

MINISTRSTVO ZA NOTRANJE ZADEVE

**ODKRIVANJE IN KARTIRANJE NOVIH ZNANJ NA
OSNOVI POIZVEDB UPORABNIKOV
SPLETNIH STRANI MNISTRSTVA ZA NOTRANJE
ZADEVE RS**

OBDOBJE OD 12.05.2009 DO 09.02.2011

AVTOR: DR. KARL PETRIČ

LJUBLJANA, 2011

Kazalo

1 Uvod	3
1.1 Cilj raziskave	3
2 Analiza poizvedb za spletne strani MNZ	3
2.1 Priprava podatkov	3
2.1.1 Slika 1: Analiza poizvedb s pomočjo AntConc	4
2.1.2 Preglednica 1: Del klasificiranih udarnih ključnih besed	5
2.2 Procesiranje podatkov skozi različna programska orodja	6
2.3 Orange Canvas	6
2.3.1 Slika 1: Nadzorna plošča programskega orodja Orange Canvas	6
2.3.2 Slika 2: Distribucija ND glede na frekvenco in UDK razredov	7
2.3.2.1 Slika 3: Distribucija LK glede na frekvenco in UDK razredov	8
2.4.2 Slika 4: Distribucija LK glede na frekvenco in UDK razrede	9
2.4.3 Slika 5: Distribucija ranga glede na frekvenco in UDK razrede	9
2.4.4 Slika 6: Distribucija dejavnika K glede na frekvenco in UDK razrede	9
2.5 Rزتreseni diagram	11
2.5.1 Slika 7: Del matrike rزتresenih diagramov	11
2.5.2 Slika 8: Statistični atributi za UDK, LK in ND	12
2.5.3 Slika 9: Polivizualizacija K, LK in ND v odnosu z UDK	13

2.5.4 Slika 10: Krožni mrežni diagram za UDK in LK področja z dejavnikom K	14
2.5.5 Slika 11: Mrežni diagram za UDK, LK in UKB z najvišjimi dejavniki K	15
3 Analiza zgradbe in vsebine spletnih strani MNZ	17
3.1 Slika 12: Omrežje spletnih strani MNZ po področjih delovanja	17
3.2 Slika 13: Analiza omrežja MNZ s pomočjo NetDraw	18
3.2.1 Slika 14: Analiza pomembnejših vozlišč s pomočjo NetDraw	19
3.3 Preglednica 2: Del indeksiranih besed s frekvencami	20
3.3.1 Slika 15: Ontološki koncepti besednih grozdov	20
3.3.2 Slika 16: Glavni in podrejeni koncepti	21
3.3.3 Slika 17: Atlas besednih grozdov	22
3.3.4 Slika 18: Vizualizacija posameznih pod- konceptov	23
3.4 Analiza obiskov in ključnih besed	24
3.4.1 Preglednica 3: Del podatkov o koristnosti spletnih strani MNZ z vidika uporabnika	24
3.4.2 Preglednica 4: Del klasificiranih podatkov o obiskanih spletnih strani	25
3.4.3 Slika 19: Površinski delež posameznih KESS	26

3.4.4 Slika 20: Razvrščanje v skupine s pomočjo vizualizacijske tehnike SOM	27
4 Spoznanja	29
4.1 Slika 21: Znanstvena strokovna področja znotraj SOM znanja	29
4.1.1 Slika 22: Spekter potrebnih vsebin s strani uporabnikov	30
4.1.3 Slika 23: Spekter določenih delov in vrste informacij	32
4.2 Preglednica 5: Področja in profili	33
4.2 Slika 24: Hierarhično grozdenje področij zanimanja	34
4.3 Slika 25: Vizualizacija 368 besednih grozdov	35
4.4 Slika 26: Kartiranje potrebnih znanj z vidika uporabnikov	36
4.5 Slika 27: Ekstrahirano omrežje za UKB, UDK, LK in ND	38
4.6 Slika 28: Miselni vzorec za razvijanje nove ideje o novi aplikaciji	39
4.7 Slika 29: Mrežni graf analiziranih predstavitvenih besedil	40
4.8 Slika 30: Slikovni preglednični seznam s povezavami	41
4.9 Slika 31: Možen model aplikacije za javna naročila na MNZ	43
5 Zaključek	45
6 Viri in programska orodja	46

1 Uvod

Metode s področja odkrivanja zakonitosti v podatkih in / ali besedilih (v nadaljevanju: OZVP/B) navadno zahteva veliko količino podatkov, da bi lahko raziskovalec uresničil določen namen ali namene. Nameni raziskovalca t.i odkritelja ali bolje razvijalca novih idej / znanj so lahko zelo različni (npr. razvijanje programskega algoritma, izboljšanje storitve). V tem delu bom poskusil odkriti nova znanja s pomočjo različnih metod s področja OZVP/B in na osnovi številnih poizvedb na svetovnih in domačih iskalnikih, ki so v nadaljevanju usmerili uporabnike na spletne strani Ministrstva za notranje zadeve (v nadaljevanju: MNZ).

1.1 Cilj raziskave

- a. odkrivanje znanja na podlagi analize poizvedb in vsebine spletnih strani MNZ
- b. kartiranje uporabnih in potrebnih znanj s pomočjo OZVP/B z vidika uporabnika

2 Analiza poizvedb za spletne strani MNZ

V tem poglavju bom predstavil analizo z obsežne raziskave, ki je bila izvedena v letu 2011.

2.1 Priprava podatkov

Iz Google analytics sem v Excel 2007 izvozil 58329 poizvedb vključno s podatki o številu obiskov, ki so bili posledica le-teh (podatki so bili pridobljeni v času od 12.05.2009 do 09.02.2011). V nadaljevanju je šlo za večplastno pripravo podatkov, pri čemer sem uporabil različna programska orodja kot npr. AntConc, Orange Canvas, Ora Casos, Ontogen. Naj pripravo in procesiranje podatkov predstavim po stopnjah:

1. stopnja: priprava poizvedb po posameznih mesecih (20 mesecev) v obliki .TXT datotek
 - a. uvoz 20 .TXT datotek v programsko orodje AntConc 3.2.1w
 - b. ustvaril sem seznam nepotrebnih besed z namenom, da bi izločil manj pomembne besede kot npr. veznike, predloge in druge besede, ki se sicer zelo pogosto pojavljajo, vendar je njihov pomen manj zgovoren (npr. MNZ).
 - c. ustvarjanje seznama besed, kamor sem pripel še seznam nepotrebnih besed.
 - d. določitev udarnih ključnih besed s pomočjo dejavnika najpomembnejših ključnih besed (*angl.: keyness*). Kot prag vrednosti sem določil povprečno vrednost, medtem ko sem kot algoritem uporabil najugodnejšega soseda (*angl.: log-likelihood*). Znotraj AntConc Keyness opcije sem tudi določil referenčni korpus. Gre za zbirko vseh 58329 poizvedb. Primerjal sem

med sabo 20 datotek s celotnim številom poizvedb. Kot izid sem dobil 828 najpomembnejših ključnih besed (v nadaljevanju: udarne ključne besede) vključno z njihovimi frekvencami in dejavnikom Keyness.

e. s pomočjo AntConc sem tudi določil besedne grozde (*angl.: word clusters*). Te sem v nadaljevanju najprej uporabil za klasifikacijo udarnih ključnih besed in kasneje tudi za ontologije.

AntConc 3.2.1w (Windows) 2007

File Global Settings Tool Preferences About

Corpus Files

AprilMaj_2010.txt
AvgustSeptember_...
AvgustSeptember_...
December2009Janu...
December2010Janu...
FebruarMarec_201...
JanuarFebruar_20...
JulijAvgust_2009...
JulijAvgust_2010...
JunijJulij_2009.1...
JunijJulij_2010.1...
MajJunij_2009.txt...
MajJunij_2010.txt...
MarecApril_2010.1...
NovemberDecember...
NovemberDecember...
OktoberNovember...
OktoberNovember...
SeptemberOktober...
SeptemberOktober...

Concordance **Concordance Plot** **File View** **Clusters** **Collocates** **Word List** **Keyword List**

Hits Keyword Types Before Cut: 18203 Keyword Types After Cut: 828

Rank	Freq	Keyness	Keyword
1	2577	300.823	prijava
2	3386	241.619	dovoljenje
3	1696	174.699	prebivališča
4	1407	145.367	varovanje
5	2054	140.699	dovoljenja
6	1117	140.621	register
7	1319	136.804	potni
8	736	133.656	prireditve
9	1362	133.431	obrazec
10	1843	127.114	list
11	1050	126.119	izkaznica
12	1346	118.372	vozniškega
13	1358	114.430	stalnega
14	1035	110.345	osebna

AntConc 3.2.1w (Windows) 2007

File Global Settings Tool Preferences About

Corpus Files

AprilMaj_2010.txt
AvgustSeptember_...
AvgustSeptember_...
December2009Janu...
December2010Janu...
FebruarMarec_201...
JanuarFebruar_20...
JulijAvgust_2009...
JulijAvgust_2010...
JunijJulij_2009.1...
JunijJulij_2010.1...
MajJunij_2009.txt...
MajJunij_2010.txt...
MarecApril_2010.1...
NovemberDecember...
NovemberDecember...
OktoberNovember...
OktoberNovember...
SeptemberOktober...
SeptemberOktober...

Concordance **Concordance Plot** **File View** **Clusters** **Collocates** **Word List** **Keyword List**

Total No. of Cluster Types: 4294 Total No. of Cluster Tokens: 10473

Rank	Freq	Cluster
1	390	prijava začasnega
2	257	prijava stalnega
3	188	prijava začasnega prebivališča
4	163	prijava prireditve
5	132	prijava začasnega bivališča
6	130	prijava stalnega prebivališča
7	122	prijava prebivališča
8	106	prijava javne
9	105	obrazec prijava
10	95	prijava bivališča
11	83	prijava javne prireditve
12	77	prijava stalnega bivališča
13	75	začasna prijava
14	68	prijava odjava

2.1.1 Slika 1: Analiza poizvedb s pomočjo AntConc

Slika 1 prikazuje analizo poizvedb s pomočjo AntConc. Na levi strani se nahajajo naložene datoteke in v sredini so določeni kazalci za udarne ključne besede in njihove besedne grozde.

Natančneje o tem v nadaljevanju tega dela

f. podatke sem izvozil v .TXT datoteko in nato v Excel.

2. stopnja: obdelava podatkov znotraj Excel

a. udarne ključne besede sem nato klasificiral s pomočjo lastno izdelane (v nadaljevanju: LK) in univerzalne decimalne klasifikacije (v nadaljevanju: UDK).

b. določil sem tudi predpostavljeni namen (v nadaljevanju: ND) določene poizvedbe z vidika uporabnika spletnih strani MNZ, pri čemer sem si prav tako pomagal z besednimi grozdi.

Podatki so tako bili pripravljene za nadaljnje analize in sinteze s pomočjo programskih orodij s sveta OZVP/B. V ta sklop lahko omenim različna odprtokodna programska orodja, ki sem jih lahko bolj ali manj koristno uporabil. V bistvu je šlo za izvedbo številnih preizkusov in za ustvarjanje predpostavljenih modelov.

Preden predstavim omenjeno snov naj še prikažem del pripravljenih podatkov v preglednični obliki.

2.1.2 Preglednica 1: Del klasificiranih udarnih ključnih besed

R	F	K	UKB	LK	UDK	ND
1	2577	300.823	prijava	3	35	Viri
2	3386	241.619	dovoljenje	3	35	Viri
3	1696	174.699	prebivališča	4	35	Posebne informacije
4	1407	145.367	varovanje	6	35	Faktografsko znanje
5	2054	140.699	dovoljenja	3	35	Viri
6	1117	140.621	register	3	35	Viri
7	1319	136.804	potni	3	35	Viri
8	736	133.656	prireditve	6	35	Posebne informacije
9	1362	133.431	obrazec	3	35	Viri
10	1843	127.114	list	3	35	Viri
11	1050	126.119	izkaznica	3	35	Viri
12	1346	118.372	voznškega	3	35	Faktografsko znanje

Preglednica 1 prikazuje del klasificiranih (LK, UDK) udarnih ključnih besed. Povrhu tega prikazuje range (v nadaljevanju: R), frekvence (v nadaljevanju: F), dejavnik Keynes (v nadaljevanju: K) in predpostavljene namene (v nadaljevanju: ND) spletnih obiskovalcev. V

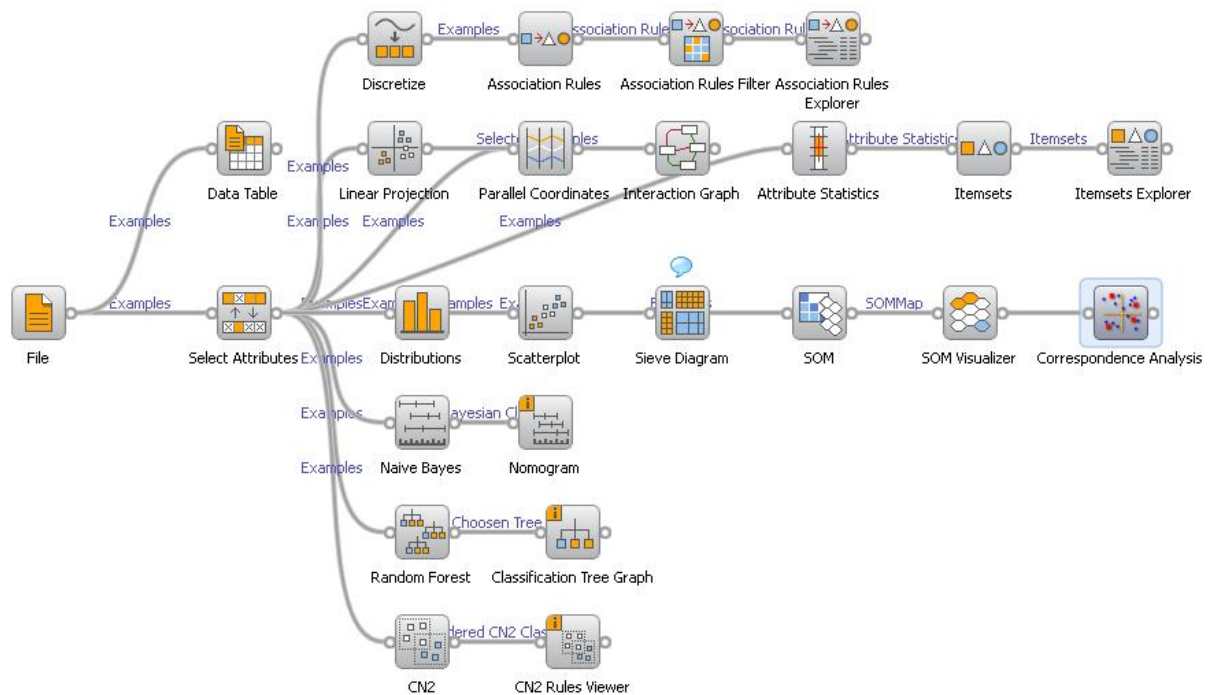
nadaljevanje sem z različnimi variacijami procesiral podatke skozi različna programska orodja.

2.2 Procesiranje podatkov skozi različna programska orodja

Podatke sem za namene OZVP/B procesiral skozi različna programska orodja od npr. Orange Canvas, Ora Casos, TMeV, Cytoscape, Visone idr.

