pa tako tvorijo enotno navidezno globalno shemo. Semantično bogatenje je pomembno tudi za druge faze integracije shem. Uporabno je pri izboljšavi kakovosti integracije kakor tudi pri spajanju shem. Kakovost integracije tako izpostavlja vprašanje načina testiranja kakovosti integracije konceptualnih shem. Za ocenjevanje se predpostavljajo kriteriji dovršenosti integracije ter minimalnosti in razumljivosti integralne sheme.

Predlog nadaljnjih raziskav obravnavanega področja, to je odkrivanja semantične sorodnosti podatkov v raznovrstnih zbirkah prostorskih, podatkov podajamo v nadaljevanju. Smiselno bi bilo izvesti preizkus uporabnosti metode vključevalnih pravil. Prednost navedene metode je zmanjšanje potrebe po skupnem jeziku, ki omejuje uporabo semantičnih metod za analizo podatkov; metoda se temu izogne z uvedbo spoznanj teorije kognitivne znanosti, kar je značilno tudi za to metodo. Potrebno bi bilo preveriti rezultate semantične integracije podatkov na grafični ravni z obsežnejšim prekrivanjem podatkovnih slojev posameznih kategorij zbirk podatkov, ki so vključene v postopek integracije. Model OPGR oziroma znanje, vsebovano v modelu, bi bilo smiselno prevesti v ekspertni sistem, s katerim bi uporabniki določali prostorsko semantiko obravnavanih izrazov; z uporabo hiperboličnega pregledovalnika konceptov bi ontologijo OPGR približali širši strokovni javnosti, ki bi z uporabo ontologije preizkusila njeno dejansko uporabnost. Za analizo kognitivne ustreznosti modela OPGR bi bila priporočena izvedba testiranja za uvrščanje izrazov v semantično omrežje prostora s strani nestrokovnjakov. Proučiti bi bilo treba potencialno uporabnost modela OPGR za podporo pri odločanju pri samodejni kartografski generalizaciji. Podrobneje bi bilo treba proučiti povezljivost modela OPGR s standardno klasifikacijo dejavnosti (SKD) in možnosti integrirane uporabe.

V Sloveniji bi bilo smiselno izdelati analizo semantične skladnosti vseh obstoječih referenčnih zbirk prostorskih podatkov. Tako bi pridobili dodatni meta podatkovni sloj za vsako jedrno zbirko

podatkov, s čimer bi temeljito obogatili ponudbo na trgu prostorskih podatkov. Obenem bi dosegli povečan izkoristek zbirk podatkov v skrbništvu državnih agencij, kar je del vsake strategije državne prostorskopodatkovne infrastrukture. S tem bi se povečala gospodarnost upravljanja zbirk prostorskih podatkov. Z objavo semantičnih meta podatkov o zbirkah prostorskih podatkov bi bilo olajšano presojanje povezljivosti zbirk podatkov oziroma posameznih konceptov, ki bi bili predstavljeni v zbirkah podatkov.

Z vidika obnavljanja podatkov o prostoru z metodami terenskega zajema podatkov so možnosti za optimizacijo in racionalizacijo še nepresežene. S povezavo neodvisnih, a semantično sorodnih zbirk podatkov bi bili lahko doseženi pomembni prihranki tako v časovnem kot v finančnem smislu. Glede tehnologije samodejnega povezovanja zbirk podatkov je pričakovati hiter napredek, še posebej zaradi gibanj za standardizacijo v geoinformacijskih skupnostih.

Smiselno bi bilo proučiti tudi zbirke prostorskih podatkov, ki niso deklarirane kot referenčne zbirke podatkov državnega pomena, saj se mnoge od njih stalno vzdržujejo in redno obnavljajo. Pri tem so mišljene predvsem različne zbirke prostorskih podatkov in katastrov (komunalnih in drugih) ter lokalnih GIS-sistemov, ki postajajo vse bolj nepogrešljivo orodje za upravljanje lokalnih skupnosti. Zgrajena ontologija prostora je potencialno sredstvo za izvedbo primerjalnih analiz podatkovnih infrastruktur različnih držav in izvedbo meddržavnih, na primer vseevropskih, povezav za zbirke prostorskih podatkov. V bodočem razvoju geografskih informacijskih sistemov bo predstavljala osrednji del skladna in empirično utemeljena ontologija prostorskih pojavov geografskih razsežnosti.

## PRILOGE

Ontološki vmesni sloj (The Ontology Inference Layer – OIL) je predlog standarda za integriranje ontologij z obstoječimi in novimi spletnimi standardi. OIL združuje modelne prvine objektno usmerjenih programskih jezikov z orodji za semantično sklepanje (opisno logiko) ter standardnimi spletnimi formati in shemami (XML, RDF). Zapis ontologije prostora geografskih razsežnosti v standardiziranem jeziku OIL smo izvedli za najvišje ravni ontologije.

```
begin-ontology
ontology-container
title Prostor Geografskih razsežnosti
creator Marjan Ceh
description SDF
description.release "1.0"
type ontology
identifier "id"
language "OIL"
```

```
ontology-definitions
class-def defined PSNN pojmovni napredni
subclass-of
  geografski prostor
  slot-constraint razseznost has-value geografski prostor
  slot-constraint pojavna oblika has-value pojmovna shema
  slot-constraint namen aktivnosti has-value napredni
// geografski prostor
```

```
class-def geografski prostor
// SUB: PSPN pojmovni prvobitni
// SUB: PSNN pojmovni napredni
// SUB: POPN predmetni prvobitni
// SUB: PONN predmetni napredni
```

```
class-def defined PSPN pojmovni prvobitni
subclass-of
  geografski prostor
```

```

slot-constraint pojavna oblika has-value pojmovna shema
slot-constraint namen aktivnosti has-value prvobitni
// geografski prostor

```

```

class-def defined PONN predmetni napredni
subclass-of
  geografski prostor
slot-constraint pojavna oblika has-value predmetni objekt
slot-constraint namen aktivnosti has-value napredni
// geografski prostor

```

```

class-def pojmovna shema
class-def prvobitni
class-def defined POPN predmetni prvobitni
subclass-of
  geografski prostor
slot-constraint razseznost has-value geografski prostor
slot-constraint pojavna oblika has-value predmetni objekt
slot-constraint namen aktivnosti has-value prvobitni
// geografski prostor
class-def predmetni objekt
class-def napredni
slot-def pojavna oblika
slot-def razseznost
slot-def namen aktivnosti
slot-def aktivnost
end-ontology

```

Določenih konceptov posameznih katalogov zbirk prostorskih podatkov ni mogoče izraziti s koncepti OPGR. Zato je bila izvedena analiza izrazov v nezdružljivih konceptih obravnavanih zbirk podatkov.

## PRILOGE

Šifra	Koncept	Izraz					
		stavba	zgradba	objekt	zemljišče	površina	ostalo/druga
<b>Katalog rabe zemljišč zemljiškega katastra</b>							
207	porušeni objekt			1			
217	stavbišče	1					
218	stavba ali zgradba	1	1				
299	nerazčiščeno grad. zemljišče				1		
399	nerazčiščene zelene površine					1	
499	nerazčiščeno nerodovitno zemljišče				1		
	skupaj	<b>2</b>	<b>1</b>	<b>1</b>	<b>2</b>	<b>1</b>	
<b>Katalog kmetijske rabe zemljišč (zajem)</b>							
14	vsa ostala kmetijska zemljišča				1		1
141	zemljišča v zaraščanju				1		
33	poslovna zemljišča				1		
35	zemljišča z mešano rabo				1		
366	druga zemljišča pod transportnimi in komunikacijskimi napravami				1		
38	sorodna odprta zemljišča				1		
422	ostala zamočvirjena zemljišča				1		1
63	ostalo (nikjer drugje op.)						1
714	ostale celinske vode						1
73	ostale vode (nikjer drugje op.)						1
	skupaj				<b>7</b>		<b>5</b>
<b>Katalog konceptov temeljnih topografskih načrtov velikih meril TNN VM 500</b>							
332300	drugo značilno znamenje						1
334400	manjša zgradba nedef. Namena		1				
334400	stavba nedefiniranega pomena	1					
	skupaj	<b>1</b>	<b>1</b>				<b>1</b>
<b>Katalog konceptov topografske baze večje natančnosti</b>							
102	opis zgradbe		1				
105	visoki objekt			1			
204	druga prometna površina						1
301	vegetacija						
302	zemljišče v posebni rabi RTP				1		
	skupaj		<b>1</b>	<b>1</b>	<b>1</b>		<b>1</b>
<b>Katalog konceptov CORINE Land Cover</b>							
11	urbane površine					1	
111	sklenjene urbane površine					1	
112	nesklenjene urbane površine					1	
133	gradbišča						
51	celinske vode						
	skupaj					<b>3</b>	
<b>Katalog konceptov prostorskih oblik (GURS)</b>							
3060002	komunalni objekt			1			
7030002	vrsta rabe zemljišča				1		
7040001	degradirano zemljišče				1		
8020006	točkovni elementi stavb	1					
8060001	manjša stavba	1					
8060002	stavba	1					
8060003	večja stavba	1					
	skupaj	<b>4</b>		<b>1</b>	<b>2</b>		
	skupaj	<b>7</b>	<b>3</b>	<b>3</b>	<b>12</b>	<b>4</b>	<b>7</b>

Preglednica 19: Analiza izrazov semantično nezdružljivih konceptov.





## TERMINOLOGIJA

Za pregled uporabljenih pomembnejših pojmov (in njihovih opredelitev) in kratic so izdelane preglednice, opisane v naslednjih dveh razdelkih. Opredelitve, označene z \*, so utemeljene v okviru monografije.

### BESEDNJAK

Pojem	Angleški izraz	Opis pomena
abstraktnost	abstract	Pojmovna, miselna predstava o nečem. Abstrakcija je načelo in proces izbire pomembnih lastnosti pojavov oziroma izpuščanje nepomembnih glede na trenutni namen.
agent	agent	Delujoči posredovalec (človek, žival, računalnik): človeški agent; programski agent.
aplikacija	aplication	Uporaba, uveljavitev: način, kako je določena tehnologija (na primer tehnologija računalniških zbirk podatkov) uporabljena za reševanje splošnih in posebnih uporabniških problemov.
celovitost predstavitev znanja	knowledge presentation integrity	Logični koncept, ki pomeni, da je za vsako dejstvo, ki je pravilno v nominalni osnovi (obravnavanem delu stvarnosti), mogoče potrditi njegovo pravilnost tudi v konceptualnem modelu.
človeško dojemanje stvarnosti	human comprehension of reality	Klasifikacijska teorija o človeškem dojemanju stvarnosti temelji na hipotezi, da pri razumevanju stvarnega sveta ljudje uporabljajo tri osnovne organizacijske metode, ki prevladujejo v njihovem mišljenju: razlikovanje konkretnih objektov in njihovih opisnih lastnosti, razlikovanje med celotnimi objekti in njihovimi sestavnimi deli, formiranje in razlikovanje med različnimi tipi objektov po njihovih lastnostih in delovanju.
dedukcija	deduction	Sklepanje iz splošnega na posamezno. Deduktivne sposobnosti sistema za upravljanje zbirk; mehanizem

TERMINOLOGIJA

		povpraševanj, ki se izražajo z uporabo logike prvega reda.
določanje položaja v prostoru	locate	Določiti lego, umestiti, namestiti, določiti meje zemljišča; umeščanje v prostoru, georeferenciranje. Glej geodezija, geometrija.
domena problema	problem domain	Stvarno, abstraktno ali hipotetično področje, ki lahko vključuje različne skupine objektov, ki se vedejo v skladu s pravili in značilnostmi obravnavanega področja. Glej področje obravave.
enakozvočnica	homonym	Dva pomensko različna koncepta sta predstavljena z isto besedo (simbolom).
enotni modelirni jezik	unified modelling language	UML je objektno usmerjen, umetni, formalni jezik, ki je neodvisen od področja in okolja uporabe. UML je kot industrijski standard razvil konzorcij OMG (Object Management Group).
entiteta	entity	Pojem za označevanje posameznih obstoječih pojavov v stvarnem okolju (oseba, živo bitje, stvar, fizični objekt, abstraktni pojem).
epistemologija	epistemology	Filozofska disciplina, ki obravnava izvor, strukturo, metodo spoznavanja in veljavnost človeškega znanja. Znanost, teorija o znanju in vedenju. Beseda epistemologija je sestavljena iz grških besed episterne (znanje) in logos (vzrok). Epistemolog je filozof, ki proučuje veljavnost človeškega znanja. Išče odgovor na vprašanje, ali stvarni svet obstaja neodvisno od človeških izkušenj in kako se človek lahko nauči značilnosti stvarnega sveta ter kako je mogoče določiti veljavnost oziroma natančnost človeških izkušenj. Sprašuje se tudi ali vse izkušnje izhajajo iz stika človeka s predmetnim (fizičnim) svetom preko zaznavnih organov. Glej kognitivno kartiranje.
etimologija	etymology	Besedotvornost, veda o zgodovini besed, posebno glede njihovega izvora in pomena.
geo-	geo-	Predpona; nanašajoč se na Zemljo.
geodetske zbirke podatkov	geodetic database	Prostorsko umeščene (geoorientirane) zbirke prostorskih podatkov, izvedene glede na geodetsko osnovo, z geodetskimi opazovanji/izmerami v prostoru.
geodezija	geodesy	Veda, ki ugotavlja natančno obliko in velikost zemeljskega površja ter zemeljska težnostna polja. Do izuma in uporabe umetnih satelitov, so geodetske meritve temeljile na zemljemerstvu s triangulacijskimi metodami v okviru geodetskega koordinatnega sistema.
geografski informacijski sistem	geographic information system	Sistem za zajemanje, shranjevanje, vzdrževanje, obdelavo, povezovanje, analizo in predstavitev prostorskih podatkov. GIS je sistem za podporo odločanju, ki vključuje integracijo zbirk