2.3 Orange Canvas

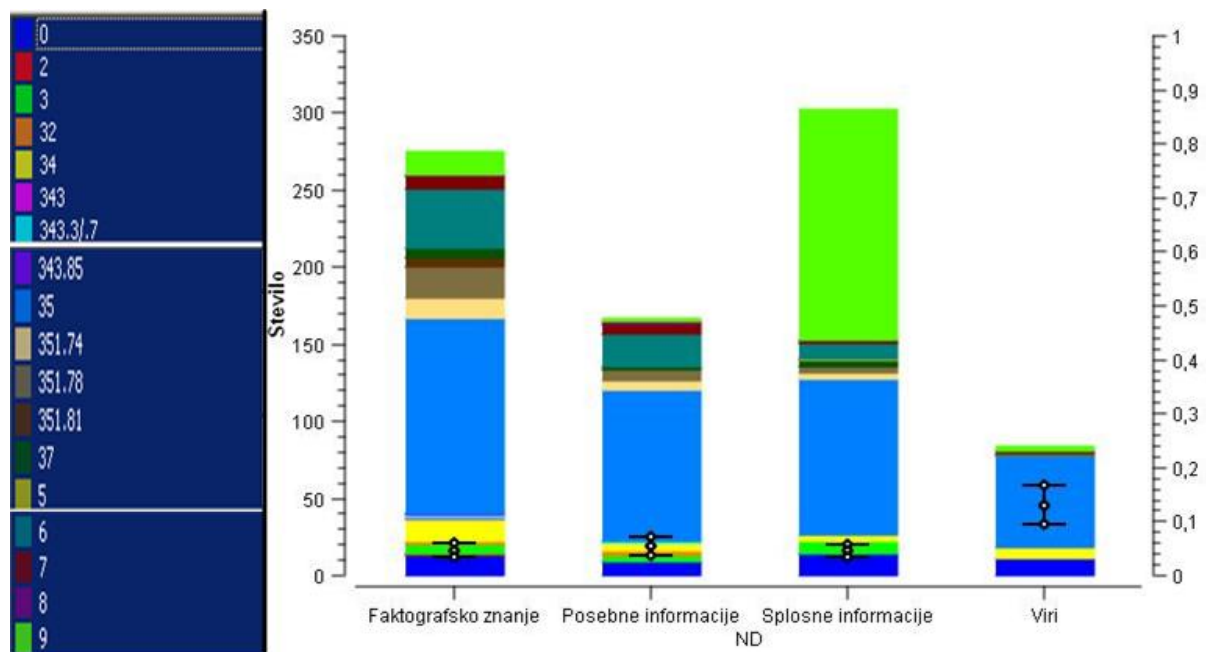
Procesiral sem vse kazalce skozi Orange Canvas, nakar sem vizualno programiral analizo podatkov.



2.3.1 Slika 2 : Nadzorna plošča programskega orodja Orange Canvas

Slika 2 prikazuje nadzorno ploščo programskega orodja Orange Canvas, s katerim sem vizualno programiral način odkrivanja zakonitosti v podatkih (določitev algoritma, vizualizacijske tehnike idr.). Po določevanju lastnosti podatkov (kazalci, razredi, meta-razredi ipd..) sem jih s pomočjo različnih vizualizacijskih tehnik (npr. distribucija, raztreseni diagram, SOM, linearna projekcija, Sieve diagram, statistika atributov) analiziral oziroma interpretiral.

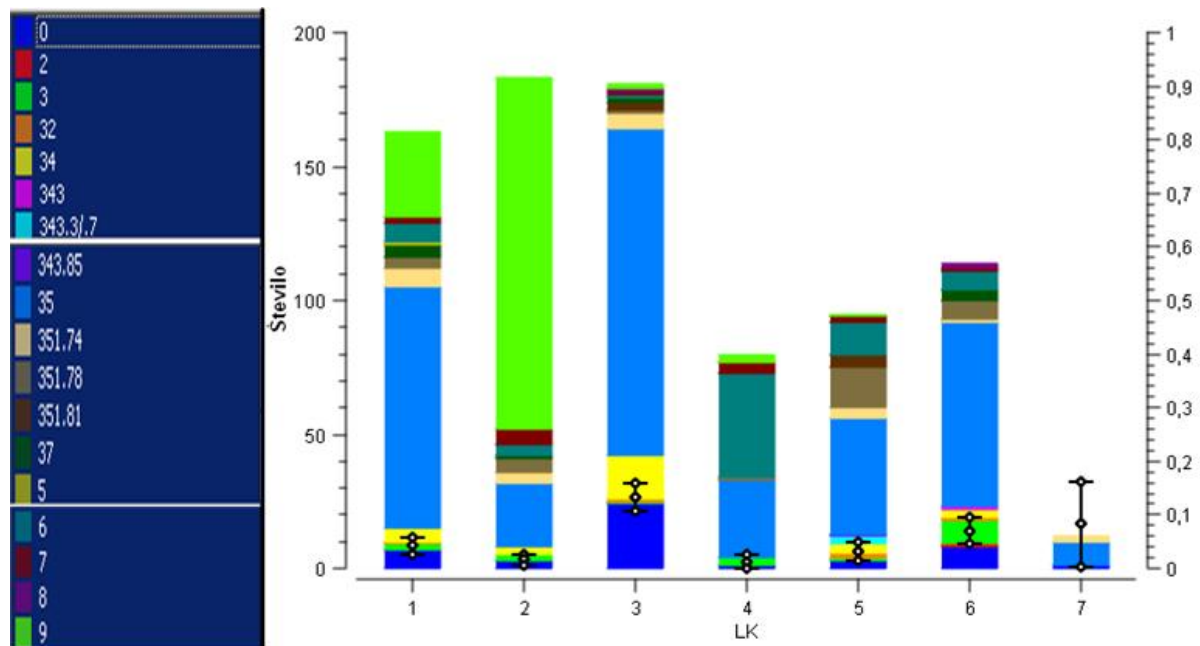
2.4 Različne oblike distribucij



2.4.1 Slika 3: Distribucija ND glede na frekvenco in UDK razredov

Slika 3 predstavlja distribucijo ND glede na frekvenco in UDK razredov (gl. obarvane stolpčne diagrame) udarnih ključnih besed. ND obiskovalcev spletnih strani MNZ je bil najpogosteje spodbujen z informacijsko potrebo po splošnih informacijah (npr. kontaktne informacije, uradne ure, oddelki, osebe), sledi informacijska potreba po pridobivanju faktografskega znanja (npr. kriminaliteta, kriminalistika, pravo, Schengenske meje, migracije), po posebnih informacijah (npr. izdaja, prebivališča, tiralice) in nenazadnje po informacijskih virih (npr. knjige, revije, zbirke podatkov, digitalne knjižnice, CD-ROM, sezname). Z vidika UDK ugotavljam, da je največje število udarnih ključnih besed klasificiranih v skupino UDK 35 (javna uprava). Ta ugotovitev velja za vse ND razen za splošne informacije. Znotraj te kategorije se nahaja največje število udarnih ključnih besed s področja UDK 9 (geografske lokacije, osebe, zgodovinski dogodki idr.). Po preučevanju besednih grozdov sem lahko razbral, da so te tvorbe mnogokrat povezane z osebami, ki službujejo znotraj javne uprave. Sorazmerno visok delež udarnih ključnih besed lahko opazimo s področja UDK 6 (zdravje, tehnologija, tehnika idr.) in UDK 351.78 (varnost, varovanje idr.), še zlasti pri ND faktografsko znanje. Zastopanost udarnih ključnih besed s področja UDK 0 (npr. knjige, zbirke podatkov, sezname, revije) je pri vseh ND kategorijah sorazmerno podobna. Kot zanimivost velja še omeniti sorazmerno visok delež udarnih

ključnih besed na področju UDK 34 (pravo) in UDK 351.74 (Policija) z ozirom na ND faktografsko znanje (npr. zakonodaja, zakoni, policaj, pridržanje idr.).



2.4.2 Slika 4: Distribucija LK glede na frekvenco in UDK razrede

Slika 4 prikazuje distribucijo LK glede na frekvence in UDK razrede. Največje število udarnih ključnih besed najdemo pri skupini LK 2 (osebe), LK 3 (informacijski viri) in LK 1 (organizirane združbe, oddelki ipd.). Pomen LK sledi v nadaljevanju:

LK 1: v to skupino sem klasificiral razne sociološke sisteme in njihove dele (npr. MNZ, Specialna knjižnica),

LK 2: gre za osebe (npr. tujci, prebivalci, Karl Petrič),

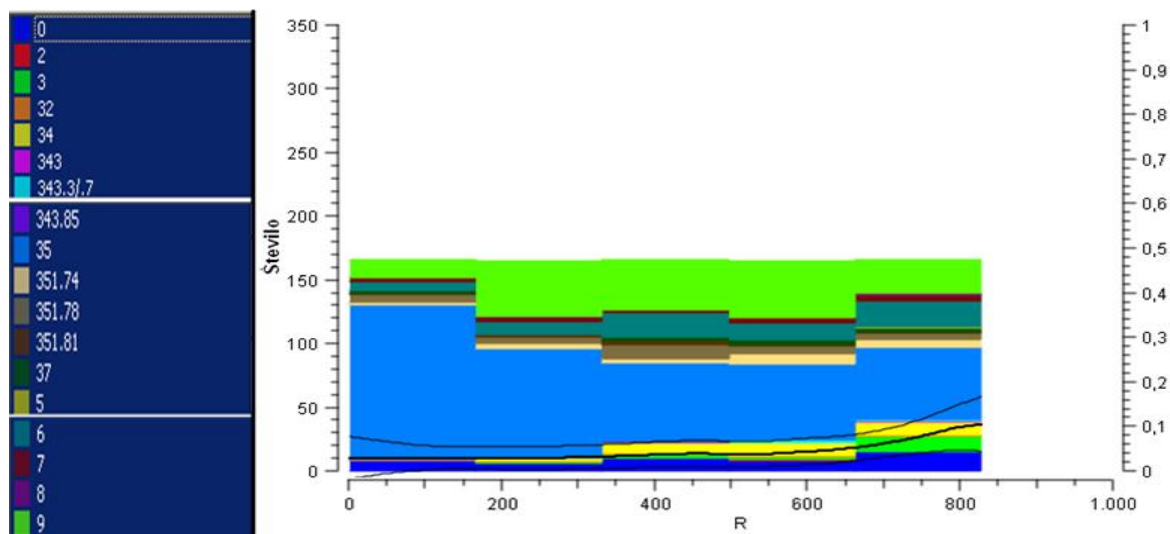
LK 3: gre za intelektualne kulturne stvaritve (npr. knjige, dokumenti, tehnični sistemi, obrazci),

LK 4: v to skupino sem umestil predmete / materiale idr. (npr. stanovanja, cena osebne izkaznice),

LK 5: v to skupino sem klasificiral kulturna področja v najširšem pomenu te besede (npr. znanost, umetnost, šport),

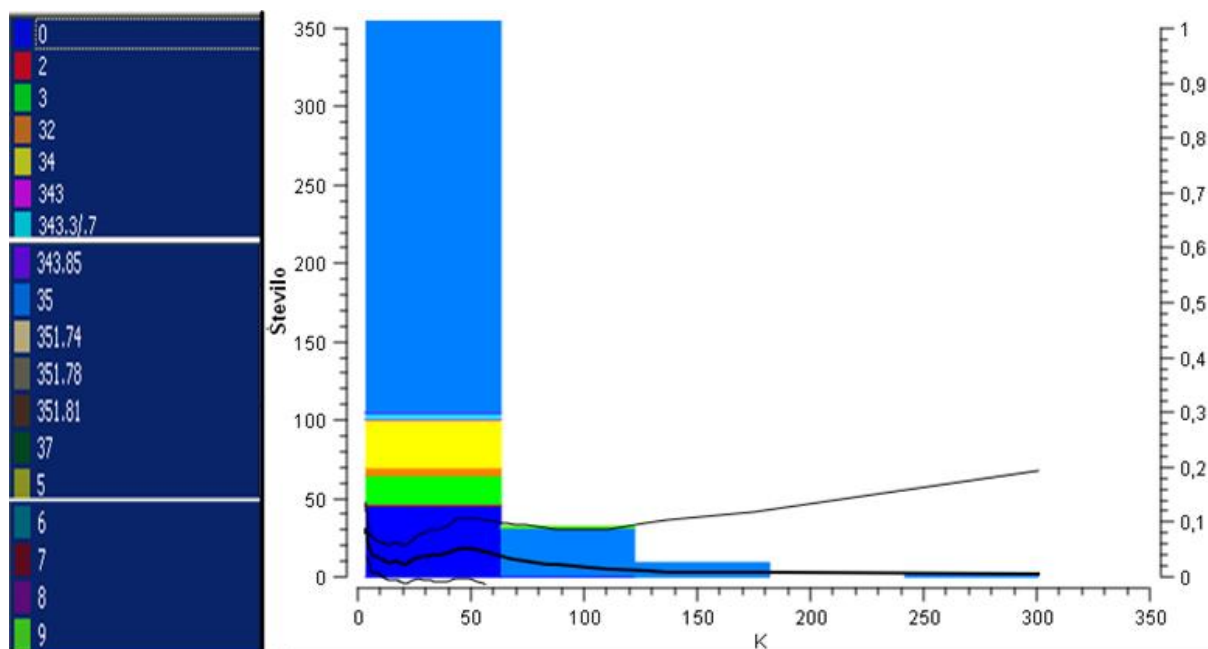
LK 6: gre za procese, dejavnosti, stanja in dogodke (npr. komuniciranje, e-anonimna prijava, proslava),

LK 7: in v to skupino sem razvrstil vprašanja, ki so jih v preteklosti zastavili spletni obiskovalci (npr. kaj potrebuje tujec za prijavo bivanja?).



2.4.3 Slika 5: Distribucija ranga glede na frekvenco in UDK razrede

Slika 5 prikazuje distribucijo ranga glede na frekvenco in UDK razrede. Najvišje range dosega udarne ključne besede s področja UDK 35. Sledijo jim besede s področja UDK 9, UDK 6, UDK 351.78, UDK 351.74, UDK 0, UDK 34 itd.

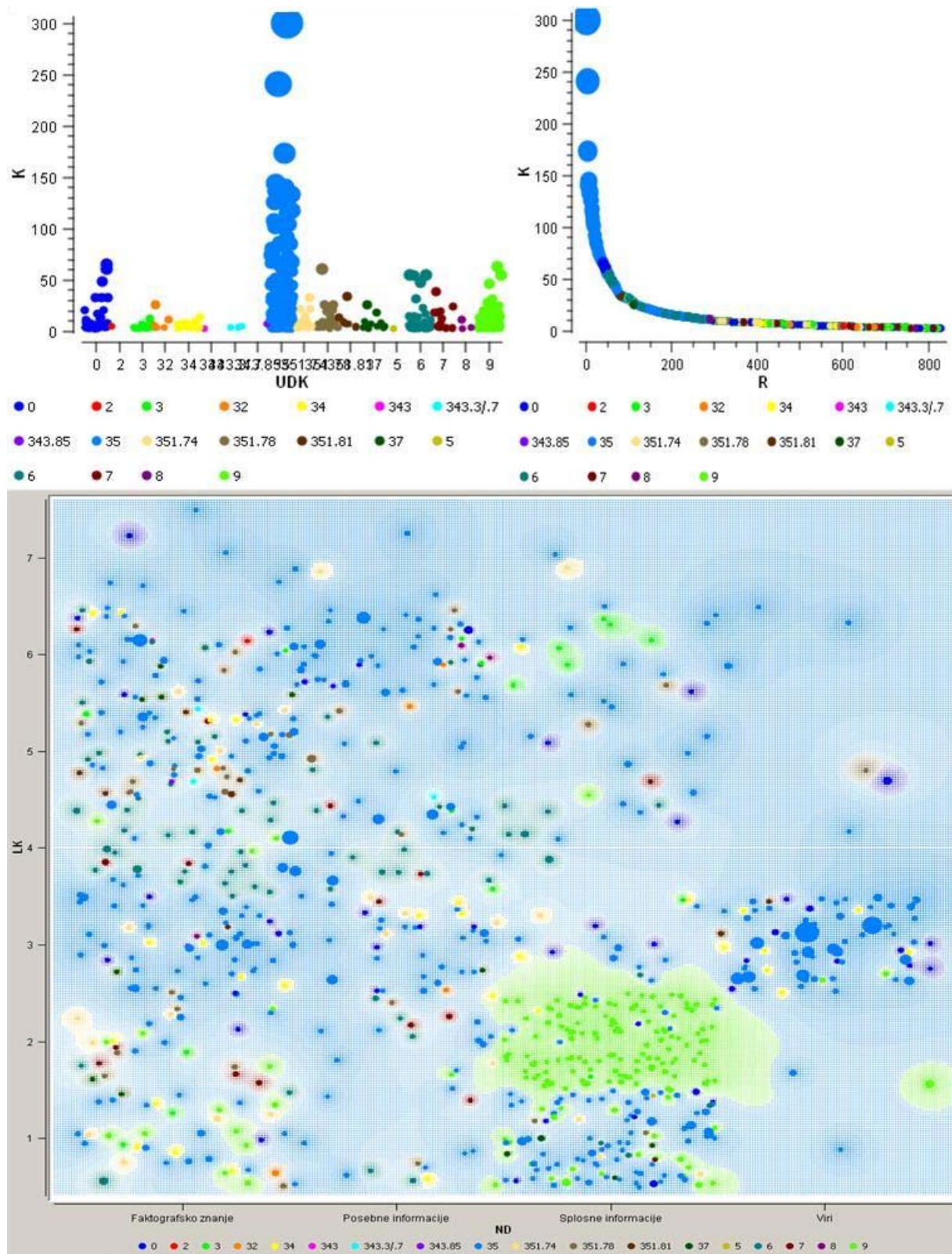


2.4.4 Slika 6: Distribucija dejavnika K glede na frekvenco in UDK razrede

Slika 6 prikazuje s pomočjo štirih stolpcev distribucijo dejavnika K glede na frekvenco in UDK razrede. Ugotavljam (gl. stolpec 1), da imamo največ udarnih ključnih besed z nižjimi UDK razrede.