TERMINOLOGIJA

		prostorskih podatkov. Poudarek je na različnih podatkovnih analizah. GIS sestavljajo strojna oprema, sistemska in posebna programska oprema, uporabniške aplikacije, integrirana baza prostorskih podatkov, vzdrževalci in uporabniki informacijskega sistema.
geografski prostor	geographic space	Geografski prostor se v geografiji lahko obravnava kot geosfero, kot okolje ali kot pokrajino. V geosferi se prepletajo: zemeljska skorja, vodovja, ozračje in biosfera. Geografsko okolje oziroma njegovi sestavni deli so enotni, čeprav so rezultat delovanja bistveno različnih prirodnih in družbenih zakonov. Pokrajina je način dojemanja zemeljske površinske sfere kot stvarne in sestavljene prostorske celote. Geografija skuša celovito obravnavati celotno zemeljsko površje kot tudi posamezne dele, teritorialno ločene in z več vidikov samosvoje celote – pokrajine. Glej prostor geografskih razsežnosti.
geomatika	geomatics	(1)* Geomatika je sklop tehnologij, metod in organizacij za zajem, upravljanje in analize prostorskih podatkov. Tehnološki del geomatike je v grobem mogoče razdeliti na geometersko tehnologijo, računalniško zasnovan informacijski sistem opremljen s primernimi programskimi paketi, žični/brezžični komunikacijski sistem ter naprave za digitalizacijo in tiskanje kart. (2) Področje dejavnosti s sistematičnim pristopom za integracijo vseh sredstev, ki so potrebna za zajemanje in upravljanje s prostorskimi podatki. Ti so potrebni kot del znanstvenih, administrativnih, pravnih ter tehničnih operacij, vključenih v proces proizvodnje, uporabe in upravljanja prostorskih informacij.
georeferenčne meritve	georeferencing	Meritve za določanje položaja stvarnega objekta v modelu prostora, glede na izbrano metodo (geodetske ali negeodetske) in njeno geometrično osnovo (model); umeščanje v modelni prostor. Geodetskim metodam za geometrično osnovo služijo geodetski, astronomski, gravimetrični, satelitski in inercialni referenčni sistemi.
informacija	information	Miselni pomen, ki ga ljudje pripisujejo podatkom z upoštevanjem uveljavljenih načel interpretacije.
integracija podatkov	data integration	Horizontalna podatkovna integracija je združevanje podatkov istega tipa za različna področja. Vertikalna podatkovna integracija je združevanje različnih tipov podatkov za isto področje.
kakovost integracije shem	schema integration quality	Kakovost integracije shem izpostavlja vprašanje kriterijev za testiranja integracije konceptualnih shem. Za ocenjevanje kakovosti integracije shem se predpostavlja kriterije popolnost, minimalnost in razumljivost.
kartiranje spajanja shem	mapping of schema merging	Pri modeliranju shem zbirk podatkov za kartiranje spajanja shem gre za modeliranje razmerij višjega reda, to je razmerij med razmerji shem. Pri tovrstnem modeliranju se skuša zadržati število

TERMINOLOGIJA

		različnih konceptov čim manjše, da bi bil model lahko čimboli enostaven in enoten. Za grafične predstavitve shem zbirk podatkov se uporablja usmerjene grafe.
klasifikacija semantičnih razmerij	semantic relations classification	Semantična sorodnost je pri integraciji shem zbirk podatkov opredeljena s klasifikacijo semantičnih razmerij. Obstaja več tipov semantičnih razmerij med koncepti: istovetno (natančno enak), enakovredno (uporaba neenakih a enakovrednih sestavin ter uporaba istih spoznanj), skladno (uporaba neenakovrednih a nenasprotujočih sestavin), neskladno (vsi elementi nasprotujoči).
kognitivna teorija	cognitive theory	Teorija o miselnih procesih spoznavanja, vedenja, ki vključuje človeško zaznavanje in presojo. V procesu presoje je obravnavani objekt razlikovan od drugih in v umu označen z enim ali več pojmi (koncepti). V sodobni kognitivni teoriji obstajata dva pristopa, in sicer pristop obdelave informacij in pristop miselne prilagoditve. V pristopu obdelave informacij se poizkuša spoznati človeške miselne postopke in postopke sklepanja. Navedene postopke se primerja z obelavo podatkov v računalniško zasnovanih sistemih. Pristop miselne prilagoditve obsega dva postopka: prisvojitve (asimilacija) in prilagoditve (akomodacija). V postopku kognitivnega prisvajanja posameznik interpretira stvarnost s pojmi lastnega notranjega modela sveta, ki temelji na predhodnih izkušnjah. V postopku kognitivne prilagoditve se notranji model sveta spreminja tako, da se prilagaja stvarnosti.
kognitivno kartiranje	cognitive mapping	Spoznavno, miselno kartiranje. Kognitivno kartiranje je miselna obdelava prostorskih informacij. Kognitivno kartiranje je miselna sposobnost pojmovnega (abstraktnega) zbiranja, organiziranja, poizvedovanja in upravljanja informacij o prostoru. Kognitivna karta je izdelek (shema) v človekovem umu, in sicer miselna predstavitev nekaterih delov prostora v določenem časovnem preseku. Za reševanje prostorskih problemov s pomočjo miselnih predstav sveta je značilna neprestana interakcija z okoljem, ki se ga kognitivno kartira.
koncept	concept	Miselna zasnova, pojem, (za)misel, predstava, zasnutek. Abstraktna miselna enota, ki se oblikuje skozi posploševanje skupnih lastnosti materialnih in nematerialnih predmetov. Koncept ni jezikovna ampak miselna enota, ki šele s poimenovanjem preide na jezikovno raven. Koncepti so neodvisni od posameznega jezika, lahko pa na oblikovanje konceptov vpliva družbeno in kulturno okolje. Razmerja med koncepti – pojmi niso razmerja med termini. Ker so pojmi abstraktno miselne enote, na ravni pojmovnega sistema sopomenskost (sinonimija) ne obstaja; to je razmerje, ki se lahko pojavi le na ravni poimenovanj oz. terminov. Glej predstavitev pojma, pojmovni sistem.
konceptualizem	conceptualism	Srednjeveška filozofska smer, ki trdi, da splošni pojmi niso samo

TERMINOLOGIJA

		besede, temveč obstajajo tudi v razumu.
konceptualna shema	conceptual schema	Formalen opis in opredelitev konceptualnega modela v jeziku za sestavljanje konceptualne sheme (jeziku konceptualne sheme). Konceptualna shema je strogi opis pojmovnega modela z formalno opisno tehniko, ki temelji na konceptualnem formalizmu. Konceptualna shema podaja vsebino, pojmovno sestavo in pravila, ki jih uporabljamo za podatke o objektih v razvitem konceptualnem modelu.
konceptualni formalizem	conceptual formalism	Niz modelnih zamisli, ki se jih uporablja za formalno opredelitev miselnega modela. Formalizem zagotavlja pravila, omejitve, pravila dedovanja, dogodke, procese in druge elemente, ki tvorijo jezik za izdelavo konceptualne sheme. Konceptualno formalizem lahko izrazimo v več formalnih opisnih tehnikah (UML, Express itd.).
konceptualni graf	conceptual graph	Konceptualni graf je struktura pojmov in razmerij med pojmi, kjer vsaka povezava povezuje konceptualno vozlišče in vozlišče konceptualnega razmerja.
konceptualni model	conceptual model	Opreljuje pojme področja obravnave ter podaja abstrakten opis izbranih stvarnih pojavov. Konceptualni model podaja ustrezno podatkovno in funkcionalno abstrakcijo izbora stvarnega sveta za določeno uporabo. Glej pojmovni sistem.
konceptualno	conceptual	(1) Pristop konceptualnega modeliranja je opis področja obravnavanja modeliranja nave s konceptualnim modelom. (2) Postopek abstrakcije in uvrščanja objektov izbranega dela stvarnega sveta v sestave, katere se lahko predstavlja kot razrede v računalniško zasnovani zbirki podatkov oziroma informacijskem sistemu.
konjunkcija	conjunction	Sestavljena izjava, katere deli se med seboj povezujejo.
kontekst	context	Vidik, miselna povezava, sobesedilo – besedilo, v katero spada obravnavani del besedila.
ločljivost	separability	Mera za podrobnost pri opazovanju skozi optično pripravo; resolucija.
logika prvega reda	first order logic	Vzdevčna (predikatna) logika imenovana tudi logika prvega reda (First Order Logic – FOL). Predikatni račun podaja jezikovne sestave za označevanje posameznih objektov, lastnosti objektov in razmerij med objekti.
medopravilni GIS	interoperable GIS	Interoperabilni GIS naj bi bil sistem neodvisen od razlik v (geo) informacijski tehnologij poleg tega pa bi moral tudi nazorno komunicirati navzven (splet), in sicer na visoki semantični ravni. Zato je treba izdelati ontologijo geoinformacijske infrastrukture. Geoinformacijska medopravilnost; skladno z mednarodnim standardom ISO 19119 – Geografski servisi, ki ga je med drugimi razvil tehnični odbor ISO TC/211 – Geografske

TERMINOLOGIJA

		informacije // Geomatika, je medopravilnost sposobnost in informacijskega sistema (GIS), da: prosto izmenjuje vse vrste in oblike prostorskih (geografskih) podatkov o objektih oziroma pojavih na, nad ali pod zemeljskim površjem; da povezano preko omrežij uporablja in poganja programsko opremo, ki je sposobna ravnati in upravljati s porazdeljenimi prostorskimi podatki in informacijami.
medopravilnost	interoperability	(1) Medizvedljivost, interoperabilnost je pogoj, ki se izpolnjuje med sistemi za elektronsko komuniciranje ali napravami za elektronsko sporazumevanje takrat, ko se podatki ali storitve med elektronskimi sistemi in/ali njihovimi uporabniki lahko izmenjujejo neposredno in zanesljivo. (2) Medopravilnost je zmožnost komuniciranja, izvajanja programov in prenosa podatkov med različnimi funkcionalnimi enotami na način, ki od uporabnika ne zahteva posebnega poznavanja tehničnih značilnosti takšnih naprav. Medopravilnost pospešuje, omogoča in podpira delitev obdelav ter porazdeljevanje podatkov med različnimi sistemi.
mereologija	mereology	Teorija o celoti in njenih delih. Mereologija za osnovo privzema prostorsko nepretrganost (kontinuum). Skrajna različica mereološkega pristopa je aristotelski mereološki potencializem, po katerem "del celote nikoli ne more biti stvaren, če je celota ni stvarna".
meta podatek	metadata	Podatek, ki vsebujejo informacije o drugih podatkih. Meta podatki ali podatki o podatkih so izpeljane informacije o zgradbi, vsebini, kakovosti, zgodovini, organizaciji, dostopnosti, vrednosti in uporabi shranjenih podatkov.
miselna karta	mind map	Ena najenostavnejših in najmanj formalnih metod predstavljanja znanja je kartiranje mišljenja (misli, spomina). Miselne karte (miselni vzorci) so popolnoma neformalni grafi, ki se jih uporablja v začetnih korakih modeliranja. Miselne karte je težko primerjati med seboj, nenazadnje ker nimajo enotnega sistema simbolov.
model	model	Namen modela je zagotovitev okolja, v katerem se je mogoče učiti dejstev o stvarnosti z analiziranjem modela namesto z neposrednim opazovanjem stvarnosti.
objekt	object	Besedo objekt se uporablja v stvarnosti in v modelih. Objekt je stvar ali pojav v stvarnem svetu. Objekti v stvarnem svetu so lahko predmetni ali pojmovni (abstraktni). Včasih se kot sopomenka (sinonim) za stvarni objekt uporablja beseda entiteta. Objekt je tudi osnovni modelni gradnik, okrog katerega so opredeljeni ostali modelni gradniki.
objektni identifikator	object identity	Za utemeljevanje objektnega sistema so potrebni jezikovni simboli, ki se nanašajo na določene stvarne objekte. Takšni simboli se imenujejo objektni identifikatorji. Objektni