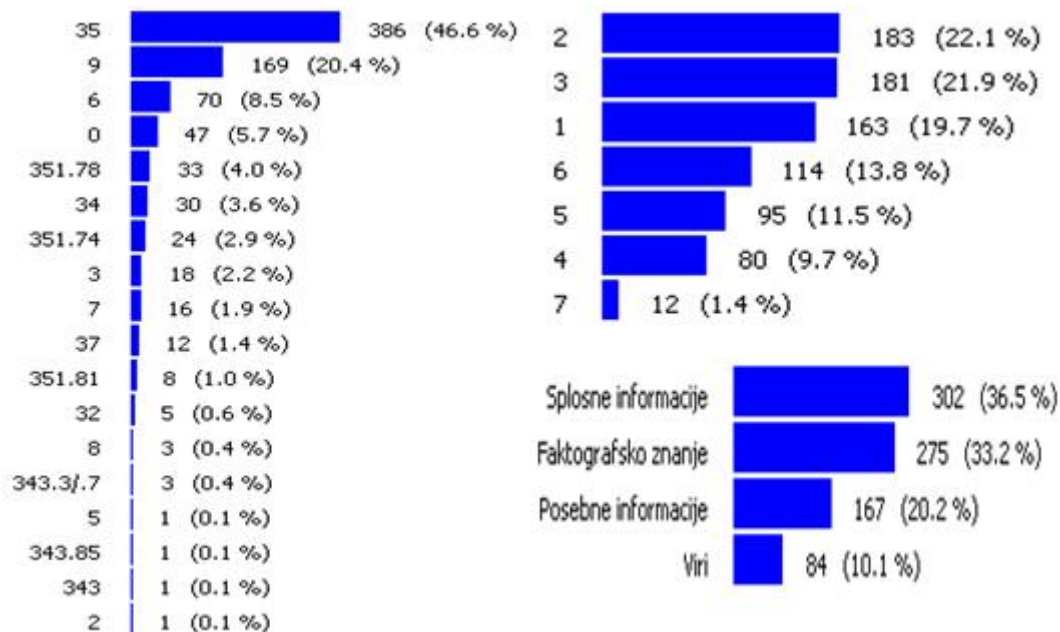
dejavniki K s področja UDK 35, nato UDK 343.85 (preprečevanje kriminalitete), UDK 343.3/.7 (kriminaliteta, kazniva dejanja), UDK 343 (kazensko pravo), UDK 34 (pravo), UDK 32 (Politika), UDK 3 (družbene vede, sociologija ipd.), UDK 2 (verstva, religija ipd.) in UDK 0 (razne knjige, dokumenti, sezname, revije, zbirke podatkov). V drugem stolpcu se nahajajo udarne ključne besede z nekoliko višjimi dejavniki K s področja UDK 9, UDK 35 in UDK 0. V tretjem stolpcu se nahajajo udarne ključne besede s še višjimi dejavniki K, ki so vse s področja UDK 35. V četrtem stolpcu se nahajata dve udarni ključni besedi (prijava – K = 300.823 in dovoljenje – K = 241.619). Obe udarni ključni besedi sem razvrstil v skupino UDK 35.

2.5 Rztreseni diagram



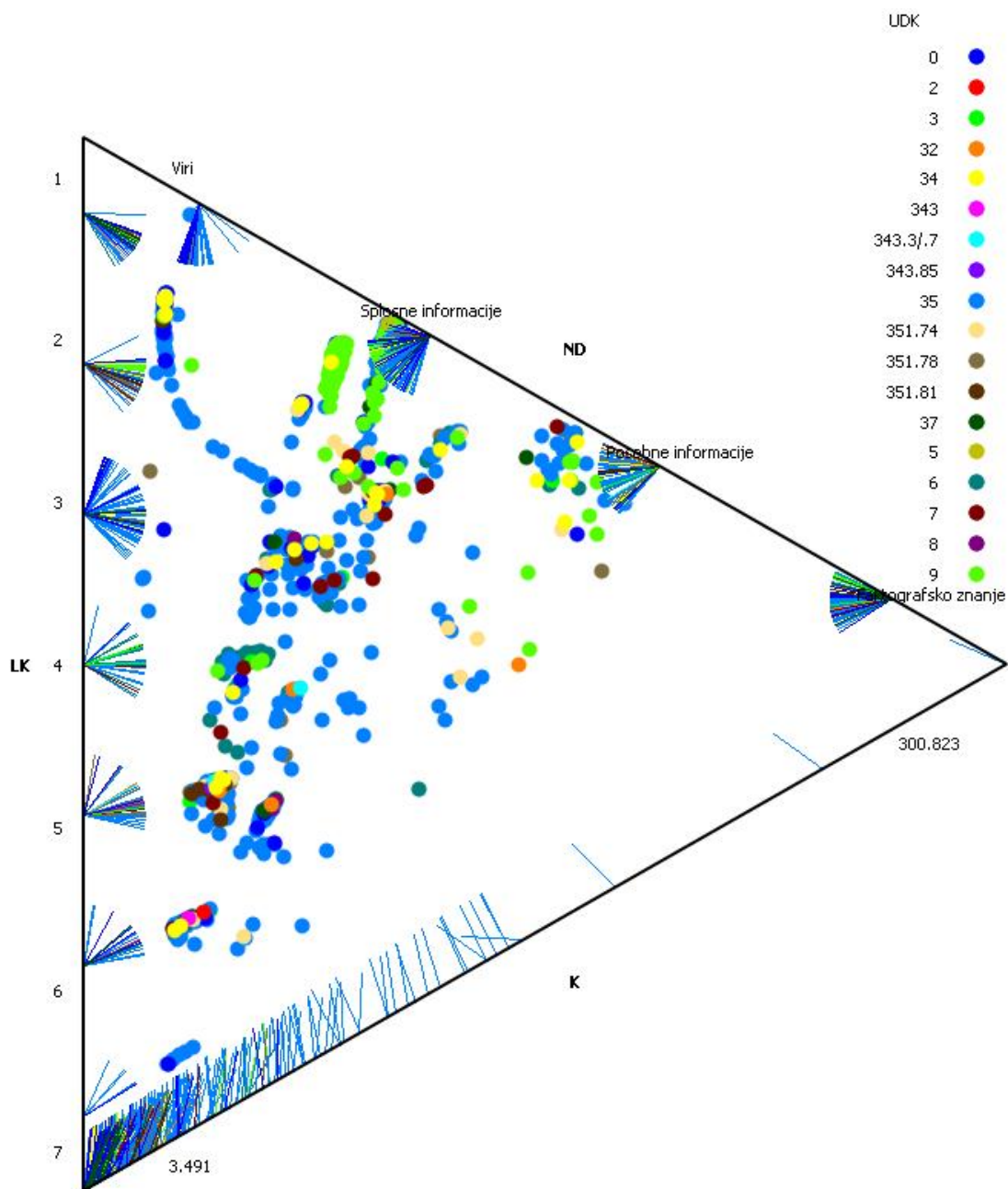
2.5.1 Slika 7: Del matrike rztresenih diagramov

Slika 7 prikazuje del matrike raztrresenih diagramov. Zgornji levi diagram ponazarja primerjavo med UDK razredi (gl. X-os) in dejavnikom K (gl. Y-os). Največ predstavnikov z visokim dejavnikom K opazimo na področju UDK 35 (prijava, dovoljenje, prebivališča, varovanje, register, dovoljenja, list, izkaznica idr.), kar še bolj nazorno prikazuje Zipfova krivulja (gl. desno stran slike 7). Spodnji del slike 7 prikazuje razmerje med ND (X-os) in LK (Y-os) s pomočjo barvne upodobitve UDK razredov ter velikostne predstavitve podatkov v obliki dejavnika K. Najbolj številčni, vendar tudi ne raznovrstni (z vidika UDK področja) in gosti del spodnjega raztrresenega diagrama se nahaja na poziciji splošnih informacij. Ta je najbolj gost pri LK 1 in LK2 (npr. pod LK 2 se je uvrstilo veliko število udarnih ključnih besed, ki so v povezavi s poizvedbami o osebah). Najbolj raznovrstne in gosto porazdeljene udarne ključne besede je možno najti pri ND faktografsko znanje (npr. varovanje, vozniškega, javna, prebivališče, državljanstvo, zakon). Sorazmerno enakomerno porazdeljene so udarne ključne besede za ND posebne informacije (npr. prebivališča, prireditve, stalnega, začasno, začasnega, cena, sprememba, vozniško). Najmočnejše udarne ključne besede je možno najti na poziciji ND viri (npr. prijava, dovoljenje, dovoljenja, register, potni, obrazec, list, izkaznica, osebna, odjava), še zlasti na področju LK 3. Sicer je predstavnikov drugih razvrščenih udarnih ključnih besed po LK za ND viri precej manj. V vseh skupinah prevladuje UDK 35 razen v tej, kjer prevladuje UDK 9.



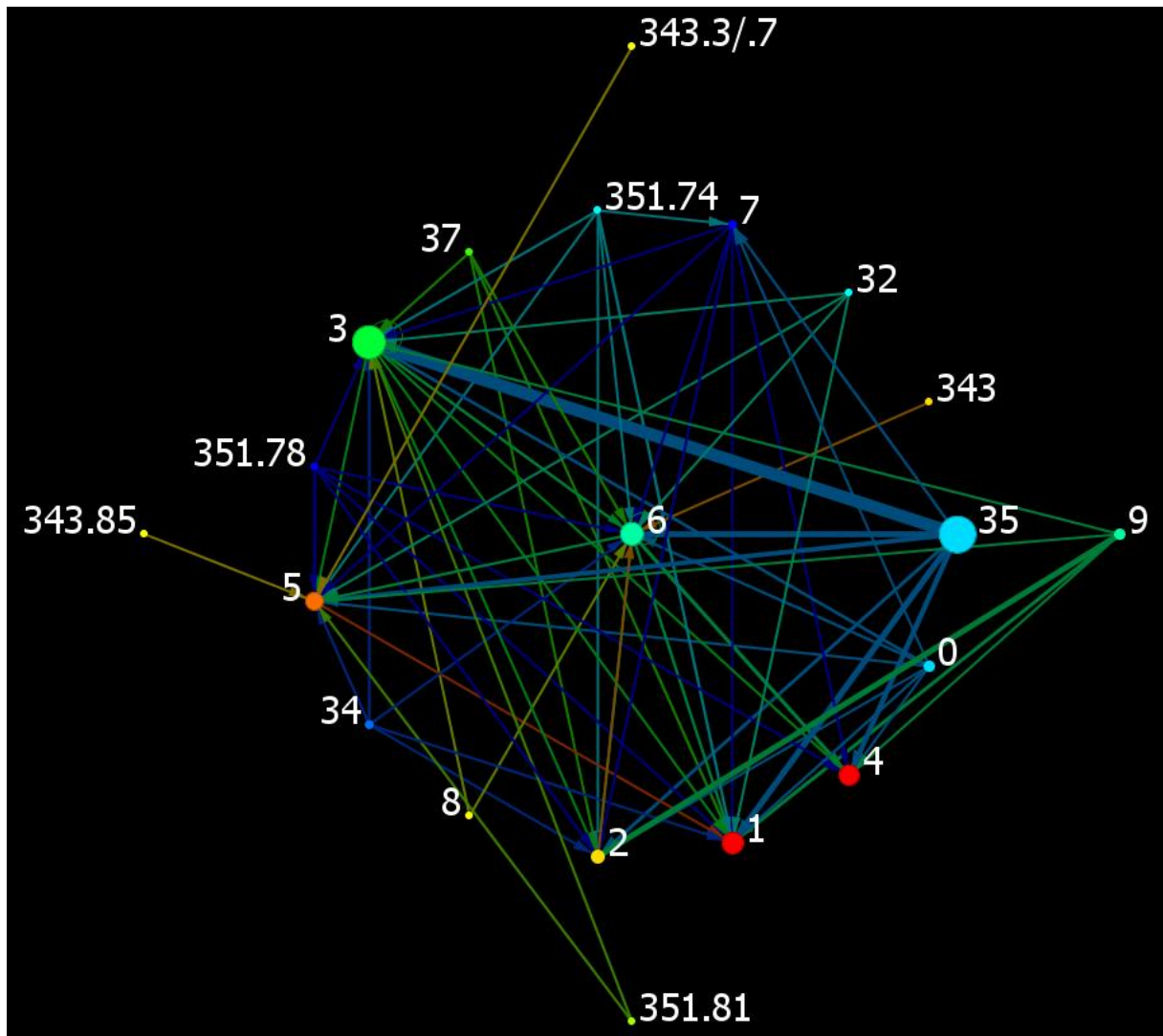
2.5.2 Slika 8: Statistični atributi za UDK, LK in ND

Slika 8 predstavlja statistične atribute za UDK, LK in ND v obliki frekvenc in odstotkovnih deležev. Največji odstotek zavzemajo udarne ključne besede s področja UDK 35, UDK 9, UDK 6, UDK 0, UDK 351.78, UDK 34, UDK 351.74, UDK 7, UDK 37, UDK 351.81, UDK 32, UDK 8, UDK 343.3/.7, UDK 5, UDK 343.85, UDK 343 in UDK 2. Z vidika LK največji odstotkovni delež zajema LK 2, LK 3, LK 1, LK 6, LK 5, LK 4 in LK 7. Desna spodnja stran slike 8 prikazuje odstotkovne deleže za splošne informacije, faktografsko znanje, posebne informacije in vire.



2.5.3 Slika 9: Polivizualizacija K, LK in ND v odnosu z UDK

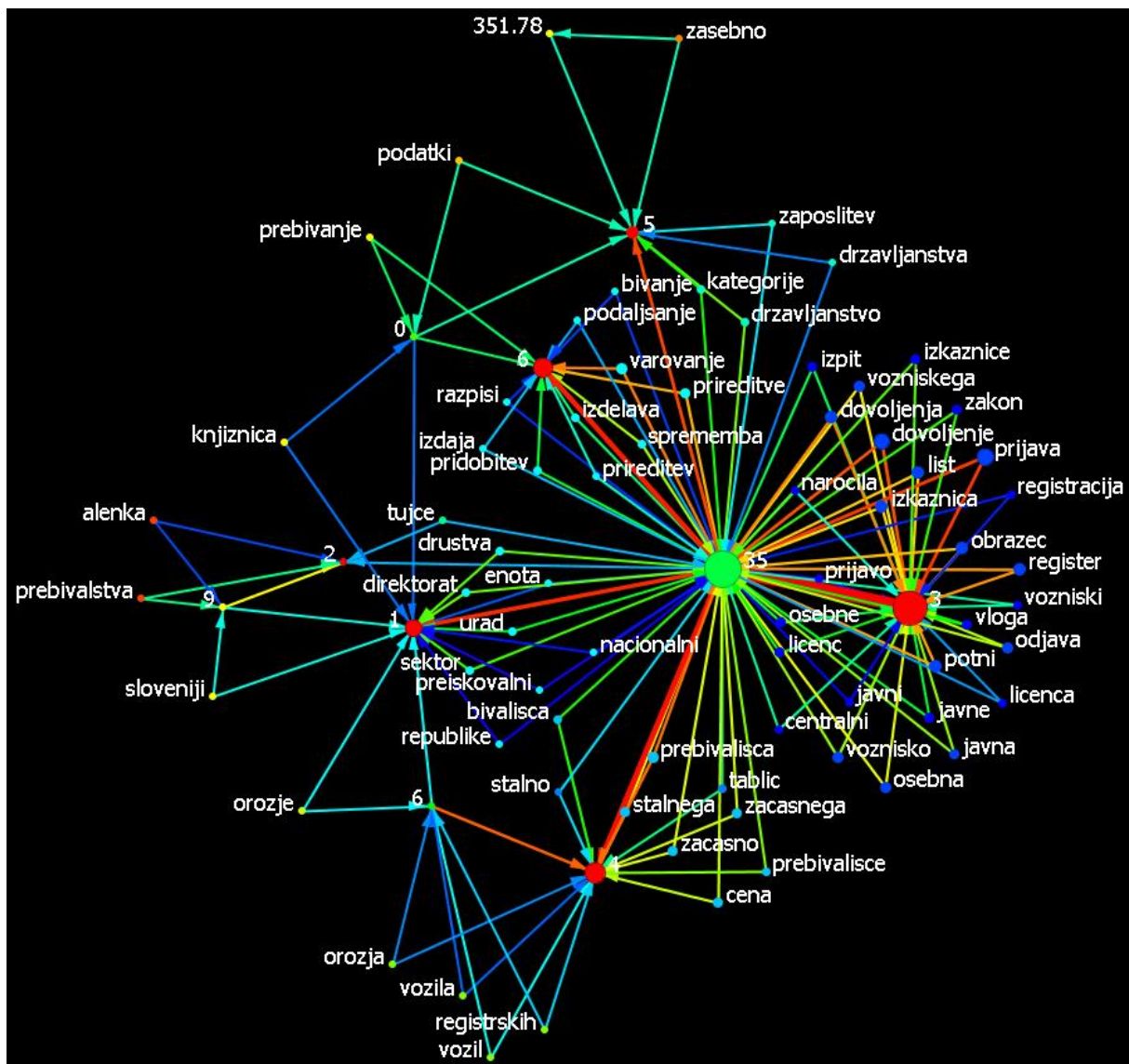
Slika 9 prikazuje kazalce K, LK, in ND v odnosu z UDK s pomočjo vizualizacijske tehnike Poliviz.



2.5.4 Slika 10: Krožni mrežni diagram za UDK in LK področja z dejavnikom K

Slika 10 prikazuje krožni mrežni diagram za najbolj udarne ključne besede, ki so razvrščene znotraj posameznih UDK področij. Pri tej analizi sem uporabil algoritem korelacijske podobnosti (za prikazovanje bližine med UDK področij) in eigenvektorjevo centralnost (za prikazovanje pomembnosti določenega UDK vozlišča znotraj omrežja). Ob tem sem izločal vsa tista vozlišča, katerih dejavnik K je bil nižji ali kvečjemu enak 40,1 ($K \leq 40,1$). Opazimo lahko, da obstaja najmočnejša povezava med UDK 35 in 3. Obe vozlišči (gl. sliko 10) sta tudi najbolj pomembni, kar je na sliki 10 upodobljeno z večjim modrim in zelenim krogom. V

sredini krožnega mrežnega diagrama se nahaja vozlišče UDK 6, iz katerega izhaja največje število povezav (ob tem naj opozorim na močnejšo povezavo med UDK 35 in 6). Obstaja tudi nekoliko močnejša povezava med LK 1, LK4 (gl. rdeče obarvana kroga na sliki 10) in UDK 35.



2.5.5 Slika 11: Mrežni diagram za UDK, LK in UKB z najvišjimi dejavniki K

Slika 11 prikazuje mrežni diagram za UDK, LK in UKB z najvišjimi dejavniki K ($K \geq 40,1$). UKB s $K \geq 40,1$ so naslednje:

$K > 100$: prijava ($K=300.823$), dovoljenje ($K=241.619$), prebivališča ($K=174.699$), varovanje ($K=145.367$), dovoljenja ($K=140.699$), register ($K=140.621$), potni ($K=136.804$), prireditve

(K=133.656), obrazec (K=133.431), list (K=127.114), izkaznica (K=126.119), vozniškega (K=118.372), stalnega (K=114.430), osebna (K=110.345), začasno (K=108.222), začasnega (K=104.269), cena (K=101.776) in odjava (K=101.517).

K < 100 (od K=92.577 do K=40.888): sprememba, vozniško, javna, prebivališče, državljanstvo, direktorat, društva, izkaznice, sektor, zakon, bivališča, kategorije, vloga, licenc, urad, javne, pridobitev, osebne, izpit, prebivanje, izdelava, tablic, prebivalstva, centralni, vozniški, podatki, zasebno, prireditev, naročila, sloveniji, vozil, orožje, zaposlitev, stalno, registrskih, izdaja, tujce, podaljšanje, licenca, orožja, državljanstva, knjižnica, vozila, enota, alenka, bivanje, razpisi, registracija, nacionalni, javni, republike, preiskovalni in prijavo. Ekstrahirani mrežni diagram v bistvu pokaže na najbolj povezana (algoritem korelacijske podobnosti) in pomembna (algoritem eigenvektorjeve centralnosti) UDK ter LK področja. Povrhu tega omenjena slika 11 prikazuje najmočnejše UKB z njihovimi povezavami do UDK in LK. Na osnovi značilnosti prikazanega omrežja je možno izpeljati naslednja spoznanja:

a. največje število UKB je s področja UDK 35 in LK 3 (med UDK 35 in LK 3 je tudi najmočnejša povezava). Obe vozlišči sta znotraj tega omrežja najpomembnejši. V tem pogledu gre še zlasti za UKB, ki opredeljujejo administrativne intelektualne izdelke kot so npr. potni list, osebna izkaznica, prijavnica, licenca, vozniško dovoljenje, vozniški izpit, centralni register, prijava prebivališča, odjava prebivališča, zakon ipd. Gre v bistvu pretežno za dokumente in zbirke podatkov, ki se uporabljajo pri poslovanju javne uprave.

b. sledi druga najmočnejša povezava med UDK 35 in LK 4 (npr. prebivališče, cena)

c. na tretjem mestu se nahaja povezava med UDK 35 in LK 6 (npr. varovanje, podaljšanje, bivanje, prireditev)

d. med močnejšo povezavo bi še omenil povezavo med UDK 35 in LK 1 (npr. društva, direktorat, knjižnica).

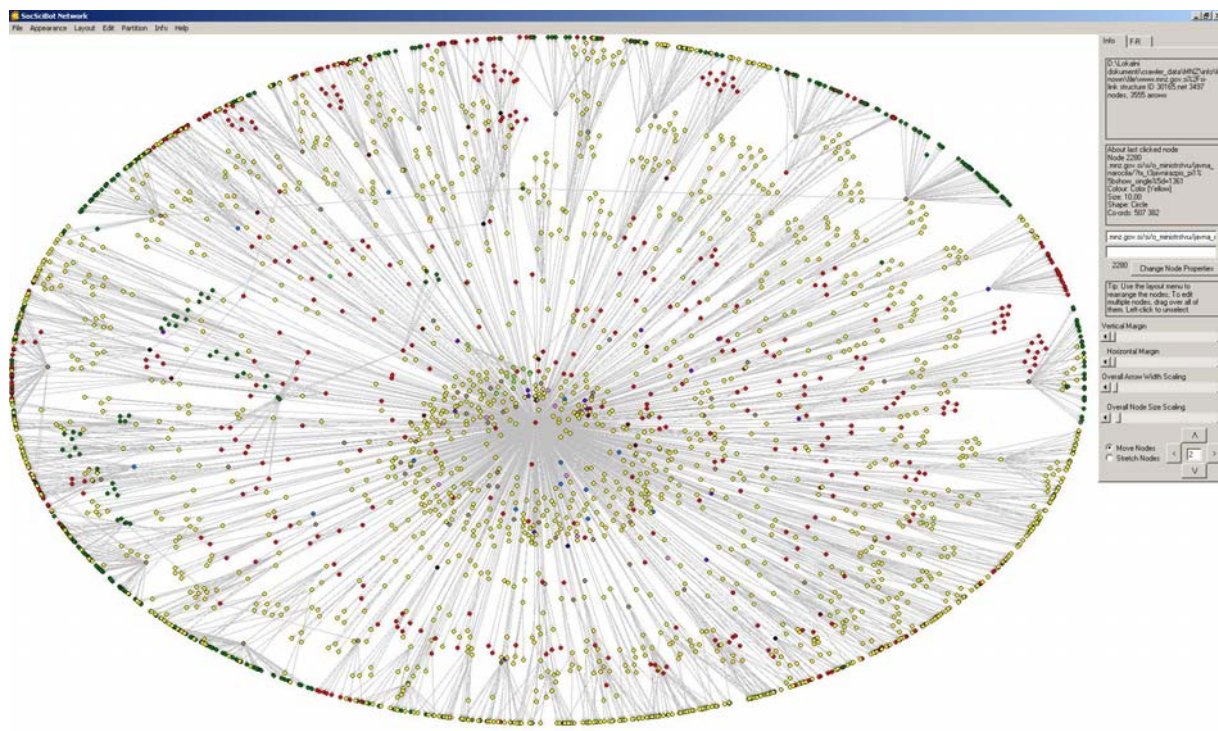
e. povezave med UDK 35 in LK 5 ter LK 2 sta najšibkejši (zasebno, alenka, podatki)

V nadaljevanju bom te ugotovitve uporabil pri analizi najpomembnejših besednih grozdov, pri čemer si bom pomagal z ustreznimi programskimi orodji.

3 Analiza zgradbe in vsebine spletnih strani MNZ

V tem poglavju bom predstavil analizo zgradbe in vsebine spletnih strani MNZ s pomočjo prikaza omrežij in besed. Glavni namen teh prizadevanj je v tem, da je možno na podlagi tega rekonstruirati in deloma napovedovati dejavnosti uporabnikov spletnih strani MNZ. Povrhu tega s to analizo ugotovimo funkcionalne in nefunkcionalne vsebine. Pričujoča analiza je bila izvedena na osnovi naslednjih dejavnosti:

- a. preiskovanje vseh spletnih strani in zunanjih povezav MNZ s pomočjo spletnega robota ali pajka SocSciBot 4 (čas pajkanja okoli pet ur; zajel okoli 15000 spletnih strani MNZ; od vseh 2812526 besed je program izbral 62575 izvirnih),
- b. upodobitev dobljenih izidov s pomočjo programskega orodja Pajek (upodobil sem omrežje, ki je vsebovalo povezave na strani in podstrani MNZ)



3.1 Slika 12: Omrežje spletnih strani MNZ po področjih delovanja

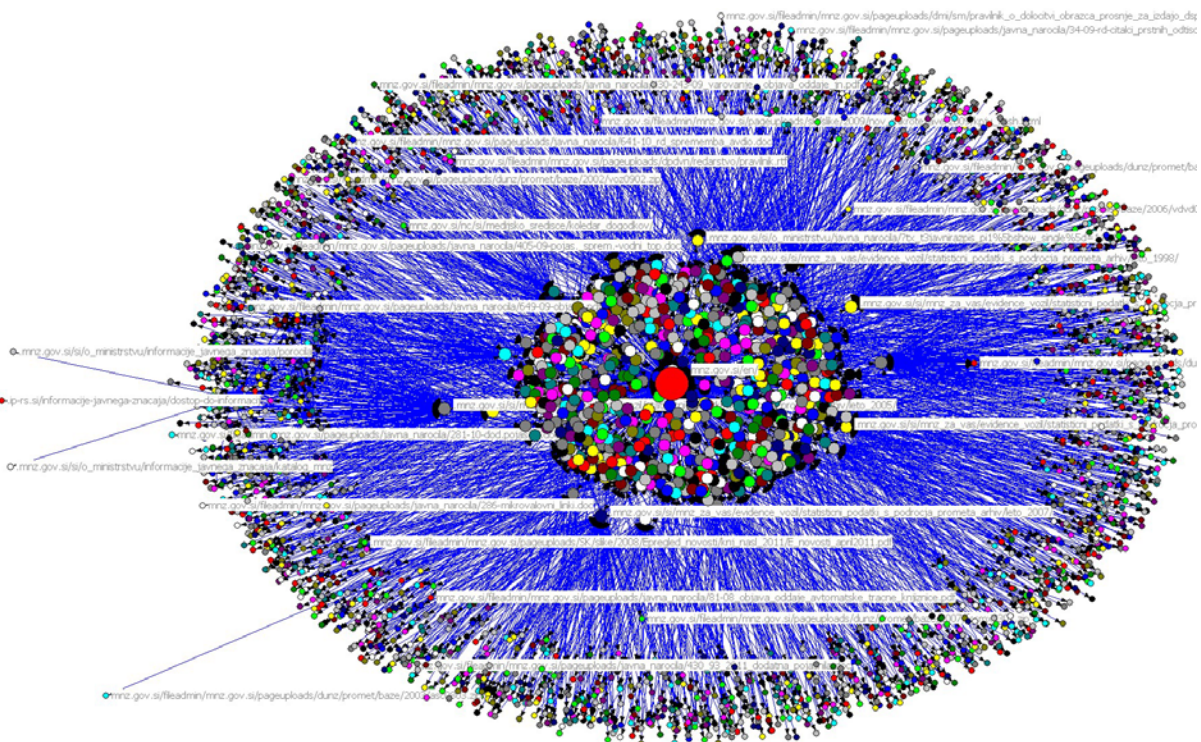
Slika 12 prikazuje omrežje spletnih strani MNZ po nekaterih področjih delovanja.