TERMINOLOGIJA

		identifikatorji zagotavljajo prijemališča objektov v objektivnem sistemu. Razliko med stvarnostjo in jezikom je mogoče predstaviti s sledečo trditvijo: objekti pripadajo svetu obravnave (del stvarnosti), objektni identifikatorji pa pripadajo jeziku, ki je uporabljen za utemeljevanje objektivnega sistema.
objektni katalog	feature catalogue	(1) Seznam, ki vsebuje opredelitve objektivnih tipov (razredov), njihovih atributov in relacij med objekti (pojavi), kateri se pojavljajo v enem ali več podatkovnih nizih, vključujoč tudi vse opredeljene operacije objektov. (2) Objektivni katalog predstavlja abstrakcijo stvarnosti, ki je predstavljena v enem ali več nizih geografskih podatkov, kot opredeljeno klasifikacijo za vsebovane fenomene. Osnovni nivo klasifikacije v objektivnem katalogu je objektivni tip. Objektivni katalog mora biti dostopen v elektronski obliki za katerikoli niz geografskih podatkov, ki vsebuje objektivne tipe. Objektivni katalog se lahko prav tako podredi specifikacijam v ustreznem mednarodnem ISO standardu ne glede na obstoječe podatkovne nize.
obnavljanje podatkov	update	Osvežitev, posodobitev; ažuriranje podatkov.
obnovljenost podatkov	currency	Ažurnost, aktualnost podatkov, oziroma informacij. Časovna natančnost je pomembna pri dinamični obravnavi prostorskih podatkov. Glej obnavljanje.
ohlapna integralna shema	loose integral schema	Če se želi z eno shemo predstaviti vse informacije druge sheme in še dodatne informacije, je potrebno, da prva shema vsebuje najmanj vse razrede druge sheme. Tako opredeljene podatkovne sheme se imenujejo ohlapne sheme.
omogočanje	affordance	Ljudje razlikujejo med objekti v svetu, ki omogočajo aktivnosti. Obravnava aktivnosti je ključni element modeliranja z vidika prostora.
ontologija	ontology	(1) Filozofska disciplina, ki obravnava osnovo, vzroke in najsplošnejše lastnosti stvarnosti (nauk o bitju). Ontologija je tesno povezana s teorijo logike. Glej semantično omrežje. (2) Ontologija v konceptualnem modeliranju je eksplicitna specifikacija konceptualizacije (uporabljene terminologije v določeni problemski domeni). Ontologija je opis konceptov in razmerij med koncepti, ki lahko obstajajo za agente oziroma skupnost agentov. Ontologija je podana kot niz formalnih opredelitev besednjaka.
ontologija aplikacije	application domain	Za vsak podatkovni niz ali posamezno aplikacijo posebna ontologija.
ontologija domene	domain ontology	Ontologija domene je rezultat soglasja o skupnem razumevanju domene in jo je zato mogoče uporabiti za različne namene, vključno za uresničevanje medopravnosti informacijskih sistemov. Pri tem se ontologija domene uporabi kot sredstvo izmenjave pomenov.

TERMINOLOGIJA

ontologija prostora geografskih razsežnosti	ontology of geographic space	Ontologija, na osnovi katere je mogoče izvajati pomensko integracijo podatkov zbirk prostorskih podatkov. OPGR vsebuje semantični slovar prostorskih izrazov in hierarhično semantično pmrežje. Za model je značilna neodvisnost od podatkovnih objektov kakor tudi od stvarnih prostorskih objektov, ker je opredeljena z nameni (prvobitni, napredni) oziroma aktivnostmi v prostoru ter značilnostmi njihovih pojavnih oblik, ki so lahko predmetne ali pojmovne; univerzalni obseg modela, ki obravnava celotno domeno prostora geografskih razsežnosti; dinamično razširjanje modela; izražanje kompleksnih konceptov s kombiniranjem osnovnih konceptov.
ontologija v inženirstvu znanja	ontology in knowledge engineering	V inženirstvu znanja se z ontologijo obvladuje jezikovno izraženo znanje. Ontološko raziskovanje presega konceptualno modeliranje. Do nedavnega ni bilo povezave med raziskavami semantike prostorskih objektov geografske razsežnosti ter med raziskavami semantike predstavitev prostorskih objektov.
ontologije geoinformacijske infrastrukture	geoinformation infrastructure ontology	Medopravilni GIS bi moral biti sistem neodvisen od razlik v (geo) informacijski tehnologiji, poleg tega pa bi moral tudi nazorno komunicirati navzven (splet), in sicer na visoki semantični ravni. Zato je potrebno izdelati ontologije geoinformacijske infrastrukture oziroma ontologije prostorskih informacijskih sistemov.
ontološka obveznost	ontologic obligation	Ontološka obveznost je dogovor/pogodba z agenti o uporabi slovarja (z ustvarjanjem trditve in povpraševanjem) na način, ki je skladen s teorijo, ki je opisana z ontologijo. Predanost agentov ontologiji, glede na povpraševanja in trditve izvedene z uporabo slovarja opredeljenega v ontologiji, je porok skladnosti (ne pa popolnosti).
opredelitev simbola	symbol definition	Opredelitev je postopek (govorjeni ali kako drugače izvedeni), v katerem se posamezniki med seboj obvestijo o pomenu obravnavanega simbola. Opredelitev navaja, na kateri koncept se nanaša simbol. Opredelitev vedno opredeljuje simbol. Opredelitev nikoli ne opredeljuje koncepta. Opredelitev se glede na namen delijo na dve vrsti: opisne in dogovorne. Za opisno opredeljevanje simbolov se uporablja opisna logika oziroma terminološka logika (Description Logic - DL).
ovojnik	wrapper	Programski paketi za prevajanje informacijskih povpraševanj iz oblike, ki je razumljiva posredovalni shemi v obliko, ki jo razumejo posamezni lokalni podatkovni viri.
platonizem	platonism	Filozofska smer, ki temelji na Platonovi filozofiji o idejah kot samostojnih bitnostih.
podatek	data	Dejstvo, ki se nanaša na stvar. Določeno dejstvo, ki omogoča določeno stvar spoznati ali o njej sklepati. Sporočilo v taki obliki, da se lahko obravnava v računalništvu. Znak, ki predstavlja surovino za ustvarjanje informacij.



TERMINOLOGIJA

podatkovni model	data model	(1) Podatkovni model se v splošnem razumeva kot opis strukture podatkov na računalniški način. Uporaba formalnih metod pri podatkovnem modeliranju omogoča uporabo konceptov, katerih pomen je formalno opredeljen. Matematična formulacija omogoča, da so vsebine predmeta obravnave mehanizirane v strojih, računalniških sistemih. (2) Abstraktni zapis in miselna interpretacija dejstev ali znanja za del zapletenega stvarnega sveta glede na določen namen.
podatkovni sloj	layer, coverage	Združuje objekte iste tematike. Digitalni sloj vsebuje le en tip geometričnih podatkov (rastrski ali vektorski).
podrobnost	detail	Ena od štirih temeljnih razsežnosti geografskih podatkov (tema, čas, podrobnost, prostor). Detajl, ki ga lahko razločimo v prostoru, času ali tematiki. Zajema, upošteva vse podrobnosti, ki z največjo mogočo popolnostjo podaja resnično stanje.
področje obravnave	universe of discourse	Torišče, domena. Celota pojavov, stvari, na katero se nanaša in je usmerjeno človekovo delovanje, ustvarjalnost, raziskovanje. Del stvarnosti, ki je predstavljen/obravnavan z modelom. (2) Pogled na stvaren ali hipotetičen svet, ki vključuje vse, kar je zanimivo za določen namen. Sestavljeno je iz niza objektnih tipov in razmerij med njimi. Področje obravnave se formalno zapiše v konceptualnih shemah.
pojmovni sistem	conceptual system	Shematska ureditev pojmov na izbranem področju; ne le terminološki pomen: pojmovni sistem nam pomaga urediti znanje in razumevanje področja v celoti ter njegovo povezovanje s sorodnimi področji; dejavniki: namen, stroka, kriteriji, razmerja med pojmi; način prikaza: grafični, tabelarni, slikovni; drevesni diagram, tabelarni diagram, krožni diagram, tabela lastnosti, puščični diagram, klasifikacija/taksonomija. Glej konceptualni model.
pomenskost	meaning ability	Sposobnost česa, da izraža pomen; simbolna pomenskost, pomenskost besede.
porazdeljena zbirka podatkov	distributed database	Porazdeljena zbirka podatkov je zbirka več logično povezanih zbirk podatkov porazdeljenih preko računalniškega omrežja.
porazdeljene geografske informacije	distributed geographic information	Porazdeljevanje prostorskih podatkov v porazdeljenih podatkovnih sistemih (predvsem internet) omogoča številne načine, na katere je mogoče porazdeljevati geografske informacije v okviru sistemske arhitekture odjemalec/strežnik.
porazdeljeni GIS	distributed GIS	Programski agent za geografske informacije v porazdeljenem informacijskem sistemu. Lahko je mobilni programski agent.
porazdeljeni podatkovni sistem	distributed database system	Porazdeljeni podatkovni sistem (PPS) = porazdeljena zbirka podatkov (PPZ) + porazdeljeni sistem za upravljanje zbirk podatkov.

TERMINOLOGIJA

porazdeljeni sistem za upravljanje zbirk podatkov	distributed database management system	Računalniški program, ki upravlja porazdeljeno zbirko podatkov (PPZ) in zagotavlja dostopni mehanizem, ki za uporabnike prikrije porazdeljenost.
pragmatika	pragmatics	Preučevanje povezav med znaki in agenti, ki jih uporabljajo za sklicevanje na stvari v stvarnosti in za sporazumevanje z drugimi agenti, ki imajo podobne ali različne namene glede istih ali različnih stvari. Pragmatika je govorništvo (retorika).
predstavitev pojma	concept representation	Koncept, miselno zasnovano, pojem lahko predstavimo na več načinov; poimenovanje: termin (eno- ali večbesedni, okrajšava), alfanumerični simbol, grafični simbol; opis: opredelitev, umestitev v razmerje del-celota; razlaga; "opisna" nejezikovna reprezentacija (npr. kompleksna formula ali graf). Glej pojmovni sistem.
primerjanje shem	schema comparison	Modele ustvarijo različni ljudje z različnimi nameni. Primerjanje shem zadeva tolmačenje (interpretacijo) modelov objektnega sistema. Vsem metodam primerjanja shem je skupno, da informacije, katere so vsebovane v lokalnih shemah niso zadovoljive za izdelavo zanesljivega niza povezovalnih trditev. Zato se v vse novejšie tehnike primerjanja shem uvaja metode semantičnega bogatenja lokalnih shem. Glej semantična bogatitev shem.
programski agent	program agent	Inteligentni posredniki med uporabnikom in programi. Stacionarni programski agenti so pasivni, čakajo na zunanjo iniciativo in imajo vlogo orodja. Nasprotno so mobilni programskimi agenti proaktivni.
prostor geografskih razsežnosti	space of geographic dimensions	Prostor geografskih razsežnosti predstavljajo prostorski objekti geografske razsežnosti; objekti, ki so večji od človeškega telesa in jih ni mogoče zaznati v celoti v okviru enega zaznavnega dejanja. Glej geografski prostor.
prostorska integracija podatkov	spatial data integration	Prostorska integracija podatkov temelji na zanesljivem geometričnem referenčnem sistemu in na razumni ravni zajemanja podrobnosti in ločljivosti različnih podatkovnih nizov. To pomeni, da mora biti za vse različne zbirke podatkov pri njihovi izdelavi uporabljen isti koordinatni sistem. Vse integrirane zbirke podatkov morajo imeti primerljivo raven ujemanja podrobnosti.
prostorska semantična hierarhija	spatial semantic hierarchy	Teorija kognitivnega kartiranja temelji na hierarhiji predstavitev prostorskega znanja, ki se imenuje prostorska semantična hierarhija. Vsaka raven hierarhije ima svojo lastno ontologijo (niz objektov in razmerij za opisovanje sveta) ter svoj niz metod reševanja problemov in sklepanj.
prostorski objekt	geographic spatial object	V splošnem se ločuje med predmetnimi (fizičnimi) objekti in nepredmetnimi objekti geografskega obsega. Geografski prostor je sestavljen iz enot, ki niso zgolj fizični objekti ampak

TERMINOLOGIJA

		v pomembnem obsegu ustvarjeni v človeškem umu. Naravni (bona-fide) prostorski objekti (planeti, luna, otoki, jezera, reke), ki obstajajo na osnovi notranjih fizičnih prekinitev v materialni sestavi Zemlje; umetni (fiat) prostorski objekti (države, regije, lastniške zemljiške parcele); ta vrsta objektov nima notranjih fizičnih prekinitev, ampak so meje utemeljene na osnovi dogovora ali so na drugačen način izdelki človeškega geografskega delovanja.
prostorski podatek	spatial data	Podatek, ki opisuje pojave v prostoru in vsebuje tudi navedbo prostorskega položaja (prostorsko referenco) v prostorskem referenčnem sistemu. Prostorski podatki so tudi podatki o razsežnostih (dimenzijah) prostorskih objektov (dolžina daljic, površina ploskev, prostornina teles) ter podatki o prostorskih razmerjih med prostorskimi objekti (preseki, prekrivanje, vsebovanost, skupna meja).
redundanca	redundacy	Presežek, prebitek, preobilica, pretek, bogatost, bogastvo. Redundanca podatkov o pomenu istega koncepta povečuje verjetnost pravilnega razumevanja informacije. Besedna redundanca se pojavi, kadar različni besedi oziroma simbola predstavljata isti objekt ali idejo. Strukturna redundanca nastopa, kadar je lahko ista ideja izražena z več različnimi stavki, celo če stavki vsebujejo iste besede.
rudarjenje prostorskih podatkov	spatial data mining	Odkrivanje uporabnih vzorcev in razmerij med podatki v zbirkah prostorskih podatkov.
semantična bogatitev shem	semantic schema enrichment	Bogatejši kot so opisi področja obravnave v določenem časovnem trenutku, bolj natančna znanja o svetu obravnave je mogoče pridobiti s tolmačenjem opisov v določenem poznejšem časovnem trenutku. Semantično obogatena shema izraža več semantike objektnega sistema kot izvirna shema. To se doseže z uvajanjem dodatnih informacij o obravnavanem področju. Pri semantični bogatitvi shem se jasno poudari različne vrste skritih informacij.
semantična dvoumnost	semantic ambiguity	Semantična dvoumnost izvira iz dovoljenih večpomenskih (polisemičnih) izrazov v zbirki znanja, kjer je en izraz povezan z več pomeni. Večpomenski izrazi se pojavijo celo, če se klasifikacija uporablja dosledno. Reševanje dvopomenskosti izrazov je mogoče razrešiti z doslednim preverjanjem konteksta sheme (ontologije).
semantična faktorizacija	semantic factoring	Postopek analiziranja – razstavljanja kategorij izvirmih ontologij v nize osnovnih kategorij, ki se imenujejo semantični faktorji. V primeru prekrivanja kategorij se izvede njihovo razdružitev v razčlenjene razrede. Skupni, prekrivajoči del tvori nov razred. S kombiniranjem semantičnih faktorjev in njihovih lastnosti se ustvari omrežje konceptov.
semantična globina prostorskega	semantic deepness of spatial	Seemantična razdalja med izhodiščno ravno ontologije domene prostora ter obravnavanim prostorskim konceptom.