Presenečajo zelo številne spletne strani in podstrani s področja javnih naročil (označene z rumeno barvo, gl. sliko 12) in prometa vključujoče s statističnimi podatki (gl. zelene kroge). Na MNZ obstaja tudi veliko število spletnih strani s področja zakonodaje in dokumentov (gl. vijolično barvo), kar pa ne preseneča. Pogosti izrazi znotraj spletnih strani povečujejo verjetnost obiska uporabnika, ki uporablja iskalnike. Prav zaradi tega imajo nekatere spletne strani po izvedeni poizvedbi na določenem zunanjem iskalniku visok rang (npr. je med prvimi

100 zadetki). Poizvedba npr. mnz javna naročila je bila glede pogostosti na 73. mestu, a UKB "javna" s K=87.410 je dosegla 21. mesto. Po drugi strani vsebuje MNZ skupine spletnih strani in podstrani, ki niso zelo pogoste, vendar so dosegle njihove poizvedbe visoke frekvence. Ob tem je potrebno dodati, da so UKB npr. prijava dovoljenje, prebivališča, varovanje, dovoljenja, register, potni, prireditve, obrazec, list dosegle visoke dejavnike K (od 300.823 do 127.114). Visoka frekvenca določenih besed znotraj spletnih strani in podstrani torej še ne pomeni zagotovilo, da bodo te strani tudi najbolj obiskane! Na podlagi analize vsebine spletnih strani MNZ lahko pridemo do zelo pomembnega spoznanja:

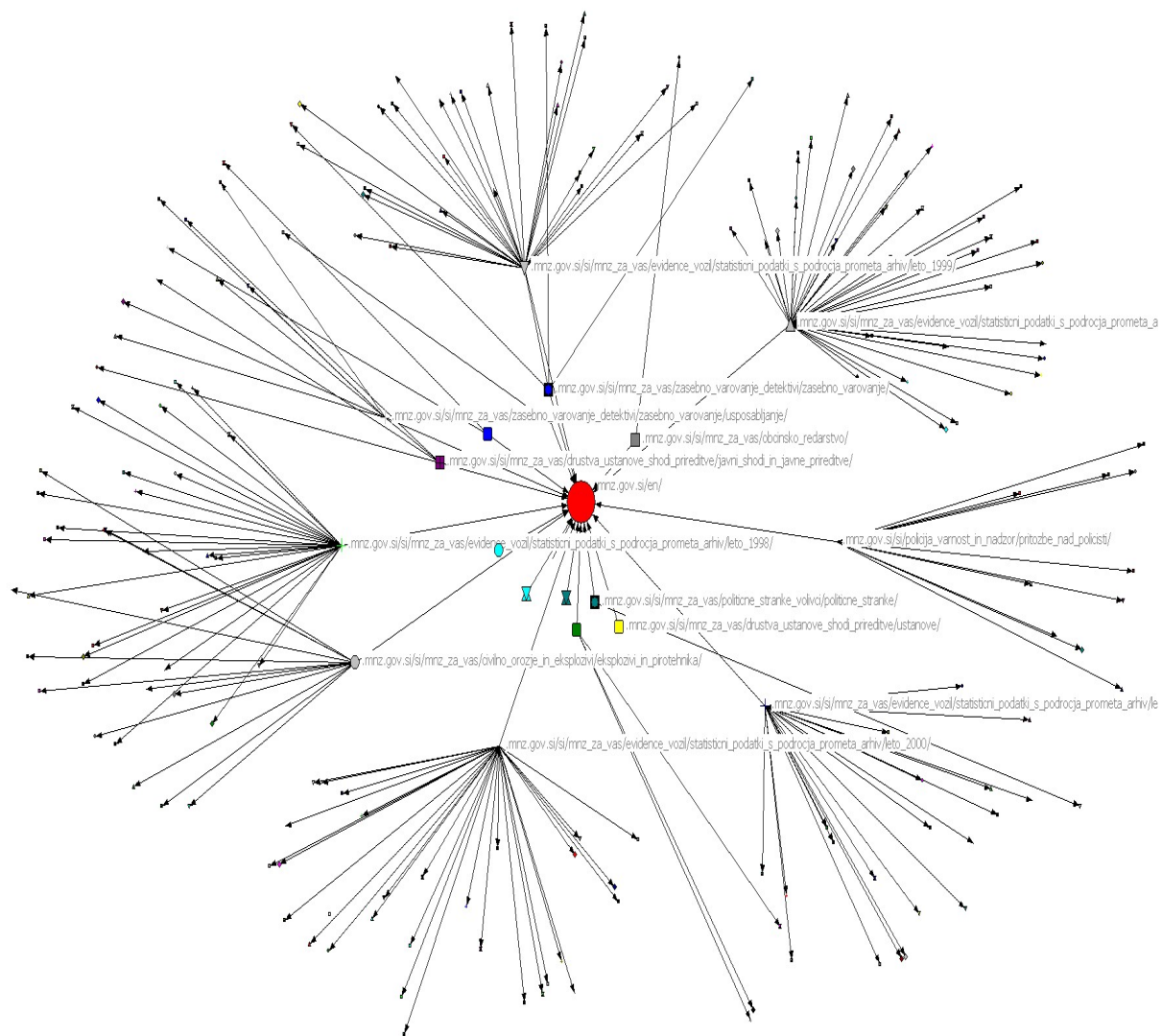
"Visoko število spletnih strani in podstrani o javnih naročilih, nam lahko daje vedeti, da so v bistvu javna naročila dokaj zapleteno področje, ki si zasluži, da bi bilo poenostavljeno in s tem racionalnejše ter koristnejše. V javnih naročilih je veliko vložkov od porabe človeške energije (npr. negativne stresne situacije), časa (npr. izvajanje zapletenih in zamudnih postopkov) pa tja do denarnih / materialnih sredstev (npr. velika poraba papirja). Skratka, javna naročila v ekstremni obliki povzročajo več stroškov kot pa koristi!" Na spletnih straneh MNZ se nahaja veliko število nedelujočih spletnih povezav. Verjetno je veliko število teh spletnih povezav nastala ob prehodu s starega na novi Typo3 sistem?

c. nadaljnja analitična obdelava s pomočjo programskega orodja NetDraw 2.111 (omrežje s 3583 vozlišči sem v nadaljevanju podvrgel algoritmu bližine in eigenvektorju).



3.2 Slika 13: Analiza omrežja MNZ s pomočjo NetDraw

Slika 13 prikazuje analizo omrežja MNZ s pomočjo NetDraw. Poleg uporabe algoritmov sem s pomočjo transformacije izločil bolj oddaljena in manj pomembna vozlišča. Pomembnejša vozlišča bom predstavil v nadaljevanju.



3.2.1 Slika 14: Analiza pomembnejših vozlišč s pomočjo NetDraw

Slika 14 prikazuje manj oddaljena in bolj pomembna vozlišča znotraj velikega omrežja, ki so naslednja: glavna stran MNZ, zasebno varovanje - evidence, vloge in obrazci, evidenca vozil, statistični podatki prometa, tujci v Sloveniji, dovoljenje za prebivanje, pritožbe nad policisti, občinsko redarstvo, prijava, prireditve, javni shodi, eksplozivi in pirotehnika. Navkljub velikemu številu spletnih strani v zvezi z javnimi naročili se le-te niso uvrstile v središče pomembnejših vozlišč. Spletne strani o javnih naročilih se nahajajo znotraj drugega spletnega zavihka, ki se imenuje O Ministrstvu, medtem ko se prej omenjene vsebine nahajajo znotraj prvega zavihka (MNZ za vas). Na osnovi te analize zelo učinkovito ugotovimo vsebinsko zgradbo spletnih strani MNZ.

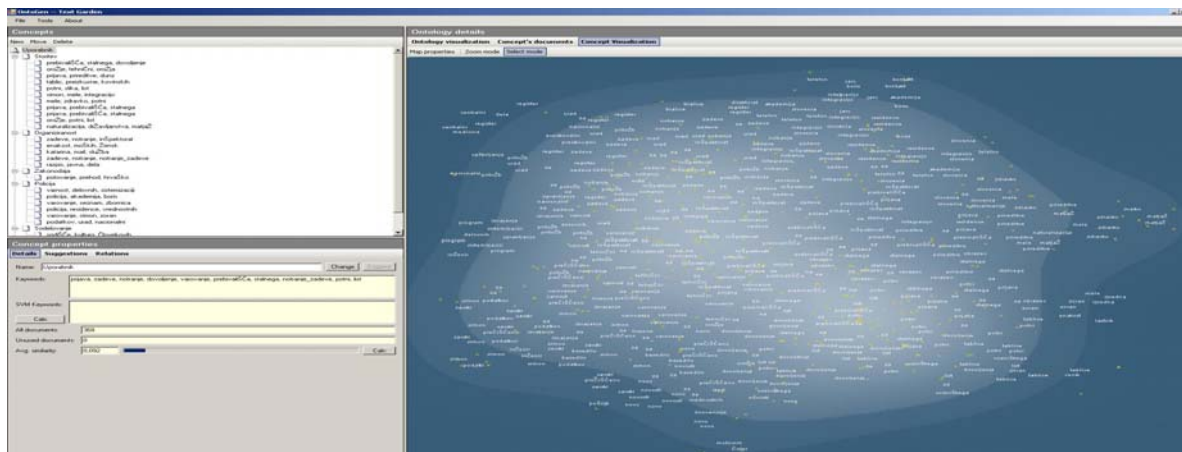
d. indeksiranje besed po frekvencah. Gre za besede, ki se bolj ali manj pogosto pojavljajo na spletnih straneh MNZ,

3.3 Preglednica 2: Del indeksiranih besed s frekvencami

Rang	Besede s spletnih strani MNZ	Število
1	Zadeve	34955
2	Notranje	34739
3	Predsednik	19065
4	Ministrstvo	13828
5	Policija	12576
6	MNZ	12110
7	Pravic	10887
8	Svet	10200
9	Vlada	10109
10	Informacijski	10021

Preglednica 2 prikazuje del indeksiranih besed s frekvencami, ki sem jih dobil s pomočjo spletnega pajka SciBot. Spoznal sem, da najbolj pogoste poizvedbe po zunanjih iskalnikih sovpadajo s tem indeksiranim seznamom (npr. MNZ, notranje, zadeve, ministrstvo).

d. izdelava ontologije z indeksiranimi besedami in UKB (UKB sem pridobil s pomočjo besedne analize številnih poizvedb) oziroma z besednimi grozdi. Besede, ki so bile mnogokrat predmet poizvedb, se dokaj pogosto pojavljajo znotraj besedil na spletnih straneh MNZ. Prav to dejstvo me je spodbudilo, da s pomočjo programskega orodja Ontogen 2.0 ustvarjam koncepte.

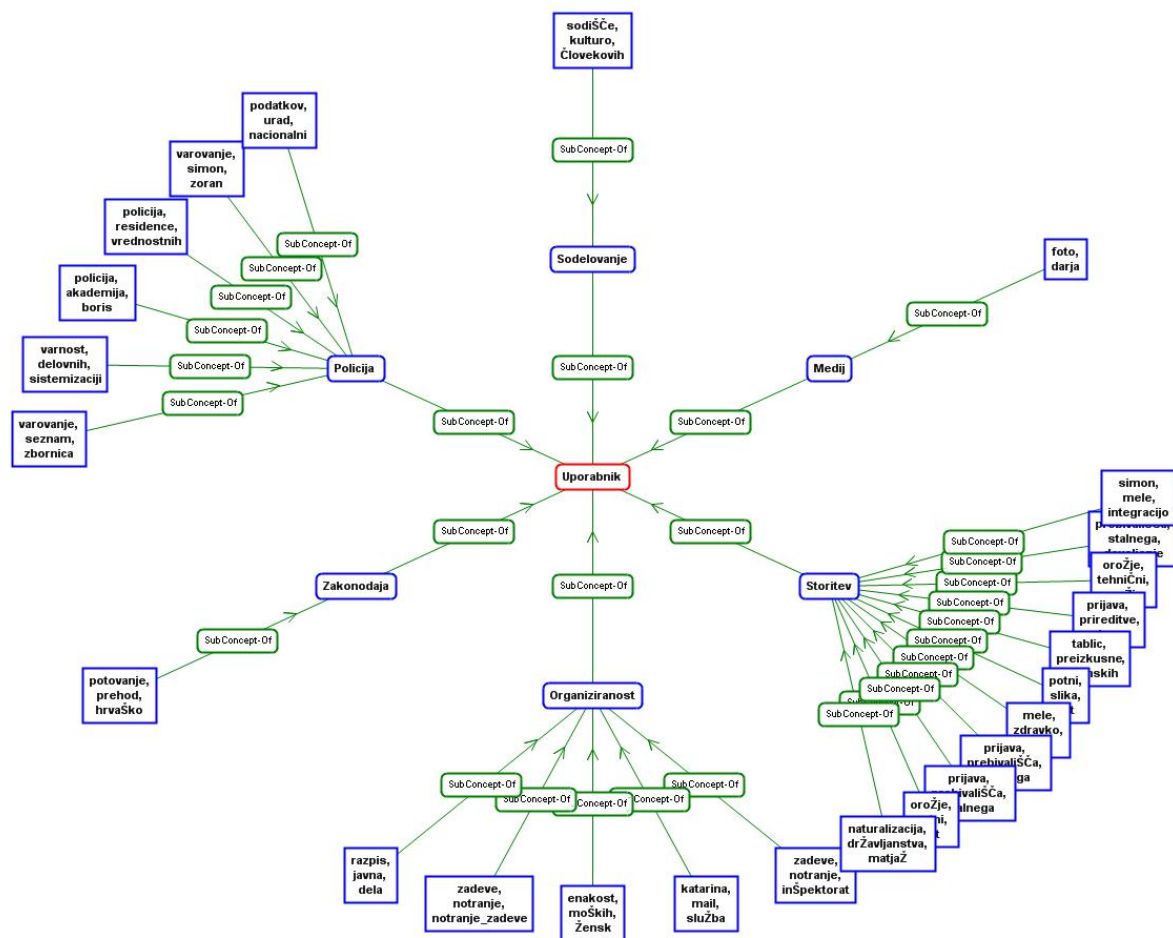


3.3.1 Slika 15: Ontološki koncepti besednih grozdov

Slika 15 prikazuje ontološke koncepte besednih grozdov. Glavne ontološke koncepte sem v bistvu izpeljal iz zavihkov spletne strani MNZ in so naslednji:

1. Storitve (gl. spletni zavihek: MNZ za vas)
2. Organiziranost (gl. spletni zavihek: o ministrstvu)
3. Zakonodaja (gl. spletni zavihek: zakondaja in dokumenti)
4. Policija (gl. spletni zavihek: policija, varnost in nadzor)
5. Sodelovanje (gl. spletni zavihek: mednarodno sodelovanje)
6. Medij (gl. spletni zavihek: medijsko središče)

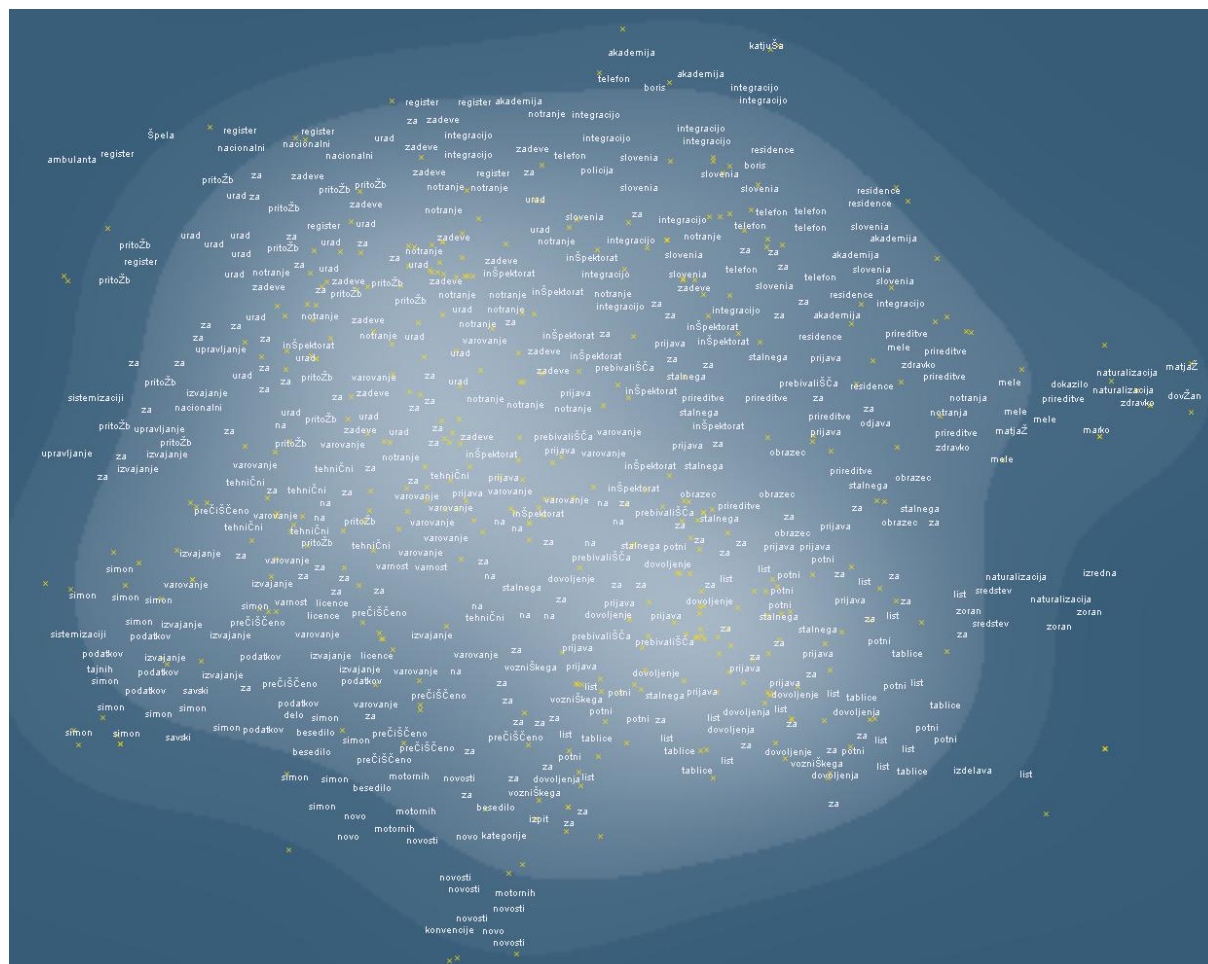
Zaradi boljšega razumevanja snovi, naj še prikažem vizualizacijo konceptov.



3.3.2 Slika 16: Glavni in podrejeni koncepti

Slika 16 prikazuje glavne in podrejene koncepte s pomočjo mrežnega grafa. V središču se nahaja glavni koncept "UPORABNIK", iz katerega vse izhaja. Na podlagi izvedene poizvedbe na domačem (npr. Najdi.si) ali tujem iskalniku (npr. Google) se uporabnik odloči za zadetek, ki je po njegovi oceni najbolj ustrezen. Skratka, odločitev ga napoti na eno izmed podrejenih konceptov, ki so lahko "Storitev" (t.j. spletna stran "MNZ za vas"), "Organiziranost" (t.j. spletna stran "O ministrstvu"), "Zakonodaja" (t.j. spletna stran

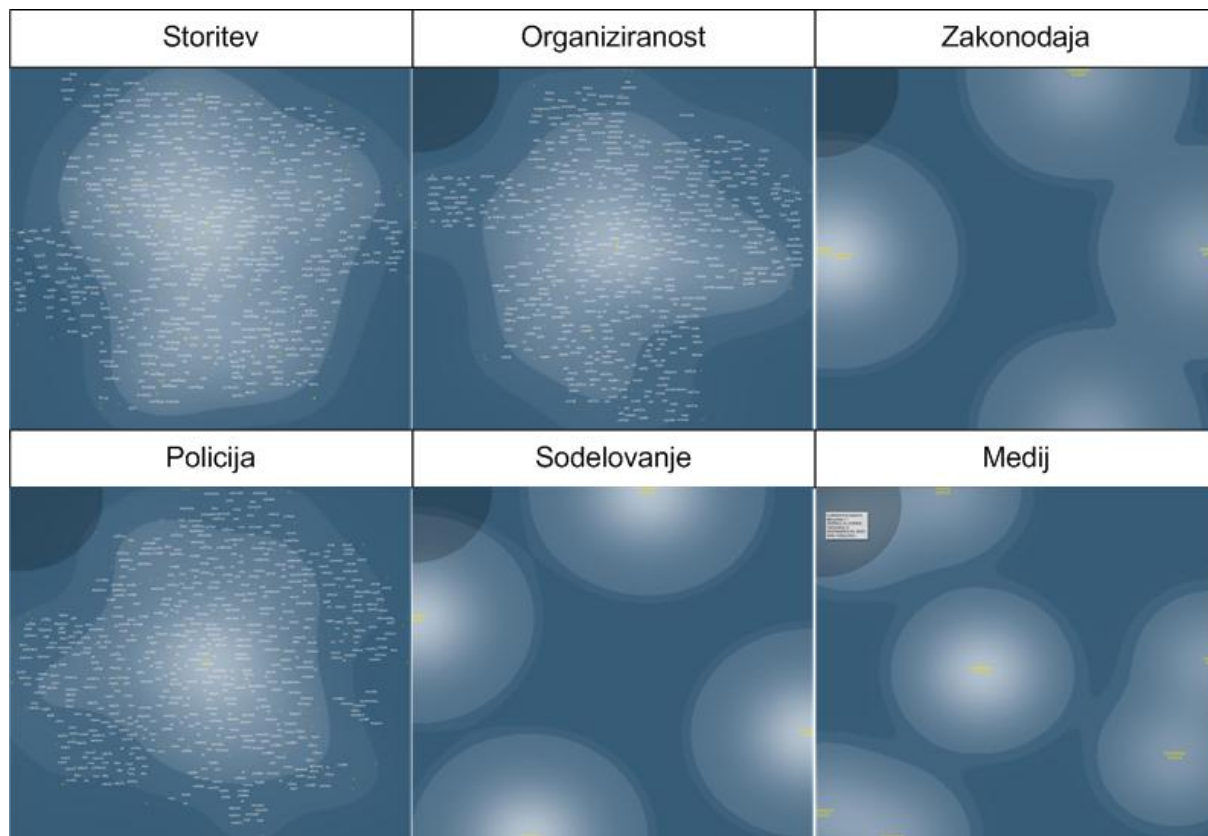
"Zakonodaja in dokumenti"), "Policija" (t.j. spletna stran "Policija, varnost in nadzor"), "Sodelovanje" (t.j. spletna stran "Mednarodno sodelovanje") in nazadnje "Medij" (t.j. spletna stran "Medijsko središče"). Na spodnji stopnji se nahajajo besedni grozdi (pod-podkoncepti), ki sem jih izpeljal iz poizvedb s pomočjo programskega orodja AntConc. Nekateri besedni grozdi so si med sabo manj oddaljeni in s tem bolj sorodni (*angl.: similarity*), kar bo nazorneje pokazala naslednja slika.



3.3.3 Slika 17: Atlas besednih grozdov

Slika 17 prikazuje atlas besednih grozdov, znotraj katerega lahko ugotovimo bližino (ali pa oddaljenost) med njimi. Pojemovna bližina med dvema konceptoma opredeljuje stopnjo vsebinske sorodnosti. Blizu sredine se nahajajo predvsem pojmi (npr. prijava, prireditve, varovanje, dovoljenje, prebivališča, stalnega, pritožb, varnost), ki imajo več sorodnih predstavnikov, medtem ko se na obrobju nahajajo pojmi z manjšim vplivom oziroma številom sorodnikov (npr. ambulanta, dela, medicine, centralni, register, list, cenik, residence, inženir, novosti, izredna, sistemizaciji).

Najbolj jasno podobo o razmerju vsebinske sorodnosti besednih grozdov dobimo na osnovi vizualizacije posameznih pod- konceptov.



3.3.4 Slika 18: Vizualizacija posameznih pod- konceptov

Slika 18 prikazuje vizualizacijo posameznih pod- konceptov. Največje število besednih grozdov vsebujejo Storitev, Organiziranost in Policija, medtem ko je precej manj besednih grozdov znotraj Zakonodaja, sodelovanje in Medij.

Prav ta ugotovitev nas lahko pripelje do zamisli o preučevanju številčnosti in intenzivnosti obiskov uporabnikov na posameznih spletnih zavihkih MNZ, ki sem jih že omenil (gl. stran 21 – 22). V ta namen bom izvozil ustrezne spletometrične podatke iz Google analytics. Ugotoviti želim, kateri spletni zavihek na MNZ je z vidika uporabnika najbolj obiskan oziroma uporaben. Spletni zavihki vsebujejo pojme, ki jih tudi najdemo znotraj besednih grozdov. Mnenja sem, da bi bilo možno vsebinsko in oblikovno podobo spletnih strani MNZ izboljšati, tako da bi uporabnik še lažje našel zaželeno informacijo. Ob tem naj dodam, da spremembe ne bi bile zelo obsežne, ker je sicer vsebinska zasnova spletnih strani MNZ dokaj dobra.

3.4 Analiza obiskov in ključnih besed

Iz Google analytics sem v Excellent analytics izvozil prvih 10000 podatkovnih vrst s poudarkom na dimenzijama "Ključna beseda (poizvedba)" in "Spletna stran" ter na tri merske enote "Čas seje na strani" "Obiskovalec" in "Obisk". Del teh podatkov kaže naslednja preglednica.

3.4.1 Preglednica 3: Del podatkov o koristnosti spletnih strani MNZ z vidika uporabnika

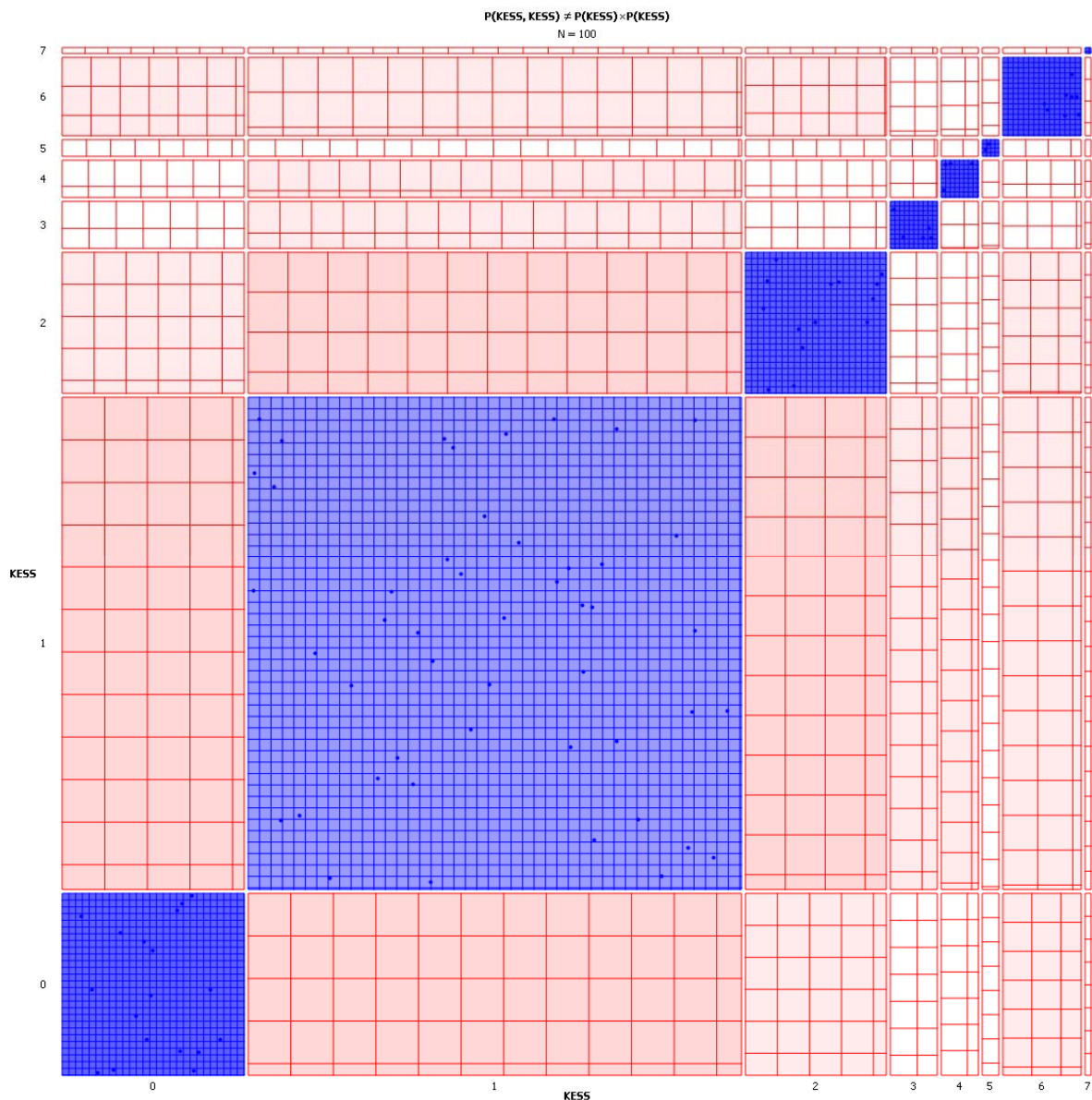
KB	SS	Tp	visitors	OB
(not set)	Ministrstvo za notranje zadeve Vlada Republike Slovenije	0:00:00	38520	150976
(not set)	MNZ za vas Ministrstvo za notranje zadeve	0:39:30	31941	139670
(not set)	Novica Ministrstvo za notranje zadeve	0:00:00	32444	28170
(not set)	404: Vsebina ne obstaja - Not found Ministrstvo za notranje zadeve	11:11:02	23138	23332
(not set)	Javna naročila Ministrstvo za notranje zadeve	11:24:27	8666	9707
(not set)	Veljavni predpisi Ministrstvo za notranje zadeve	17:35:48	4989	3271
(not set)	Javna naročila po postopku zbiranja ponudb Ministrstvo za notranje zadeve	14:43:00	4675	3164
(not set)	Revija Varnost Ministrstvo za notranje zadeve	19:00:55	2418	3141
(not set)	Statistični podatki s področja prometa Ministrstvo za notranje zadeve	9:19:44	1799	2753
(not set)	Ministry of the Interior Government of the Republic of Slovenia	3:00:09	4689	2079
alenska klepac	Ministrstvo za notranje zadeve Vlada Republike Slovenije	4:56:12	6	1739

Preglednica 3 prikazuje del podatkov o koristnosti spletnih strani MNZ z vidika uporabnika. V prvem stolpcu se nahajajo ključne besede (KB), v drugem spletni strani (SS), v tretjem so podatki o času trajanja spletne seje (Tp) na spletnih straneh MNZ, v četrtem število obiskovalcev in v petem so podatki o številu obiskov na posameznih spletnih straneh MNZ. Podatek z izrazom "(not set)" nam pove, da je uporabnik brez izvedene poizvedbe neposredno prispel na določeno spletno stran MNZ. 100 /10000 podatkovnih vrst sem klasificiral od 0 do 7.

3.4.2 Preglednica 4: Del klasificiranih podatkov o obiskanih spletnih strani

KB	SS	Tp	visitors	OB	KESS
(not set)	Ministrstvo za notranje zadeve Vlada Republike Slovenije	0:00:00	38520	150976	0
(not set)	MNZ za vas Ministrstvo za notranje zadeve	0:39:30	31941	139670	1
(not set)	Novica Ministrstvo za notranje zadeve	0:00:00	32444	28170	6
(not set)	404: Vsebina ne obstaja - Not found Ministrstvo za notranje zadeve	11:11:02	23138	23332	7
(not set)	Javna naročila Ministrstvo za notranje zadeve	11:24:27	8666	9707	2
(not set)	Veljavni predpisi Ministrstvo za notranje zadeve	17:35:48	4989	3271	3
(not set)	Javna naročila po postopku zbiranja ponudb Ministrstvo za notranje zadeve	14:43:00	4675	3164	2

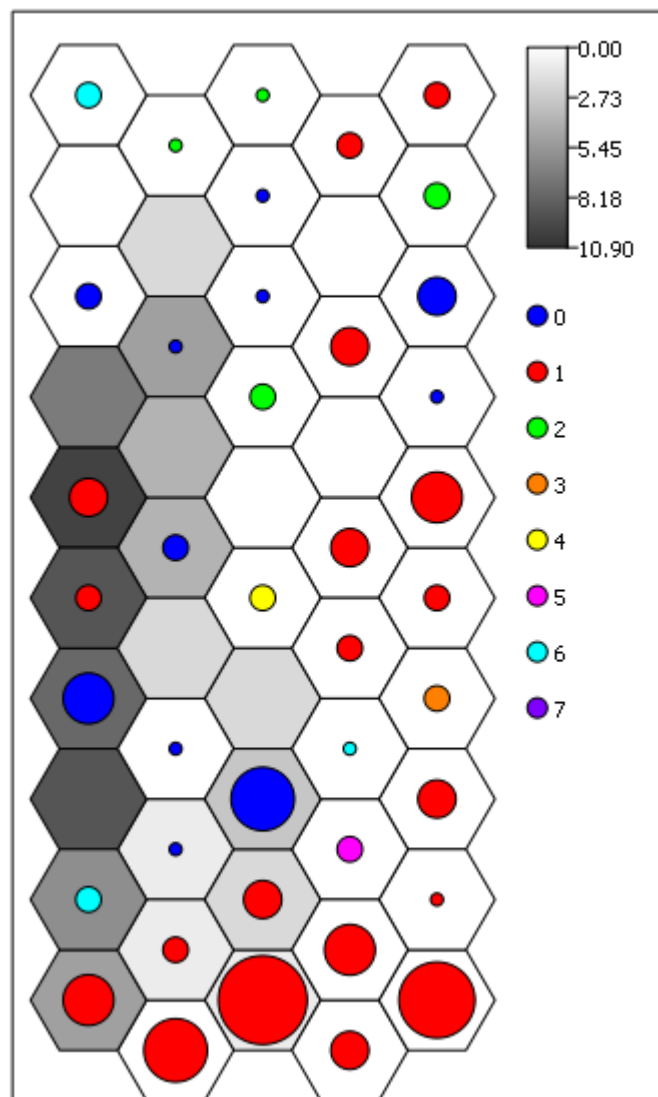
Preglednica 4 prikazuje klasificirane (v nadaljevanju: KESS) podatke o obiskanih spletnih straneh MNZ. Obiskane spletne strani MNZ sem razvrstil od ena do šest na osnovi besednih grozdov oziroma spletnih zavihkov MNZ (gl. sliki 17 in 18). Obiske na glavno stran sem razvrstil kot nič, medtem ko sem kot obisk na nedostopne (gre predvsem za spletne strani, ki so spremenile lokacijo) vsebine razvrstil v skupino sedem. Poskusni vzorec sem procesiral skozi programsko orodje Orange Canvas in preizkusil različne vizualizacijske tehnike. Najbolj učinkovite v tem primeru bodo prikazane v nadaljevanju.