TERMINOLOGIJA

koncepta *	concept	Večja globina predstavlja večjo semantično podrobnost prostorskega koncepta.
semantična globina zbirke prostorskih podatkov*	semantic deepness of spatial database	Semantična globina zbirke prostorskih podatkov glede na izhodiščno raven ontologije prostora geografskih razsežnosti. Utemeljena je na osnovi meritev semantičnih razdalj med koncepti zbirke podatkov in izhodiščne ravni ontologije domene prostora.
semantična heterogenost	semantic heterogeneity	Pomenska raznovrstnost podatkov se pojavlja, kadar obstaja neskladje med osnovno razlago (interpretacijo) podatkov in med nameravano uporabo podatkov, lahko pa tudi, kadar je isti pojav modeliran v dveh shemah na različen način, četudi sta shemi oblikovani na osnovi istega podatkovnega modela.
semantična integracija zbirk podatkov	database integration	Poenotenje podatkov v spoznavnem (kognitivnem) smislu. Kognitivno poenotenje ima za integracijo informacij temeljni pomen, kajti izvedba analiz je zaradi integriranih informacij mogoča na masovni (molarni) ravni brez potrebe po takojšnji presoji (molekularnih) podrobnosti. Postopek spoznavanja besed je rezultat integracije različnih gesel, miselnih zvez, namigov. Na višji ravni je razumevanje pomenov terminov (besed) tudi rezultat integracije miselnih zvez. Integracija se povezuje z vrednotenjem v verigo integracijskih postopkov. Integracijski postopek na eni ravni zasnuje postopek vrednotenja za drugo raven obravnave. Združitev lokalnih shem več obstoječih zbirk podatkov v globalno shemo.
semantična kakovost razmerja	relation quality	Stopnja razmerja sorodnih konceptov (po pomenu). Štiri stopnje semantičnih razmerij: šibko razmerje, znosno razmerje (kompatibilno), enakovredno razmerje (ekvivalentno) in spojitveno razmerje (mergable).
semantična medopravilnost	semantic interoperability	Semantično medopravilnost na podatkovni ravni pomeni skladno interpretacijo pomena terminov (simbolov) pri povezovanju zbirk podatkov.
semantična natančnost agenta	semantic accuracy of agent	* Semantična natančnost agenta je natančnost, s katero agent uvršča simbol/izraz v primeren koncept ontologije. Meriti jo je mogoče pri ponavljanju uvrščanja istih izrazov. Natančnost klasifikacije.
semantična pristranost koncepta	semantic bias of concept	* Semantična pristranost koncepta je posledica podrejanja meto- dam zajemanja podatkov pri opredeljevanju konceptov prostora. Še posebej je značilna za koncepte v katalogih prostorskih podatkov, ki so zajeti z metodami daljinskega zaznavanja in razpoznavani z metodami fotointerpretacije. V opredelitvah semantično pristranskih konceptov je značilna uporaba izrazov za opisovanje snovi (pesek, kamen, beton, asfalt, vegetacija).

TERMINOLOGIJA

semantična projekcija	semantic projection	Semantična projekcija omogoča kartiranje iz bolj kompleksnih v bolj enostavne podatkovne modele.
semantična razdalja	semantic distance	Semantična ustreznost se izraža s semantično razdaljo med pomeni znakov (simbolov, terminov) o istem stvarnem objektu, pridobljenih iz različnih podatkovnih virov. Semantična razdalja služi ocenjevanju primernosti uporabe določenih podatkov za povezovanja zbirk podatkov (na primer za namen samodejnega obnavljanja podatkov v medopravilnih zbirkah podatkov).
semantična razpršenost zbirke prostorskih podatkov	semantic diffusion of spatial database	*Razprostranjenost konceptov zbirke podatkov v okviru ontologije prostora geografskih razsežnosti. Utemeljena na osnovi meritev semantičnih razdalj med koncepti zbirke podatkov. Podatek o razpršenosti zbirke podatkov je uporaben meta podatek (glej meta podatek), še posebej za integracijo zbirk podatkov v okolju porazdeljenih geografskih informacijskih sistemov.
semantična skladnost	semantic accordance	Pomensko ujemanje (ustreznost, neustreznost) različnih predstavitev (znakov, terminov, simbolov) istih pojavov stvarnega sveta. Pri analizi podatkovnih shem se išče obstoj razmerij med shemami. Ko so ta razmerja odkrita se odkriva stopnjo podobnosti. Pogoj za semantično povezovanje zbirk podatkov.
semantična skladnost zbirk prostorskih podatkov	semantic connectedness of spatial databases	*Stopnja podobnosti zbirk prostorskih podatkov glede na dva kriterija (zapisana kot meta podatka) vsake primerjane zbirke prostorskih podatkov (semantična globina, semantična razpršenost). Semantična skladnost je predpogoj za semantično povezovanje zbirk podatkov.
semantična translacija semantična ustreznost ontologije	semantic translation semantic adequacy of ontology	Semantična translacija je opredeljena kot matematični proces transformacije iz enega v drug semantični referenčni sistem. *Ocena stopnje primernosti oziroma zanesljivosti določene ontologije za izražanje konceptov zbirk podatkov.
semantične prvine	semantic primitives	Semantične prvine so: obstoj, sklicevanje, razmerje, sovpadanje, razveljavitev. Semantični operatorji so: posploševanje, posledičnost in ločevanje. Vsak zapis, s katerim je mogoče izraziti navedene prvine, v vseh mogočih kombinacijah, vsebuje logiko prvega reda kot podniz. Različni jeziki lahko uporabljajo različne zapise za predstavitev navedenih prvin.
semantični datum	semantic datum	Semantični datum opredeljuje pomen osnovnih izrazov izven sistema. Semantični datum je opredeljen kot mehanizem za učvrstitvev / interpretacijo osnovnih izrazov semantičnega referenčnega sistema.
semantični opis	semantic description	Brez poznavanja miselne povezave z ontologijo je težko prepoznati ali dve besedi predstavljata isti koncept ali ne. V pomoč je lahko semantični opis, ki poleg opredelive simbola podaja natančnejši opis koncepta, ki ga simbol predstavlja, predvsem

TERMINOLOGIJA

		dovoljene operacije za koncept. Semantični opis se uporablja pri semantični bogatitvi shem.
semantični slovar	semantic dictionary	Semantični slovar je podpora integracijskemu orodju, za pregledovanje simbolov v lokalnih shemah, ki se jih povezuje. Semantični slovar predstavlja zgodovinski pregled že izvedenih integracij, saj lahko beleži ugotovitve, katere je mogoče uporabiti v prihodnjih izvajanih integracije.
semantični referenčni sistem	semantic reference system	Elementa semantičnega referenčnega sistema sta semantični datum ter semantično referenčno omrežje. Semantični datum opredeljuje pomen osnovnih izrazov izven sistema, semantično referenčno omrežje pa deluje kot koordinatni sistem, ki formalno opredeljuje pomen izrazov. Glej semantično referenčno omrežje.
semantično referenčni podatki	semantic reference framework	Semantično omrežje, ki je sestavljena predstavitev semantičnih razmerij med objekti ali pojavi v stvarnem svetu (objektnem sistemu). Za izdelavo semantičnega slovarja je potrebna osnova, ki se imenuje semantično omrežje. Deluje kot koordinatni sistem, ki formalno opredeljuje pomen izrazov. Glej semantični referenčni sistem.
semantično referenciranje	semantic referencing	Semantično referenciranje: osnovni postopek vzpostavljanja razmerijih med izrazi podatkovnega modela ter sematičnim referenčnim omrežjem.
semantika	semantics	Pomensko, nanašajoč se na pomen, pomenoslovno. Semantika je logična primernost (logic proper), ki je formalna znanost o pogojih pravih predstavitev stvarnosti. Semantika je preučevanje povezav med znaki (termini, simboli) in stvarmi (v stvarnosti).
sematika tarskega	Tarski semantics	Trditvena (propozicijska) logika; atomarne dele jezika (besede, fraze) se predstavi s trditvenimi znaki – črkami, trditvam se pripisuje pomen z določanjem pogoja pravilnosti trditve.
semiotika	semiotics	Semiotika je veda o znakih in je del teorije informacij, ki obravnava komunikacijo med različnimi vrstami agentov preko jezika in drugih simbolnih sistemov. Semiotika je sorodna kognitivni znanosti v tem, da je interdisciplinarna veda, katere področje se nanaša na jezikoslovje, logiko in filozofijo.
simbol	symbol	Znamenje, znak ali več povezanih znakov, ki predstavljajo kak pojem. Jezikovni simbol – beseda, termin; alfanumerični simbol; grafični simbol.
skladnja	syntax	Nauk o stavku. Sintaksa je preučevanje povezav med znaki; s sintakso se obravnava načine sestavljanja stavkov iz manjših sestavnih delov kot so znaki abecede, besede in fraze; odkriva splošna pravila za sestavljanje stavkov; sintaksa je po Peirceju čista slovnica.

TERMINOLOGIJA

skladnostne trditve	integration assertions	Določitev pogojev, izraženih s skladnostnimi (integracijskimi) trditvami, pri katerih sta lahko dve shemi smiselno združeni. To pomeni določitev načina za spajanje dveh shem v integrirano (globalno) shemo, ki bo imela enako informativno sposobnost kot izvirni shemi. Integralna shema mora predstavljati toliko informacij kot izvirni shemi.
sopomenke	synonym	Dva različna termina (simbola) predstavljata isti pojem (koncept).
spajanje shem	schema integration	Metode povezovanja/integracije shem za spajanje lokalnih in porazdeljenih informacij. Spajanje shem je problem povezovanja (kombiniranja) več shem zbirke podatkov v enoten pogled oziroma integralno shemo.
standardna klasifikacija dejavnosti	statistic nomenclature for economic activities	Standardna klasifikacija dejavnosti (SKD) je obvezen nacionalni standard za evidentiranje, zbiranje, obdelovanje, analiziranje, posredovanje in izkazovanje podatkov, pomembnih za spremljanje stanj in gibanj na ekonomskem in socialnem področju ter na področju okolja in naravnih virov. Uporablja se za določanje dejavnosti in za razvrščanje poslovnih subjektov in njihovih delov za potrebe različnih uradnih in drugih administrativnih zbirk podatkov ter za potrebe statistike in analitike v državi in mednarodno. SKD temelji na klasifikaciji dejavnosti NACE Rev.1 (Nomenclature statistique des activités économiques dans la Communauté européenne), ki je obvezen statistični standard Evropske unije in je neposredno povezana z mednarodno klasifikacijo dejavnosti Združenih narodov ISIC Rev. 3 (International Standard Industrial Classification). SKD vsebuje 16 področji, na prvi od šestih hierarhični ravni.
temeljna ontologija	core ontology	Ontologija, ki obsega vse ravni, od posamezne entitete do kategorije najvišje ravni domene. Sistem integracije zbirk prostorskih podatkov vsebuje poleg prostorske ontologije še ontologijo geoinformacijske infrastrukture (tipoloških elementov).
temeljne zbirke podatkov	reference data-base/basic framework data	Ključne (referenčne) zbirke prostorskih podatkov državne podatkovne infrastrukture.
termin	term	Beseda, besedna zveza; jezikovno poimenovanje določene-ga pojma v okviru strokovnega področja. Glej pojem, predstavitev pojma.
topologija	topology	Teorija o stiku, razdruževanju in povezanosti. Topologija podaja povezljivost in opisuje sosedske odnose med objekti v prostoru. V GIS topologija opredeljuje relativne odnose in geometrično povezanost grafičnih gradnikov, ki predstavljajo nek geografski pojav, brez uporabe njihovega absolutnega položaja oziroma korodinat. V vektorski sestavi je topologija

TERMINOLOGIJA

		temelj geometrijske organizacije podatkov.
translacija shem	schema translation	Prevajanje shem; kadar so sheme na različnih vozliščih izražene z različnimi podatkovnimi modeli, je neobhodno zagotoviti prevajanje med modeli.
transparentnost	transparency	V smislu informacijskih sistemov gre za prikrito porazdeljenost. Posredovalni sistemi omogočijo storitve kot so prikrit dostop do oddaljenih podatkov in povezovanje aplikacij v porazdeljene datotečne sisteme ali porazdeljene zbirke podatkov.
večnamenskost	multipurpose	Večnamenskost podatkov v podatkovni zbirki. Uporabnost istih podatkov za različne namene.
vzdrževanje podatkov	data maintenance	Vzdrževanje sistema za upravljanje in shranjevanje podatkov (ohranjanje podatkov). Glej obnavljanje podatkov.
zbirka podatkov	database	(1) Formalno organizirana zbirka podatkov; če je avtomatizirana z računalniškim sistemom, se uporablja ime podatkovna baza. (2) Podatkovna baza je zbirka neodvisnih del, podatkov ali drugega gradiva v kakršnikoli obliki; gradivo, ki je sistematično urejeno in posamično dostopno z elektronskimi ali drugimi sredstvi, pri čemer pridobitev, preveritev ali predstavitev vsebine podatkovne baze zahteva znatno naložbo.
zbirka prostorskih podatkov	spatial database	(1) Formalno organizirana zbirka podatkov, katere del so tudi podatki o položaju v prostoru, umeščeni (georeferencirani) glede na izbrani prostorski referenčni sistem (globalni, lokalni, geografski, geodetski). (2) Organizirana zbirka podatkov o prostoru ne glede na topološko in podatkovno strukturo. Zbirke podatkov o prostoru niso nujno umeščene v prostor (geoorientirane) glede na geodetski referenčni sistem.
zvezno porazdeljena zbirka podatkov	federated database	Arhitektura porazdeljenega niča s čimer se poudarja, da se med procesorje delovnih postaj ne porazdeljujejo niti spominske enote niti hitri pomnilnik, vendar je med številne geografsko porazdeljene podatkovne strežnike porazdeljena zbirka podatkov.