3.4.3 Slika 19: Površinski delež posameznih KESS

Slika 19 prikazuje površinski delež posameznih KESS s pomočjo vizualizacijske tehnike Sievov diagram. Ugotovimo lahko, da je najmočnejši predstavnik skupina 1, ki jo lahko poistovetimo s pod- konceptom "Storitev" oziroma s spletnim zavihkom "MNZ za Vas". Na

osnovi majhnega klasificiranega vzorca (zajel sem prvih 100) sem lahko videl, da lahko skupino 1 povežemo s kazalci kot so največje število obiskov, obiskovalcev in celo z najdaljšim časom spletnih sej. Sledijo ji skupina 0 (glavna stran MNZ), 2 ("Organiziranost" oziroma spletni zavihek "O Ministrstvu"), 6 ("Medij" oziroma "Medijsko središče"), 3 ("Zakonodaja" oziroma "Zakonodaja in dokumenti"), 4 ("Policija" oziroma "Policija, varnost in nadzor"), 5 ("Sodelovanje" oziroma "Mednarodno sodelovanje") in na zadnje mesto se je uvrstila skupina 7 (spletne strani so spremenile lokacijo). Naj v nadaljevanju prikažem še isto snov z vizualizacijsko tehniko SOM (*angl.: self-organizing maps*).



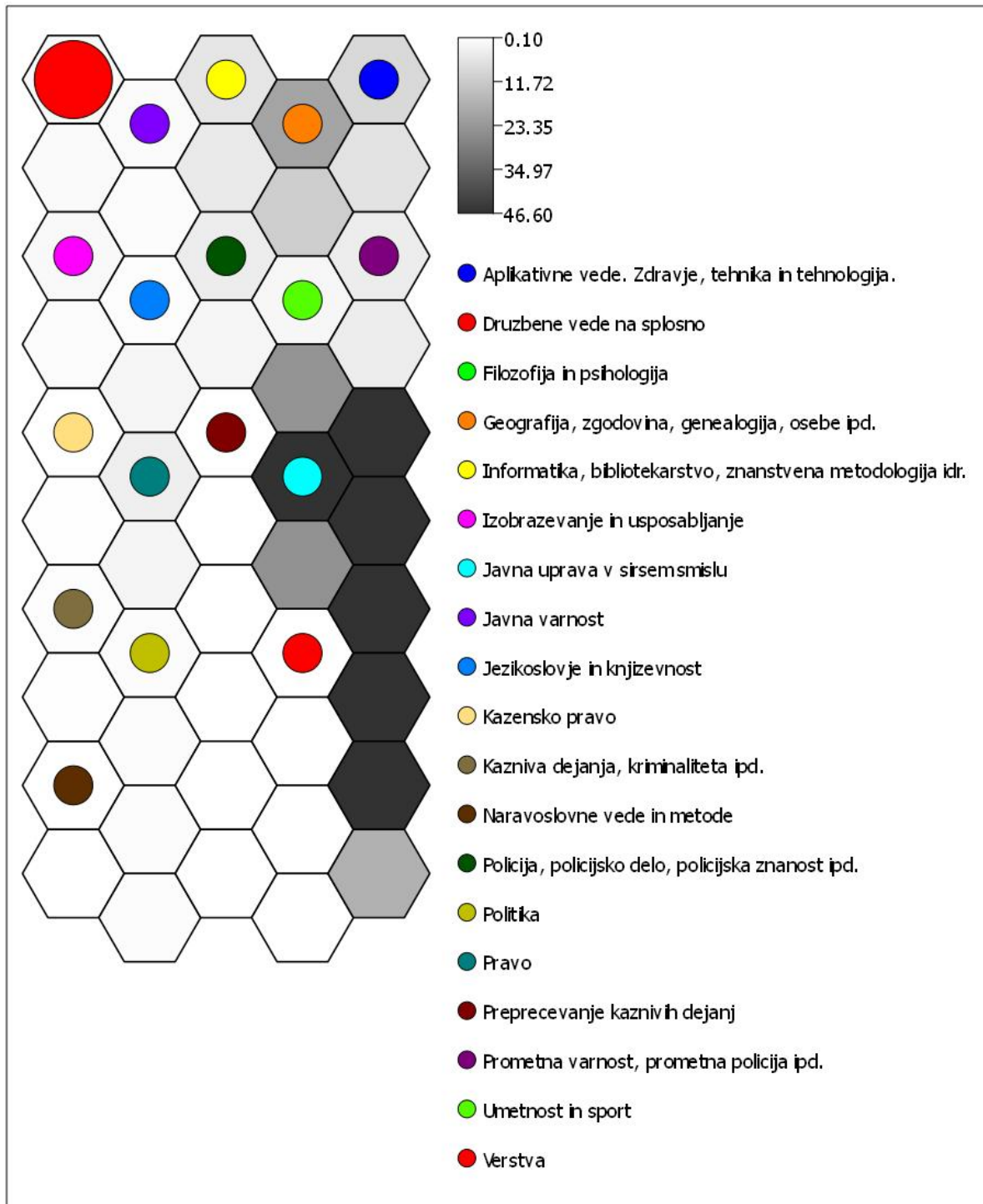
3.4.4 Slika 20: Razvrščanje v skupine s pomočjo vizualizacijske tehnike SOM

Slika 20 prikazuje razvrščanje v skupine s pomočjo vizualizacijske tehnike SOM (heksagonalna topologija). Pri razvrščanju v skupine s SOM sem uporabil algoritem Gaussove sosesčine s poudarkom na večjem deležu predstavnikov. Primerjal sem med sabo ključne besede (KB) in KESS. V najbolj temnosivem šesterokotniku se nahaja rdeč krog, ki ima najvišjo vrednost in je razvrščen kot skupina 1. Nasploh je možno opaziti, da je skupina 1 najmočnejši predstavnik, ki ima tudi največ najbližjih sosedov. V bistvu lahko ugotovimo podobno kot pri Sievovem diagramu.

Skratka, največje število odjemalcev informacij na spletnih straneh MNZ izhaja iz vrst zunanjih uporabnikov oziroma državljanov. Njihova potreba po natančnih informacijah je pretežno v predznaku reševanja problemov iz vsakdanjega življenja (npr. prijava prebivališča, vozniška dovoljenja, osebna izkaznica, potni listi). Prav ta krepko izražena informacijska potreba državljanov in drugih prebivalcev v bistvu napeljuje k zamisli, da bi bilo smiselno nekoliko reorganizirati informacijsko ponudbo. Gre za težnjo, da bi bilo čim več koristnih in uporabnih informacij na enem spletnem mestu. Osnovna zamisel je v tem, da bi glavno stran MNZ premestili na spletni zavihek "Medijsko središče", medtem ko bi spletni zavihek "MNZ za vas" namestili na glavno stran. Ključna pridobitev tovrstne optimizacije spletnih strani MNZ, bi se kazala v tem, da bi prebivalci Republike Slovenije še hitreje prišli do potrebnih informacij. Povrhu tega bi takšna optimizacija poskrbela za še večjo preglednost z vidika zunanjega uporabnika (preglednost in prijaznost spletne strani bi bilo možno še okrepiti s pomočjo ustreznih oziroma prepoznavnih ikon).

4 Spoznanja, ideje in nova znanja

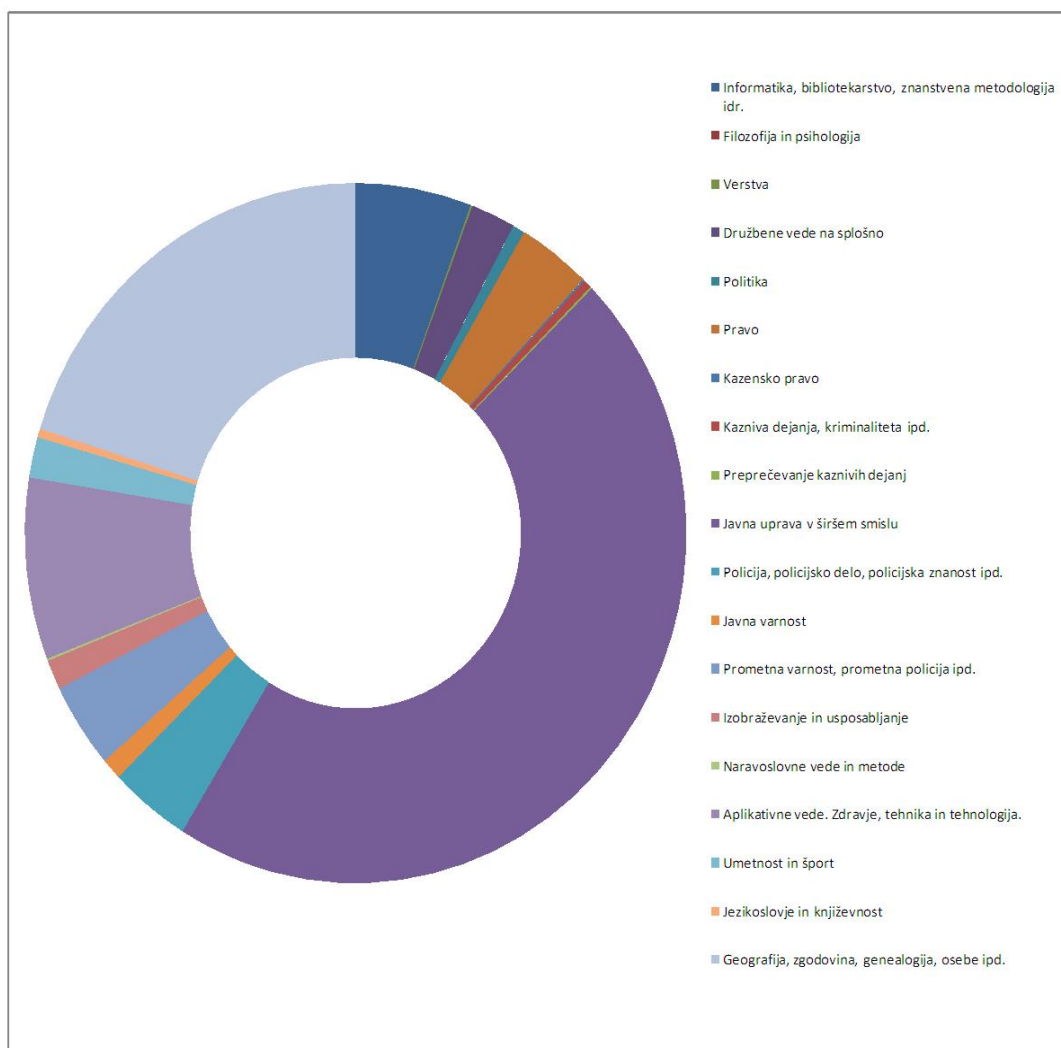
V tem poglavju bom navedel spoznanja, ideje in nova znanja do katerih se je bilo možno dokopati.



4.1 Slika 21: Znanstvena strokovna področja znotraj SOM znanja

Slika 21 prikazuje znanstvena strokovna področja znotraj SOM znanja s pomočjo merske