SEZNAM UPORABLJENIH KRATIC

V seznamu uporabljenih kratic so navedene samo pomembnejše kratice, ki so uporabljene v monografiji.

Slovenska kratica	Angleška kratica	Razlaga kratice
CORBA	CORBA	splošna arhitektura posrednika objektnih povpraševanj / Common



## TERMINOLOGIJA

CASE	CASE	Object Request Broker Architecture računalniško podprto sistemsko inženirstvo/Computer Assisted System Engineering
UZDG	DBKM	Upravljanje znanja na osnovi dokumentarnih gradiv/Document-Based Knowledge Management
OL	DL	opisna logika/Description Logic
FOL	FOL	logika prvega reda / First Order Logic
GIS	GIS	geografski informacijski sistem / Geographic Information System
GML	GML	geografski označevalni jezik / Geographic Markup Language
GPS	GPS	sistem globalnega določanja položaja / Global Positioning System
KG	CG	konceptualni graf/Conceptual Graph
KDD	KDD	odkivanje znanja v zbirkah podatkov / Knowledge Discovery in Databases
KIF	KIF	izmenjevalni format znanja / Knowledge Interchange Format
OOAD	OOAD	objektno orientirana arhitektura in oblikovanje/Object Oriented Architecture and Design
PPO	DRS	predstavitve področja obravnave/Discourse Representation Structures
OIL	OIL	jezik za ontološko specifikacijo / Ontology Interchange Language
NOP	OLTP	neposredna obdelava poslov/On-Line Transaction Processing
OMG	OMG	Object Management Group
RBAC	RBAC	nadzor pristopa osnovan na vlogah / Role-Based Access Control
RDF	RDF	omrežje opisovanja virov (Resource Description Framework)
SDM	SDM	rudarjenje prostorskih podatkov / Spatial Data Mining
XML	XML	razširljivi označevalni jezik / Extensible Markup Language

## SEZNAM UPORABLJENIH OZNAK

Oznaka	Razlaga oznake
apso	absolutna prostranost semantičnega ogrodja
apzp	absolutna prostranost zbirke podatkov
arso	absolutna razsutost semantičnega ogrodja
arzp	absolutna razsutost zbirke podatkov po semantični globini
asgpz	absolutna semantična globina zbirke prostorskih podatkov
gnr	globina neobhodne ravni
r	semantična razdalja
rpzp	relativna prostranost zbirke podatkov
rrzp	relativna razsutosti zbirke podatkov po semantični globini
sgk	semantična globina koncepta
srzp	semantična razpršenost zbirke podatkov
unr	utež neobhodne ravni
urg	utež relativne globine

## TERMINOLOGIJA

# STVARNO KAZALO

- A  
agent 66, 67  
aktivnosti  
aktivnosti v prostoru 66  
    človeške aktivnosti 66  
    v prostoru 67  
arhitektura porazdeljenega niča 95  
Aristotelova metoda 39  
atributiranje 29  
avtonomnost 97
- B  
bogate lokalne sheme 23, 114, 115
- C  
celovitost predstavitve znanja 39
- D  
deduktivni pristop 103  
državna informacijska strategija 103, 104  
državna infrastruktura prostorskih podatkov 103, 104
- E  
enakovrednost 30  
enakozvočnice 21  
entiteta 27, 36  
epistemologija 62
- F  
filozofija  
    mejni spori 48  
    teorija večrazsežnostne nepretgranosti 46  
filozofske teorije  
    mereologija 46
- mereološki aktualizem 46  
    mereološki potencializem 46  
    skrajni idealizem 46  
    skrajni realizem 46  
    zmerni realizem 46
- G  
generalizacija 29, 39  
geografska dejstva 51  
geografski informacijski sistem 13  
geografski podatkovni nizi 70, 111  
geoinformatika  
    človeške aktivnosti 64  
    človeške aktivnosti v prostoru 66  
    hierarhija aktivnosti 66, 67  
    izgradnja ontologij 64  
    koncept omogočanja 66, 67  
    konceptualizacija sveta 66, 67  
    metapodatkovni sistem 64  
    ontologija geoinformacijske infrastrukture 107  
    ontologija tipoloških elementov 66
- GIS 66  
globalna shema 115
- H  
Helsinško načelo 99  
hierarhija abstrakcij 38, 73
- I  
identifikacija 29  
informacije 13, 17, 50, 59  
    dodatne informacije 23  
    geografske 97, 99  
    standardi 97  
    geografske informacije 64

- informacije o objektih 21
  - informacije o prostoru 51
  - informacijska podpora odločanju 23
  - integracija informacij 62
    - iskanje informacij z ontologijo 61
    - izmenjava informacij 61
    - ključne informacije 51
    - miselne obdelave podatkov 48
    - odvisnost informacij od konteksta 66
    - odvisnost pomena 67
    - porazdeljevanje informacij 62, 111
    - prostorske informacije, temeljni viri 51
    - selektivna izbira informacij 50
    - semantične informacije 25, 111
    - skrite 115
    - skrite informacije 19, 114
    - teorija informacij
    - semiotika 51
    - tok informacij 59
    - vrste informacij 50
  - informacijski sistem
    - ontologija 61
  - integracija
    - horizontalna 99
  - integracija podatkov
    - hibridni pristop 17
    - navidezni pristop 17
    - snovni pristop 17
  - integracija pogledov 19
  - integracija shem 11
    - analiza skladnosti 30
    - ciklične generalizacije 21
    - časovni vidik shem 24
    - faze integracije shem 24
    - globalna shema 23, 25, 114
  - integracijska moč 27
    - kakovost integracije shem 24
    - kartiranje shem 27
    - kartiranje spajanja shem 30
    - lokalna shema 25
    - minimalnost integralne sheme 24
    - najmanjši zgornji obseg 24
    - popolnost integracije 24
    - prilagajanje shem 24
    - primerjava shem 25
    - pristop navideznega pogleda 24
    - razumljivost integralne sheme 24
    - relativnost skladnostnega razmerja 29
    - semantično poenotenje 25
    - skladnostne trditve 29
    - skladnostni tip 29
    - spajanje shem 23, 24, 27, 30
    - transformacija sheme 24
    - vzorci shem, metoda kopičenja 19
    - združevalno omrežje 27
    - združljivost 29
  - integracija zbirk podatkov 17, 19
  - istovetnost 21, 30
- J
- javni interes 104
  - jezik 99
    - besedna dvoumnost 53
    - besedna redundanca 53
    - formalen jezik 41
  - jezik in pravilnost dokazovanja 53
  - jezik za predstavitev pomenalogika 59
  - jezikovni simbolibesede 53
    - načelo kompozicije 53
    - naravni jezik 53
    - semantika Tarskega 53
    - strukturna nejasnost 53
    - strukturna redundanca 53
    - trditvena logika 53
    - umetni jezik 53
    - večpomenski izrazi 27
- K
- karte
    - izvrševalna uporaba geografskih kart 48
  - kartografske 50
  - kognitivne 48
    - miselne karte 39
  - kartiranje shem 30
  - klasifikacija 17, 29, 51, 111, 113, 115
    - standardna klasifikacija dejavnosti (SKD) 114
  - kognitivna klasifikacija 64
  - kognitivne aktivnosti 36, 46, 70, 111
  - kognitivno kartiranje 48
    - besedno izražanje 50
    - funkcionalna pomembnost 50
    - interakcija z okoljem 50
  - kognitivne podobe 50

- osebni pomen 50
- prostorska semantična hierarhija 50
- razločevalnost/likovitost 50
- selektivnost 50
- koncept 21, 23, 24, 29, 34, 36, 38, 39, 50, 67, 73, 81, 91, 107, 113, 115
- konceptualni formalizem 99
- konceptualni graf 23, 36, 41, 56, 59
- konceptualni model 99
- konceptualni sistem 36
- konceptualno modeliranje 36
  - jezik 53
  - nominalna osnova 36
  - standardi 97
- Konzorcij odprtega GIS (OGC) 99
- L
- logika
  - odvisnost od nedavne omembe 56
  - ontologija 56
- lokalna shema 17, 25, 114, 115
- lokalni model semantike 17
- M
- medopravilnost 61, 99
  - GIS 97
- medopravilni GIS 64
  - opredelitev ontologije za 13
  - razmerja agentov 13
  - standardi 97
- meje 46, 48
- merjenja semantične podobnosti 21, 114, 115
- metapodatki 61, 99, 107
  - katalog 99
- metoda sinteze 39
- miselna predstavitev 48
- miselna zasnova 36, 39
- model 30, 36, 99, 114, 115, 147, 150
- N
- neposredna obdelava poslov 17
- neposredno analitično procesiranje 17
- nerazločnost konceptov 21
- O
- objekt
  - enostavni objekt 48
  - homogeni prostorski objekti 46
  - naravnibona-fideprostorski objekt 46
  - nepredmetni 68
  - objektadministrativna entiteta 46
  - objektpojavniz neizrazito zunanjo mejo 48
  - politično-administrativni objekt 48
  - predmetni 68
  - prostorski objekt 46
  - prostorski objekti geografskih razsežnosti 68
  - umetni fiat prostorski objekt 46
  - objektni identifikator 41
    - jezikovni simbol 41
  - objektni sistem
    - dogodek 41
  - obnavljanje podatkov 13
  - obrnjeno inženirstvo 19
  - odkrivanje pomena simbolov 21
  - odkrivanje semantičnih podobnosti 25
  - ontologij
    - analiza aktivnosti domene 67
  - ontologija 70, 76, 89, 91, 111, 115, 119
    - aplikacijska 70
    - avtor ontologije 62
    - formalna ontologija 62
    - geoprostorska stvarnost 48
    - gradnja 70
    - iskanje informacij in ontologija 62
    - izboljšava komunikacije 61
    - količina pomena in ontologija 62
    - lahka ontologija 62
    - logika in ontologija 56
    - ontologija in sistemski inženiring 61
    - ontologija domene 23
    - ontologija geografskih struktur 11
    - ontologija geoprostorske stvarnosti 64
    - ontologija GIS 64
    - ontologija in konceptualni grafi 59
    - ontologija in use case 61
    - ontologija v geoinformatiki 64
    - ontološka obveznost 64
    - ontološke predpostavke 56
    - opredelitev
      - filozofska opredelitev ontologije 62
      - informacijska opredelitev ontologije 62
      - porazdeljevanje ontologij 61
      - pravila modeliranja 70

- ravni ontologije 61
- različnost ontologij 36
- razvoj ontologije 66
- scenarij uporabe ontologije 61
- splošna ontologija 62
- univerzalna ontologija 62
- uporaba ontologije 62
- večnivojska 67
- večnivojska ontologija 66
- visokonivojska ontologija 13
- vloge igralcev v ontologiji 62
- značilnosti ontologije 61
- zrelost ontologije 61
- ontologija prostora 11
  - geografskih razsežnosti 68
  - gradnja ontologije 70
  - koncepti razsežnosti 70
  - lokacija 70
  - natančna 70
  - približna 70
  - naravno učenje 76
- ontologija prostora geografskih razsežnosti (OPGR) 107
  - ravni ontologije 70
  - semantične mere 85
  - semantični prameter 107
  - semantično omrežje 68
  - stvarne ravni 76
  - vključevalna pravila 70
  - zapleteni objekti in sheme 73
  - zdravorazumsko mišljenje 70
- ontologija prostora geografskih razsežnosti 107
  - opredelitev 38, 56, 70, 113
    - namen agenta 56
    - neformalna opredelitev 62
    - zdravorazumsko mišljenje 70
  - opredelitev geometričnih struktur 13
  - opredeljevanje simbolov 38
    - dogovorne opredelitve 38
  - opisne opredelitve 38
  - ovojnik 17
- P
- podatki
- porazdeljevanje 62
- podatkovni viri 13, 17
- podobnost 29, 30, 111
- področje obravnave 36, 39, 41, 99
- pojavnne oblike 73
- pomenska neskladnost 21
- porazdeljevanje 17, 61, 62, 97, 111, 113
- porazdeljevanje podatkov 17
- posredovalni sistem 17, 68
- pragmatika 53, 61
  - govorni nastop 61
- prameter 107
- predintegracija 25
- primerjava shem 114, 115
  - dedukcija z dedovanjem 27
  - indukcija 27
  - metoda enakovrednosti atributov 27
  - metoda mehke terminologije 27
  - metoda porazdeljenih ontologij 27
  - modelna dvoumnost 27
  - največji razpon drevesa 27
  - semantična dvoumnost 27
  - semantično poenotenje 27
  - stopnja podobnosti 27
  - stopnja skladnosti 29
  - terminološko drevo 27
- pristop navideznega pogleda 17
- posredovalni sistem 30
- programski agent 95
- prostorske pojmovne sheme 73
- prostorski položaj 48
- prostorski predmetni objekti 73
- protokol s pristopom poriva 97
- protokol s pristopom potega 97
- protokol zakupa 97
- R
- računalniški sistemi 36
- raziskovanje znanja
  - iskanje v globino 76
  - iskanje v širino 76
- raznolikost semantike 21, 23
- raznovrstnost 17
- redundanca 13, 17
- referenčne zbirke 104
- S
- samodejna integracija informacij 17, 113
- semantična bogatitev shem 19, 23, 114, 115
- semantična dvoumnost 113

- semantična medopravilnost 13, 70  
 semantična pristranost 78  
 semantična razmerija 68, 89, 99, 115  
 semantična razmerja 19, 21, 70, 79, 113
  - agregatni razred 70
  - enakovredni razredi 70
  - klasifikacija 113
 semantična raznovrstnost 21  
 semantična skladnost 78  
 semantična sorodnost 113, 115  
 semantična združljivost 113  
 semantične mere 79
  - absolutna prostranost 85
  - normalizirana semantična razdalja 29
  - razpršenost 86
  - razsutost 86
 semantična globina 78, 81  
 semantična razdalja 29  
 semantična razpršenost 78
  - utež neobhodne ravni 85
  - uteži semantičnih razdalj 85
 semantični podatkovni model 36  
 semantični slovar 19, 23, 76
  - razširjeni slovar sopomenk 23
 semantično omrežje 13, 19, 79, 114, 115  
 semantika 53
  - predpostavke o podatkih 97
 semantična dvoumnost 27  
 semiotika 51  
 shema  
 standardna 99  
 simbol 21, 23, 27, 62, 73, 89, 119  
 simbol v trditveni logiki 53  
 sintaksa 53  
 skladnostna trditev 29  
 smerno razmerje 34  
 sosedski graf 34  
 splet 95
  - dvoumnost 56
  - nadzorovani naravni jezik 58
  - poimenovalni dogovori 56
 standard
  - družina ISO 19100 97
  - družina standardov STEP 61
  - načela konceptualnega modeliranja 99
  - okolja odprtih istemov
  - OSE 97
  - standard XML 48
  - standard za konceptualne grafe 41
  - standardni vmesnik za doseganje podatkov SDAI 17
    - za geografske podatke in geomatiko 97
    - za metodologijo objektnih katalogov 99
    - za modeliranje konceptualnih shem ISO/IEC 14481 (CSMF) F) 97
      - za referenčni model ISO 19101 97
  - strukturna neskladnost 21
  - stvarnost 36, 38, 51, 64, 70, 111
- T
- tehnika opredmetenja razmerij 24  
 teorija konceptualnih grafov 23  
 teorija mehkih množic 23  
 topologija 34
  - sosedska razmerja 34
 transparentnost 95  
 trinivojska arhitektura 95
- U
- ujemanje oblik meja 34  
 umetne nevronske mreže 19  
 usmerjeni graf 30
- V
- večpomenski izrazi 113  
 vozlišča 38
- Z
- združljivost 115  
 zgodovinska analiza shem 24  
 znanje
  - celovitost predstavitve znanja 39
  - povzemanje znanja iz pisnih listin 58
  - pridobivanje semantičnega znanja 23, 25
    - prostorsko 51
    - prostorsko znanje
    - konfiguracijsko znanje 51
    - postopkovno 111
    - semantično znanje integracijsko .....
      - orodje 23
      - sistem za povzemanje znanja 58
  - zvezno porazdeljeni informacijski sistem 17

## KAZALO SLIK

1	Prikaz osnovnih pojmov.....	10
2	Omrežje za doseganje semantične medopravnosti zbirk podatkov.....	18
3	Ravni znanja pri integraciji shem.....	19
4	Metamodel za skladnostne trditve .....	29
5	Razmerje med ontologijo, entitetno-relacijskim modelom in shemami .....	36
6	Richards-Ogdenov semantični trikotnik .....	37
7	Predstavitev koncepta s simbolom koncepta .....	39
8	Simetrični diagram.....	52
9	Primer uporabe ACE.....	58
10	Tok informacij od dokumentov do računalniške predstavitve .....	58
11	Spekter ontologij.....	60
12	Splošna ontologija predstavitev znanja.....	63
13	Konceptualno omrežje prostora geografskih razsežnosti.....	72
14	Ontologija prostora geografskih razsežnosti (OPGR).....	74
15	Grafični prikaz semantičnih parametrov zbirk prostorskih podatkov.....	77
16	Grafični prikaz semantične globine zbirke prostorskih podatkov. ....	80
17	Prikaz semantične globine koncepta.....	82
18	Prikaz absolutne semantične globine.....	82
19	Semantične razdalje razmerij med koncepti .....	83
20	Relativna prostranost in relativna globina zbirke podatkov.....	84
21	Medopravnost informacijskih sistemov.....	91
22	Grafična predstavitev vloge konceptualnega modeliranja .....	97
23	Sestavni deli deduktivne objektno usmerjene zbirke prostorskih podatkov .....	101
24	Različne razsežnosti geografskih podatkov v državni informacijski strategiji.....	103
25	Ocenjevanje skladnosti konceptov in povezovanje katalogov dveh zbirk.....	106
26	Povezovanje konceptov geografskega prostora in konceptov tehnologije GIS.....	108



## KAZALO PREGLEDNIC

1	Semantične prvine izražene v naravnem jeziku.....	54
2	Opredelelitev logičnih operatorjev .....	54
3	Ontologija prostora geografskih razsežnosti - prva raven.....	73
4	Ontologija prostora geografskih razsežnosti - prva in druga raven.....	75
5	Pripomoček za izračun semantičnih globin konceptov.....	81
6	Določanja semantične globine koncepta na osnovi govorečih šifer.....	81
7	Primer izračuna absolutne semantične globine.....	82
8	Določitev globin neobhodnih ravni semantičnih razmerij.....	83
9	Izračun semantičnih razdalj (r) semantičnih razmerij.....	84
10	Izračun uteži neobhodne ravni (unr) semantičnih razmerij.....	85
11	Izračun uteženih semantičnih razdalj.....	85
12	Izračun uteži relativne globine (urg) semantičnih razmerij.....	87
13	Izračun uteži semantične razpršenosti (usr) zbirke podatkov.....	87
14	Izračun uteženih semantičnih razdalj.....	88
15	Izračun semantične ustreznosti OPRG.....	88
16	Analiza izrazov semantično nezdržljivih konceptov.....	89
17	Primerjava med operacijskimi sistemi POS, MOS in VOS.....	93
18	Lestvica možnih semantičnih razmerij.....	105
19	Analiza izrazov semantično nezdržljivih konceptov.....	117
20	Določitev semantičnih razmerij med simboli katalogov.....	118

## LITERATURA

Abdelmoty, A., Williams, M., Paton, N., 1993: Deduction and Deductive Databases for Geographic Data Handling. *Advances in Spatial Databases, 3rd International symposium SSD'93*, Singapore, Springer Verlag, Berlin, Nemčija.

Abiteboul, S., Hull, R., 1987: IFO: a formal semantic database model. *Transactions on Database Systems*, Year 12, No. 4, ACM, New York, ZDA.

Albrecht, J., 1999: Towards interoperable geo-information standards: A comparison of reference models for geo-spatial information. *The Annals of Regional Science* No. 33, Springer Verlag, Berlin, Nemčija.

Anderson, N., 1981: Foundations of information integration theory. *Methods of information integration theory*, Academic Press, New York, ZDA.

Assenova, P., 1999: Formalism and Schema Evolution. Department of Computer and Systems Sciences, Stockholm University, Royal Institute of Technology, Stockholm, Švedska.

Assenova, P., Johannesson, P., 1996: Improving Quality in Conceptual Modelling by the Use of Schema Transformations. *15th International Conference on Conceptual Modelling, ER '96*, Springer Verlag Berlin, Nemčija.

Banares, J., Zaragaga, F., Nogueras, J., Gutierrez, K. and Muro-Medrano, P., 2000: Construction and Use of Concept Hierarchies from Word Taxonomies for Searching Geospatial Data. *EuroConference - Ontology and Epistemology for Spatial Data Standards*, La Londe-les-Maures, Francija.

Barja, M., Bratvold, T., Myllymaki, J., Sonnenberger, G., 1998: Informia: a Mediator for Integrated Access to Heterogeneous Information Sources. *7th International Conference on Information and Knowledge Management, CIKM '98*, Bethesda, Maryland, ZDA.

Battini, C., Ceri, S., Navathe, S., 1992: *Conceptual Database Design*. Benjamin-Cummings Publications, ZDA.

Battini, C., Lenzerini, M., Navathe, S., 1986: A comparative Analyses of Methodologies for Database Schema Integration. *ACM Computing Surveys*, Vol. 18, ZDA.

Becam, A., Miquel, M., Laurini, R., 2000: A Distributed Environment using Ontology for the Interoperability of Urban Data and Models. *EuroConference - Ontology and Epistemology for Spatial Data Standards*, La Londe-les-Maures, Francija.

Benslimane, D., Leclercq, E., Savonnet, M., Terrasse, M., Yetongnon, K., 2000: On the definition of generic multi-layered ontologies for urban applications, *Computers, Environment and Urban Systems* 24, Pergamon.

Bishr, Y., 1997: *Semantic Aspects of Interoperable GIS*. Doctor Thesis, ITC Publication Series, No. 56, Enschede, Nizozemska.

Bittner, T., 1999: Rough location. Department of Geoinformation, TU Vienna, Avstrija.

Bittner, T., Frank, A., 1999: On design of formal theories of geographic space. *Journal of geographical systems*, Springer Verlag, Berlin, Nemčija.

Boman, M., Bubenko, J., Johannesson, P., Wangler, B., 1997: *Conceptual Modelling*. Prentice Hall, London, Velika Britanija.

Bouguettaya, A., Benatallah, B., Elmagarmid, A., 1998: *Interconnecting Heterogeneous Information Systems*. Kluwer Academic Publishers, Dordrecht, Nizozemska.

- Bouzeghoub, M., Comyn-Wattiau, I., 1990: View Integration by Semantic Unification and Transformation of Data Structures. 9th International Conference on Entity-Relationship Approach, Lausanne, Švica.
- Bratko, I., Džeroski, S., Kompare, B., Walley, W., 1998: Analyses of Environmental data with machine learning methods, Center for knowledge transfer in information technologies, Institut Jožef Štefan, Ljubljana.
- Bray, T., 1998: "RDF and Metadata" (<http://www.xml.com/pub/a/98/06/rdf.html>).
- Brentano, F., 1988: Philosophical Investigations on Space, Time and the Continuum. Crom Helm, London, Velika Britanija.
- Bruegger, B., 1996: Spatial Theory for the Integration of Resolution - Limited Data. Doktorska disertacija, ETH, Swiss Federal Institute of Technology, Švica, The Graduate School, University of Maine, Maine, ZDA.
- Bucher, B., 2000: Users access to geographic information resources: a model of tasks and roles to specify intentional uses regarding available resources. 2000, EuroConference - Ontology and Epistemology for Spatial Data Standards, La Londe-les-Maures, Francija.
- Buneman, P., Davidson, S., Kosky, A., Van Inwegen, M., 1992: A basis for Interactive schema Merging. Hawaii International Conference on System Sciences, Hawaii, ZDA.
- Camara, G., Monteiro, V., Paiva, J., Souza, R., 2000: Action-driven ontologies of the geographic space. GIScience 2000, AAG, Savannah, GA, ZDA.
- Camara, G., Thome, R., Freitas, U., Miguel, A., Monteiro, V., 1999: Interoperability in practice: Problems in Semantic Conversion from Current Technology to Open GIS. Interoperating Geographic Information Systems, Second International Conference INTEROP '99, Zürich, Switzerland, Springer Verlag, Berlin, Nemčija.
- Casati, R., Smith, B., Varazi, A., 1998: Ontological Tools for Geographic Representation. Guarino (ed.), Formal Ontology in Information Systems, IOS Press, Amsterdam, Nizozemska.
- Casati, R., Varazi, A., 1994: Holes and Other Superficialities. Cambridge, MIT Press, MA, ZDA.
- Casati, R., Varazi, A., 1995: The structure of spatial localization. International Journal of Philosophical studies, 82 (2), Routledge, Taylor & Francis Group, London, Velika Britanija.
- Cattell, R., Douglas, B., 2000: The Object Data Standard: ODMG 3.0, Morgan Kaufmann Publishers, San Francisco, ZDA.
- Ceri, S., Pagelatti, G., 1984: Distributed Databases, Principles and Systems. McGraw Hill, New York, ZDA.
- Chisholm, R., 1993: Spatial Continuity and the Theory of Part and Whole. Brentano Studien, No. 4, (<http://www.roell-verlag.de>).
- Couloris, G., Dollimore, J., Kindeberg, T., 2001: Distributed systems concepts and design. Addison Wesley, Publishers Ltd. Wokingham, Anglija.
- Convey, C., Karpenko, O., Tatbul, N., Yan, J., 2001: Data Integration Services. ([www.cs.brown.edu](http://www.cs.brown.edu)).
- Crown, D., 1990: GIS versus CAD versus DBMS: what are the differences? Introductory readings in Geographic information systems, Taylor and Francis, London, Velika Britanija.
- Čeh, M., Ferlan, M., 2001: Integration of spatial data bases as support to agriculture information system maintenance. 4th International Symposium Turkish-German Joint Geodetic Days, Berlin, Nemčija.
- Čeh, M., 2002: Analiza geodetskih podatkovnih zbirk za potrebe kmetijstva. Doktorska disertacija, Fakulteta za gradbeništvo in geodezijo, Univerza v Ljubljani, Slovenija.
- Devitt, M., Sterelny, K., 1987: Language and Reality. Basil Blackwell Publishing, Oxford, Velika Britanija.
- Devogele, T., Parent, T., Spaccapietra, S., 1998: On spatial database integration. International Journal Geographic Information Science, Vol. 12, No. 4, Taylor and Francis, London, Velika Britanija.
- Downs, R., Stea, D., 1977: Maps in mind, reflections on cognitive mapping. Harper and Row Publishers, New York, ZDA.
- Egenhofer, M., 1991: Reasoning About Binary Topological Relations. Advances in Spatial Databases, SSD '91, LNCS 525, Springer Verlag, Berlin, Nemčija.
- Egenhofer, M., Mark, D., 1995: Naive geography. COSIT '95, LNCS 988, Springer Verlag, Berlin, Nemčija.

- Ekenberg, L., Johannesson, P., 1995: Conflictfreeness as a Basis for Schema Integration. CISM0D'95, LNCS, Springer Verlag, Nemčija.
- Embley, D., Campbell, D., Smith, R., 1998: Ontology-Based Extraction and Structuring of Information from Data-Rich Unstructured Documents, CIKM '98, Bethesda, Maryland, ZDA.
- Ester, M., Kriegel, H., Sander, J., 1997: Spatial Data Mining: A Database Approach. Advances in Spatial Databases, 5th International Symposium, SSD '97, Berlin, Nemčija.
- Fankhauser, P., 1997: A Methodology for Knowledge - Based Schema Integration. Doktorska disertacija, Fakultät Informatik, Technischen Universität Wien, Wien, Avstrija (<http://www.darmstadt.gmd.de/oasys/reports/ftp/pdf/P1997-34.pdf>).
- Fankhauser, P., Neuhold, E., 1993: Knowledge based integration of heterogenous databases. Interoperable Database Systems, (DS-5), North-Holland, Amsterdam, Nizozemska.
- Fayyad, U., Piatetsy-Shapiro, G., Smyth, P., Uthursamy, R., 1996: Advances in knowledge discovery and data mining. AAAI/MIT Press, Cambridge, MA, ZDA.
- Fellbaum, C., 1999: WordNet - an electronic lexical database, MIT Press, Cambridge, MA, ZDA.
- Fensel, D., Horrocks, I., Harmeln, F., Decker, S., Erdmann, M., Klein, M., 2000: OIL in a Nutshell. European Workshop on Knowledge Acquisition, Modeling and Management, Lecture Notes in Artificial Intelligence, Springer Verlag, Berlin, Heidelberg, Nemčija.
- Fischer, M., Leung, Y., 1998: A Genetic Algorithms based evolutionary computational neural network for modelling spatial interaction data. The Annals of Regional Science, No. 32, Springer Verlag, Berlin, Nemčija.
- Fowler, M., 1997: Analyses Patterns - Reusable Object Models. Addison-Wesley, New York, ZDA.
- Frank, A., 1999: Spatial Ontology: A geographical Information Point of View, Spatial and Temporal reasoning. Kluwer Academic Publishers, Dordrecht, Nizozemska.
- Frank, A., Raubal, M., van der Vlugt, M., 2000: PANEL-GI Compendium, A Guide to GI and GIS, Pan European link for Geographical Information, European Commission, Italija.
- Frege, Gottlob, 1879: Begriffsschrift. English translation in J. van Heijenoort, ed. (1967) From Frege to Gödel, Harvard University Press, Cambridge, MA, ZDA.
- Freundschun, S., 1991: Spatial Knowledge Acquisition of Urban Environments from Maps and Navigation Experience. Doktorska disertacija, Department of Geography, State University of New York at Buffalo, ZDA.
- Fuchs, N. E., Schwertel, U., Schwiter, R., 1998: Attempto Controlled English - not just another logic specification language. Proceedings LOPSTR '98, Manchester, Velika Britanija.
- Gallagher, L., Carnahan, L., 2000: A general purpose registry/repository information model. NIST Information Technology Laboratory, Washington, DC, ZDA.
- Gamut, L., 1991: Logic, Language, and Meaning. The University of Chicago Press, ZDA.
- Gerbé, O., Perron, M., 1995: Presentation definition language using conceptual graphs. Conceptual Structures: Applications, Implementation, and Theory, Lecture Notes in AI 954, Springer Verlag, Berlin.
- Gerbé, O., 1997: Conceptual graphs for corporate knowledge management. Conceptual Structures: Fulfilling Peirce's Dream. Lecture Notes in AI 1257, Springer Verlag, Berlin, Nemčija.
- Gerbé, O., Guay, B., M. Perron, 1996: Using conceptual graphs for methods modeling. Conceptual Structures: Knowledge Representation as Interlingua, Lecture Notes in AI 1115, Springer Verlag, Berlin.
- Gerbé, O., Keller, R., Mineau, G., 1998a: Conceptual graphs for representing business processes in corporate memories. Conceptual Structures: Theory, Tools, and Applications, Lecture Notes in AI 1453, Springer Verlag, Berlin, Nemčija.
- Gerbé, O., Kerhervé, B., 1998b: Modeling and metamodeling requirements for knowledge management. OOPSLA '98 Workshops, Vancouver, Kanada.
- Ghidini, C., Serafini, L., 1998: Model Theoretic Semantics for Information Integration. AIMSA '98,

- LNAI 1480, Springer Verlag, Berlin, Heidelberg, Nemčija.
- Gibson, J., 1986: *The Ecological Approach to Visual Perception*. Lawrence Erlbaum, Hillsdale, NY, ZDA.
- Giger, C., Najar, C., 2003: *Ontology-based integration of data and metadata*. AGILE, Lyon, Francija.
- Gomez, A., Benjamins, R., 1999: *Overview of Knowledge Sharing and Reuse Components: Ontologies and Problem-Solving Methods*. Workshop on Ontologies and Problem-Solving Methods, IJCAI '99, Stockholm, Švedska.
- Gruber, T., 1993b: *Toward principles for the design of ontologies used for knowledge sharing*. Workshop on Formal Ontology, Padova, Italija.
- Gruber, T., 1993a: *A translation approach to portable ontologies*. Knowledge Acquisition, No. 5, Academic Press, New York, ZDA.
- Guarino, N., 1994: *The ontological level*. Philosophy and the cognitive science, Hoelder-Pichler-Tempsky, Vienna, Avstrija.
- Guarino, N., Masolo, C., Vetere, G., 1999: *Ontoseek: Using large linguistic ontologies for accessing on-line yellow pages and product catalogs*. IEEE Intelligent Systems, No. 14.
- Hakkarainen, S., 1999: *Dynamic Aspects and Semantic Enrichment in Schema Comparison*. Doktorska disertacija, Department of Computer and Systems Science, Stockholm University, Royal Institute of Technology, Stockholm, Švedska.
- Harvey, F., Kuhn, W., Pundt, H., Bishr, Y., Riedmann, C., 1999: *Semantic interoperability, A central issue for sharing geographic information*. The Annals of Regional Science No. 33, Springer Verlag, Berlin, Heidelberg, Nemčija.
- Hayne, S., Ram, S., 1990: *Multi-user View Integration System (MUVIS): An Expert System for View Integration*. 6<sup>th</sup> IEEE Conference on data engineering, Los Angeles, ZDA.
- Held, R., Rekosh, J., 1963: *Motor-sensory feedback and geometry of visual space*. Science, No. 141, New York, ZDA.
- Hull, R., King, R., 1987: *Semantic data base modelling: survey, applications and research issues*. Computing Survey, Vol. 19, No. 3, ACM, New York, ZDA.
- ISO/IEC 13249-1, 2000: *SQL Multimedia and Application Packages (SQL/MM)*, International organisation for standardisation, Geneve, Švica.
- ISO/TC 211, 2000: *Draft international standard ISO/DIS 19101 (Geographic information – Reference model)*. International organisation for standardisation Geneve, Švica.
- ISO/TC 211, 2000: *Draft international standard ISO/DIS 19110 (Geographic information – Feature cataloguing methodology)*. International organisation for standardisation Geneve, Švica.
- ISO/TC 211, 2001: *Draft international standard ISO/DIS 19113 (Geographic information – Quality principles)*. International organisation for standardisation, Geneve, Švica.
- Jasaper, Z., Uschold, M., 1999: *A Framework for Understanding and Classifying Ontology Applications*. IJCAI '99 Ontology workshop proceedings, Stockholm, Švedska.
- Johannesson, P., 1993: *Schema integration, Schema Translation, and Interoperability in Federated Information Systems*. Doktorska disertacija, KTH, Stockholm University, Stockholm, Švedska.
- Kamp, H., 1981b: *A theory of truth and semantic representation*. Formal Methods in the Study of Language, Mathematical Centre Tracts, Amsterdam, Nizozemska.
- Kangassalo, H., 1993: *COMIC: A System and Methodology for Conceptual Modelling and Information Construction*. Data and Knowledge Engineering, No. 9, North-Holland, Nizozemska.
- Kaptelinin, V., Nardi, B., Macaulay, C., 1999: *The activity checklist: a tool for representing the space of context*. Human-Computer Interaction. Vol. 1, No. 4, Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey, ZDA.
- Keim, D., Kriegel, H., Seidl, T., 1994: *Supporting Data Mining of Large Databases by Visual Feed-*

- back Queries. Proceedings of 1994 IEEE 10th International Conference on Data Engineering, Houston, IEEE Comput. Soc. Press, Los Alamitos, California, ZDA.
- Kemppainen, H., 1995: On the formalization of the semantics of spatial objects. Reports of the Finnish Geodetic Institute, Helsinki, Finska.
- Knorr, E., Ng, R., Shilvok, D., 1997: Finding Boundary Shape matching Relationships in Spatial Data, Advances in Spatial Databases. 5th International Symposium, SSD '97, Berlin, Nemčija.
- Koepsell, D., 2001: Prospectus of Center for Commercial Ontology. Business Webs and Ontology, (<http://www.acsu.buffalo.edu/~koepsell/center.htm>).
- Kononenko, I., 1997: Strojno učenje. Fakulteta za računalništvo in informatiko, Univeza v Ljubljani, Ljubljana.
- Kosky, A., 1994: Formal model for Databases with Applications to Schema Merging. Specifications of Database Systems, Harper and Norrie Editors, Glasgow, Škotska.
- Kuhn, W., 1996: Semantics of geographic information. Geo-info series, Vol. 7, Technical university Vienna, Dept. of Geoinformation, Avstrija.
- Kuhn, W., 2001: Ontologies in support of activities in geographical space. International Journal of Geographic Information Science, Vol. 15, No. 7, Taylor Francis Ltd., London, Velika Britanija.
- Kuhn, X., Raubal, M., 2003: Implementing semantic reference systems. AGILE, Lyon, Francija.
- Kuipers, B., 1978: Modelling spatial knowledge. Cognitive Science, No. 2. Cognitive Science Society, University of Texas, Austin, ZDA.
- Kuipers, B., 1994: Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge. MIT Press, Cambridge, MA, ZDA.
- Lake, R., 2002: Impact of GML on Data Development (<http://www.galdosinc.com/Gml.htm>).
- Land Data Bank Legislation, 2000: Swede National Survey Schema, Gavle, Švedska.
- Laurel, B., 1993: Computers as Theatre. Addison-Wesley, Reading, Massachusetts, ZDA.
- Leung, Y., 1997: Intelligent spatial decision support systems. Springer Verlag, Berlin, Nemčija.
- Leung, Y., 2000: Some thoughts on spatial data structure analyses. Geographical Systems, Springer Verlag, Berlin, Nemčija.
- LGSD, 2000: Licence agreement for Geografiska Sverigedata. Lantmateriet, Gavle, Švedska.
- Liben, L., 1988: Conceptual issues in the spatial development of spatial cognition. Spatial Cognition: Brain Bases and Development, Hillsdake, New Jersey, ZDA.
- Lindland, O., Sindre, G., 1994: Understanding Quality in Conceptual Modelling. Information Systems, Vol. 11, No. 2, IEEE Software.
- Lu, W., Han, J., Ooi, B., 1993: Discovery of General Knowledge in Large Spatial Databases. Far East Workshop on GIS, Singapore, Kitajska.
- MacEachern, A., 1992: Application of environmental learning theory to spatial knowledge acquisition from maps. Annuals No. 82, Association of American Geographers, ZDA.
- Mark, D., 1993: Human Spatial Cognition, Human factors in Geographical Information Systems. Belhaven Press, New York, ZDA.
- Mark, D., 2000: Ontology of geographic object categories. EuroConference on Ontology and Epistemology for Spatial Data Standards-La Londe-les-Maures, Francija.
- Mark, D., Khun, W., Smith, B., Turk, A., 2003: Ontology, natural language, and information systems: implication of cross-linguistic studies of geographic terms, AGILE, Lyon, Francija.
- Masser I., 1998: Governments and Geographic Information. Taylor and Francis Ltd., London, Velika Britanija.
- Masser, I., Campbell, H., Craglia, M., 1996: GIS Diffusion: The Adoption and Use of Geographic Information Systems in Local Government in Europe. Taylor & Francis, London, Velika Britanija.
- McGuinness, D., 1998: Ontological issues for knowledge-enhanced search. Formal Ontology in Information Systems, Trento, Italija.

- Metais, E., Kedad, Z., Comyn-Watiau, I., Bouzeghoub, M., 1996: Implementation of a Third Generation View Integration Tool. Application of Natural Language to Information Systems, IOS Press, Amsterdam, Nizozemska.
- Miller, R., Ioannidis, Y., Ramakrishnan, R., 1994: Schema Equivalence in Heterogeneous Systems: Bridging Theory and Practice. Information Systems, Vol. 19, No. 1.
- Mirabel, I., 1995: Semantic integration of Conceptual Schemes. Workshop on Applications of Natural Language to Databases, NLDB '95, Versailles, Francija.
- Monmoyer, M., 1991: How to Lie With Maps. The University of Chicago Press, Chicago, ZDA.
- Motro, A., 1987: Superviews: Virtual Integration of Multiple Databases, Transactions on Software Engineering, Vol. 13, No. 7. IEEE Publications.
- Mylopoulos, J., 1998: Information Modeling in the Time of Revolution. Information Systems, Vol. 23, No. 1, Elsevier Science Ltd., Velika Britanija.
- Nyegres, 1992: Representing graphical meaning. Map generalization: rules for knowledge representation, Longman Scientific and Technical, Essex, Anglija.
- ODMG, 2000: The Standard for Storing Objects. Object Data Management Group, Burnsville, ZDA (<http://www.odmg.org/>).
- OGC, 2001: Geography Markup Language (GML) 2.0. OpenGIS Implementation Specification DN. 01-029 (<http://www.opengis.net/gml/01-029/GML2.html>).
- Open GIS Consortium - OGC, 1998: Technical Specifications.
- Openshaw, S., 1997: Artificial intelligence in geography. John Willey and Sons, New York, ZDA.
- Ordnance Survey, 1996: Results of the Consultation Exercise on "The national interests in mapping". Southampton, Velika Britanija.
- Østensen, O., 2001: The expanding agenda of Geographic information standards. ISO Bulletin, July 2001, International organisation for standardisation, Geneve, Švica (<http://www.iso.ch/iso/en/commcentre/pdf/geographic0107.pdf>).
- Oszu, M., Valduriez, P., 1999: Principles of Distributed Database Systems, Prentice Hall Publishers Ltd., New Jersey, ZDA.
- Peirce, C., 1885: On the algebra of logic. American Journal of Mathematics No. 7, Johns Hopkins University Press, Baltimore, ZDA.
- Phipps, S., 2000: Meaning, not Markup. Vol. 1, No. 1, XML Journal, SYS-CON Media, Inc., ZDA (<http://www.sys-con.com>).
- Pinker, S., 1994: The Language Instinct. William Morrow, New York, ZDA.
- Pogačnik, M., Privošnik, M., 1998: Inteligentni vs. Mobilni Agenti ([http://www.extra.research.philips.com/euprojects/storit/reports/infos98/pogacnik\\_infos98.ppt](http://www.extra.research.philips.com/euprojects/storit/reports/infos98/pogacnik_infos98.ppt)).
- Ravi, 1995a: Coordinating Body for Geographic Information in the Netherlands. Half-year report on Ravi Netherlands, Council for Geographic Information, Amersfoort, Nizozemska.
- Ravi, 1995b: Geoinformation terrain model. A Dutch standard for: Terms, definitions and general rules for classification and coding of objects related to the earth's surface (NEN3610), Ravi Netherlands Council for Geographic Information, Amersfoort, Nizozemska.
- Rhind, D., 1995: Spatial data from government. The 1995 AGI Source Book for Geographic Information Systems, Wiley, Chichester, Velika Britanija.
- Robnik - Šikonja, M., 1997: Razvoj hevrstike za usmerjanje učenja regresijskih dreves. Magistrska naloga, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana.
- Rodriguez, F., Bravo, G., Navarro, A., 1996: An ontology-based approach to spatial information modelling. SDH '96, International Geographical Union, Vol. II, Delft, Nizozemska.
- Sainsbury, M., 1991: Logical Forms. Basil Blackwell Publ., Oxford, Velika Britanija.
- Sester, K., Anders, K., Walter, V., 1998: Linking objects of different spatial data sets by integration and aggregation. Geoinformatica, Vol. 2, No. 4. Kluwer Academic Publishers, Dordrecht,

- Nizozemska.
- Shet, A., Larson, J., 1990: Federated Database Systems for managing Distributed, Homogeneous and Autonomous Databases. *Computing Surveys*. Vol. 22, No. 3, ACM, New York, ZDA.
- Sheth, A., Kashyap, V., 1993: So far (schematically) yet so near (semantically). IFIP '93, Elsevier Science Publishers B. V., North Holand, Nizozemska.
- Simons, P., 1987: *A study in Ontology*. Calarendon Press, Oxford, Velika Britanija.
- Skuce, D., 2000: Integrating web-based documents, shared knowledge bases, and information retrieval for user help. *Computational Intelligence* Vol. 16, No. 1, Blackwell Publishers, Oxford, Velika Britanija.
- Smith, B., 1992: Formal Ontology, Common Sense and Cognitive Science. *International Journal of Human Computer Studies*, No. 43, Academic Press, New York, ZDA.
- Smith, B., 1993: Ontology and the Logistic Analyses of Reality. International workshop on formal ontology in conceptual analyses and knowledge Representation, Analytic Phenomenology, Kluwer, Dordrecht, Netherlands.
- Smith, B., 1994: Fiat Objects, Parts and Wholes: Conceptual Part-Whole Relations and Formal Mereology. European Committee for Artificial Intelligence, Amsterdam, Nizozemska.
- Smith, B., 1995: More Things in Heaven and Earth. *Grazer Philosophische Studien*, No. 50, Buffalo, ZDA.
- Smith, B., Mark, D., 1998: Ontology and geographical kinds. 8<sup>th</sup> International Symposium on Spatial Data Handling, IGU Commission on GIS, Vancouver, Kanada.
- Smith, B., Mark, D., 2001: Geographical categories: an ontological investigation? *International Journal of Geographic Information Science*, Vol. 15, No. 7, Taylor Francis Ltd., London, Velika Britanija.
- Song, W., 1995: Schema integration, Principles, Methods and Application. Department of Computer and Systems Science, Stockholm University and Royal Institute of Technology, Stockholm, Švedska.
- Song, W., Johannesson, P., 1992: Semantic Similarity Relations in Schema Integration. International Conference on the Entity Relationship Approach, Karlsruhe, Nemčija.
- Sowa, J., 1984: *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley, Reading, MA, ZDA.
- Sowa, J., 1987: *Semantic Networks*. Encyclopedia of Artificial Intelligence, Wiley and Sons, New York, ZDA.
- Sowa, J., 1998: *Conceptual Graphs*. Draft proposed American National Standard, NCITS.T2/98-003 (<http://www.ncits.org/>).
- Sowa, J., 2000a: *Knowledge Representation, Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co., Pacific Grove, California, ZDA.
- Sowa, J., 2000b: *Ontology, Metadata, and Semiotics*. *Conceptual Structures: Logical, Linguistic, and Computational Issues*, Lecture Notes in AI No. 1867, Springer Verlag, Berlin, Nemčija.
- Spaccapietra, S., Parent, C., Dupont, Y., 1992: Model independent assertions for integration of heterogenous schemas. *VLDB Journal*, Vol. 1, Springer Verlag, Nemčija.
- Spanoudakis, G., Constantopoulos, P., 1994: Similarity for Analogical Software Reuse: a Computational Model. 11<sup>th</sup> European Conference on Artificial Intelligence, John Wiley and Sons, New York, ZDA.
- Stuckenschmidt, H., 2000: Knowledge based system for integration of geo data. Euro conference on Ontology and Epistemology for spatial data standards, La Londe-les Maures, Francija.
- Sutcliffe, A., Maiden, N., 1994: Domain Modelling for Reuse. International Conference Software Reuse, Rio de Janeiro, Brazilija.
- Šumrada, R., 1993: Uporaba CASE tehnologije in orodij za načrtovanje zemljiškega informacijs-



- kega sistema (IIS). Doktorska disertacija, Fakulteta za arhitekturo, gradbeništvo in geodezijo, Univerza v Ljubljani, Slovenija.
- Šumrada, R., 1999: Uporaba UML-ja za modeliranje sestave geografskih informacijskih sistemov. *Geodetski vestnik*, letnik 43, št. 4, Ljubljana.
- Tanenbaum, A., Steen, M., 2002: *Distributed systems Principles and Paradigms*. Prentice Hall Publishers Ltd., New Jersey, ZDA.
- Tarski, A., 1993: *Introduction to Logic and to the Methodology of Deductive Science*. Oxford University Press, Velika Britanija.
- Technical Committee ISO/TC 211 Geographic Information/Geomatics, 1999: International standard ISO 15046-1, *Geographic Information - Part 1: Reference model*. Norwegian Technology Standards Institution, Oslo, Norveška.
- Troyer, O., 1993: *On Data Schema Transformations*. Doktorska disertacija, Katholieke Universiteit Brabant, Nizozemska.
- Twardowski, K., 1979: *Actions and Products: Comments on the Boreder Area of Psychology, Grammar and Logic*. Semiotics in Poland, Reidel, Dordrecht, Nizozemska.
- Uitermark, H., 1996: The integration of geographic databases. Realising geodata interoperability through the hypermap and a mediator architecture, Second Joint European Conference, Exhibition on Geographical Information (JEC-GI, 96), Vol. 1, IOS Press, Barcelona, Španija.
- Uitermark, H., 2001: *Ontology-Based Geographic Data Set Integration*. Doktorska disertacija, Thesis Enschede, Nizozemska.
- Uitermark, H., Oosterom, P., Mars, N., Molenaar, M., 1993: *Onthology-Based Geographic Data Set Integration* (<http://ooa.kadaster.nl/~oosterom/STDBM993.PDF>).
- Uschold, M., 1998: Knowledge level modelling: Concepts and terminology. *Knowledge Engineering Review*, No. 13, Edinburgh, Škotska.
- Vckovski, A., 1998: *Interoperable and Distributed Processing in GIS*. Taylor and Francis, London, Velika Britanija.
- Veijalainen, J., 1988: *Multidatabase systems in ISO/OSI Environmets*. Standards in Information Technology and industrial Control, ed. Malagardis and Williams, North Holland, Nizozemska.
- Vet, P., Mars, N., 1998: Bottom-up construction of ontologies. *IEEE Transactions on knowledge and data engineering*, Vol. 10, No. 4, IEEE Computer Society (<http://www.computer.org/tkde>)
- Wegner, M., Masser, I., 1996: *Brave new GIS worlds. The Adoption and Use of Geographic Information Systems in Local Government in Europe*, Taylor & Francis, London, Velika Britanija.
- Wijngaarden, F., Putten, J., Oosterom, P., Uitermark, H., 1997: *Map integration. Update propagation in a multi-source environment*, ACM-GIS '97, Las Vegas, Nevada, ZDA.
- Winter, S., 2001: *Ontology: buzzword or paradigm shift in GI science?* *International Journal of Geographic Information Science*, Vol. 15, No. 7, Taylor Francis Ltd., London, Velika Britanija.
- Wen, L., 1998: *Constructing Information Systems Based on Schema Reuse*. Doctoral dissertation, Department of EECS, Northwestern University, ZDA.
- Wohed, R., 1997: *A Language for Enterprise and Information System Modelling*. Doktorska disertacija, Department of Computer and System Sciences, Stockholm University and Royal Institute of Technology, Stockholm, Švedska.
- Zdonik, S., Maier, D., 1990: *Readings in object-oriented database systems*. Morgan Kaufmann Publishers, San Mateo, California, ZDA.
- Zhan, C., Cox, M., Jones, D., 2000: *An Environment for Managing Enterprise Domain Ontology. Evolution to Buisness E-Commerce*, HICSS33, Maui, Hawaii, ZDA.
- Žalik, B., Podgorelec, D., 2001: *Geografski informacijski sistemi - integrativna informacijska paradigma*. Mednarodni posvet g-Slovenija v e-Evropi, *Geodetski vestnik* 46, št. 3. Zveza geodetov Slovenije, Ljubljana.



**ZRC SAZU, Založba ZRC**, Novi trg 2, 1001 Ljubljana, p. p. 306;  
tel.: 01/470 64 64; faks.: 01/425 77 94; e-pošta: [zalozba@zrc-sazu.si](mailto:zalozba@zrc-sazu.si);

[www.zrc-sazu.si/zalozba](http://www.zrc-sazu.si/zalozba)

ISBN 961-6500-03-1



9 789616 500036

ZALOŽBA ZRC, LJUBLJANA



ZALOŽBA  
Z R C