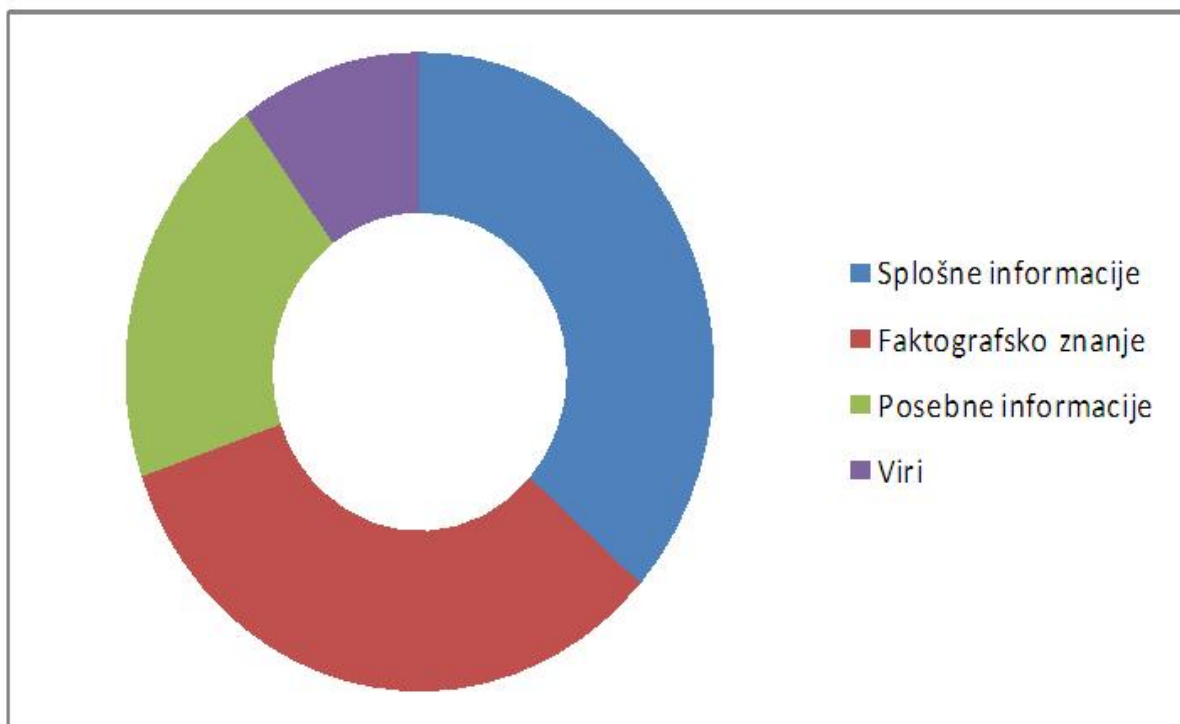
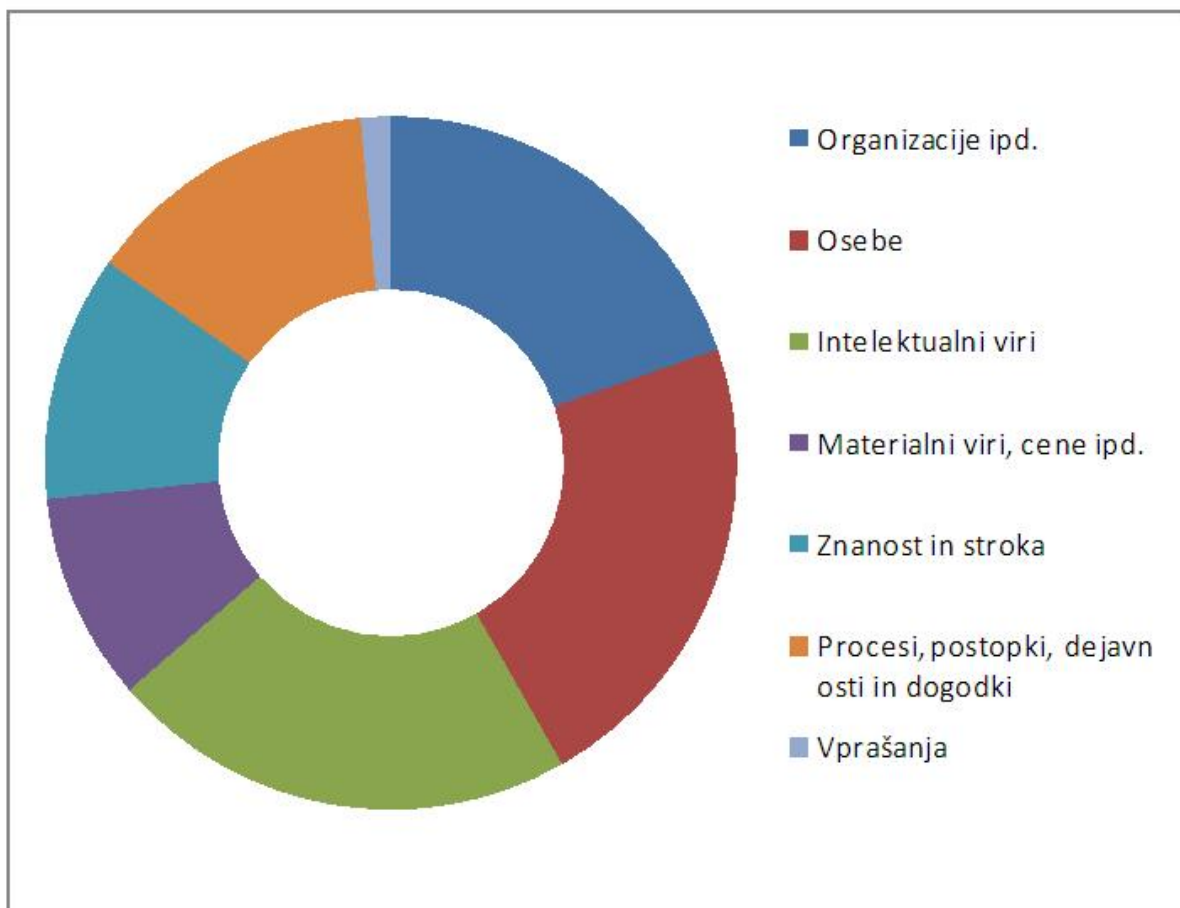
enote odstotkovnih deležev (gl. mersko skalo v levem zgornjem kotu slike 21 od 0,10 do 46,60). Na prvem mestu lahko opazimo, da se nahaja mapa javne uprave (UDK 35) v širšem smislu (gl. svetlomoder krog znotraj temnosivega šesterokotnika. Temu področju sledijo geografija, zgodovina, genealogija, osebe idr. (UDK 9), aplikativne vede, zdravje, tehnika in tehnologija (UDK 6), informatika, bibliotekarstvo in znanstvena metodologija idr. (UDK 0), javna varnost (UDK 351.78), pravo (UDK 34), policija, policijsko delo, policijska znanost ipd. (UDK 351.74), družbene vede na splošno oziroma v širšem smislu (UDK 3), umetnost in šport (UDK 7), izobraževanje in usposabljanje (UDK 37), prometna varnost, prometna policija idr. (UDK 351.81), politika (UDK 32), jezikoslovje in književnost (UDK 0), kazniva dejanja, kriminaliteta ipd. (UDK 343.3/.7), naravoslovne vede in metode (UDK 5), preprečevanje kaznivih dejanj (UDK 343.85), kazensko pravo (UDK 343) in verstva (UDK 2). Na podlagi naslednje slike bo možno še videti razmerja, ki jo zastopajo obravnavana področja.



4.1.1 Slika 22: Spekter potrebnih vsebin s strani uporabnikov

Slika 22 prikazuje spekter potrebnih vsebin s strani uporabnikov po različnih področjih znanja. Kolobarni diagram nazorno pokaže, da so uporabniki najpogosteje potrebovali vsebine s področja UDK 35 (javna uprava v širšem smislu) t.j. 46,6 % (gl. vijolično barvo na sliki 22). Sledijo vsebine s področja UDK 9 (geografija, zgodovina, genealogija, osebe idr.) t.j. 20,4 %, UDK 6 z 8,5 % (aplikativne vede, zdravje, tehnika in tehnologija), UDK 0 s 5,7 % (informatika, bibliotekarstvo, znanstvena metodologija idr.), UDK 351.78 s 4,0 % (javna varnost), UDK 34 s 3,6 % (pravo), UDK 351.74 z 2,9 % (policija, policijsko delo, policijska znanost ipd.), UDK 3 z 2,2 % (družbene vede na splošno), UDK 7 z 1,9 % (umetnost in šport), UDK 37 z 1,4 % (izobraževanje in usposabljanje), UDK 351.81 z 1,0 % (prometna varnost, prometna policija), UDK 32 z 0,6 % (politika), UDK 8 z 0,4 % (jezikoslovje in književnost), UDK 343.3/.7 z 0,4 % (kazniva dejanja, kriminaliteta ipd.), UDK 5 z 0,1 % (naravoslovne vede in metode), UDK 343.85 z 0,1 % (preprečevanje kaznivih dejanj), UDK 343 z 0,1 % (kazensko pravo) in UDK 2 s prav tako 0,1 % (verstva). Uporabniki niso niti enkrat povpraševali po UDK 1 (filozofija in psihologija), kar je nekoliko presenetljivo. Skratka, z vidika (zlasti zunanjih) uporabnikov, so najbolj pomembne informacije o storitvah javne uprave (npr. svetovanje o prijavah, navodila) in deloma o njenih javnih uslužbencih (gre predvsem za pridobivanje kontaktnih informacij). V drugo skupino pomembnih informacij bi lahko uvrstili zdravstvene storitve (npr. medicina dela, zdravstveni pregledi), tehnična sredstva (npr. pirotehnika), materiali (npr. eksplozivi) in tehnologijo (npr. video- nadzor, formx). V isto skupino z ozirom na pomembnosti informacij, bi lahko uvrstili potrebo uporabnikov po različnih publikacijah, podatkih in knjižničnih storitvah. V tej drugi skupini gre pretežno za (notranje) uporabnike, ki so zaposleni na MNZ (vključuje Policijo in Inšpektorat za notranje zadeve). V tretjo skupino pomembnih informacij je možno uvrstiti informacije o nadzoru, varnosti, policiji, statistiki in zakonih. Prav tako lahko za tretjo skupino trdim, da gre predvsem za notranje uporabnike oziroma za uslužbence MNZ. Na zadnje lahko še predstavim četrto skupino pomembnih informacij z vidika uporabnikov, kjer gre pretežno za informacijske potrebe zunanjih uporabnikov (npr. izobraževanje študentov, avtošole, kazni za prekršek).

V nadaljevanju naj še na kratko predstavim uporabnikove potrebe po določenih delih in vrstah informacij.



4.1.3 Slika 23: Spekter določenih delov in vrste informacij

Slika 23 prikazuje spekter določenih delov in vrste informacij. Uporabniki (gre predvsem za zunanje) so najpogosteje potrebovali informacije o osebah (22,1 %), kar sovпада s potrebo po

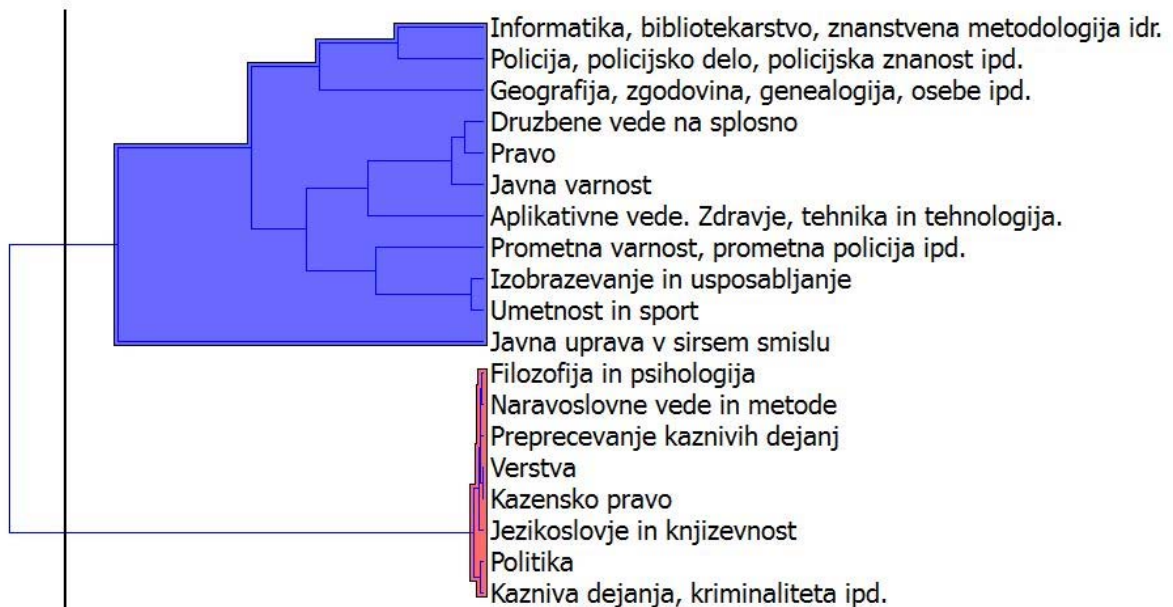
splošnih informacijah (36,5 %). Do podobnega spoznanja je možno priti, če preučimo UKB po organizacijah (19,7 %).

Zelo pogosto so tudi potrebovali informacije o različnih virih (npr. obrazec za prijavo, dovoljenja)(21,9 %), kar je možno povezati tako z ND Viri (10,1 %), ND Faktografsko znanje (33,2 %) kot tudi z ND Posebne informacije (20,2 %). Največjo potrebo uporabnikov po faktografskem znanju je možno opaziti (33,2 %) pri skupini LK 5 (11,5 %). V tem pogledu je možno trditi, da so uporabniki potrebovali več vsebin s področja UDK 351.74, 351.78 in 351.81 kot pri drugih skupinah. Zanimiva je tudi skupina LK 6 (13,8 %), znotraj katere prevladujejo informacijske potrebe po posebnih informacijah (siceršnji delež je 20,2 %).

4.2 Preglednica 5: Področja in profili

Področje	UDKN	Profili
Informatika, bibliotekarstvo, znanstvena metodologija idr.	47	A
Filozofija in psihologija	0	D
Verstva	1	D
Družbene vede na splošno	18	A
Politika	5	D
Pravo	30	A
Kazensko pravo	1	D
Kazniva dejanja, kriminaliteta ipd.	3	D
Preprečevanje kaznivih dejanj	1	D
Javna uprava v širšem smislu	386	A
Policija, policijsko delo, policijska znanost ipd.	24	A
Javna varnost	33	A
Prometna varnost, prometna policija ipd.	8	C
Izobraževanje in usposabljanje	12	B
Naravoslovne vede in metode	1	D
Aplikativne vede. Zdravje, tehnika in tehnologija.	70	A
Umetnost in šport	16	B
Jezikoslovje in književnost	3	D
Geografija, zgodovina, genealogija, osebe ipd.	169	A

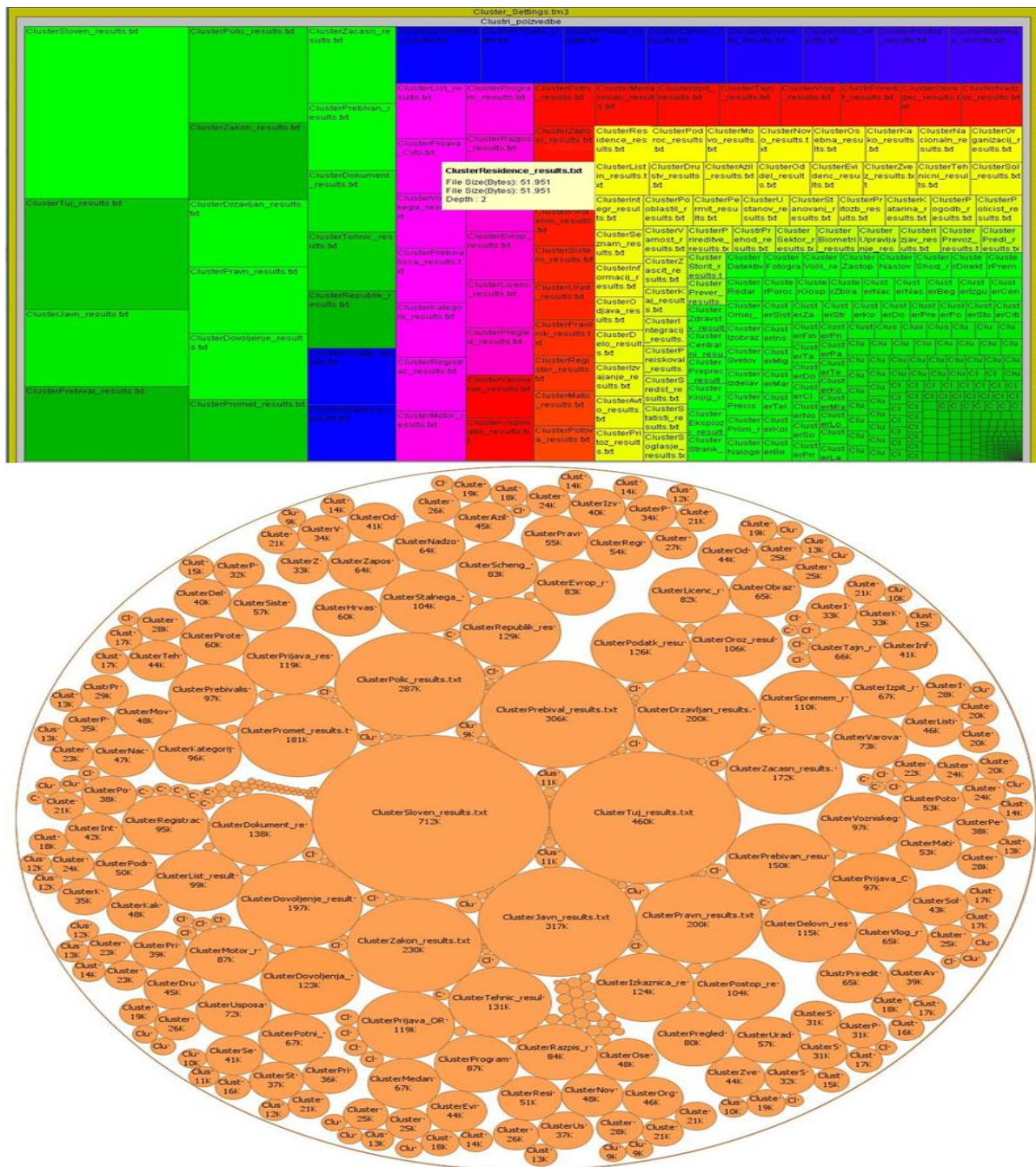
Preglednica 5 prikazuje področja in razvrščene profile uporabnikov v skupine kot so A, B, C in D. Profile uporabnikov sem ustvaril z ozirom na številčnost pojavljanja po področjih. Na osnovi teh podatkov sem izvedel postopek hierarhičnega grozdenja. Izid je viden v obliki dendrograma, ki ga kaže naslednja slika.



4.2 Slika 24: Hierarhično grozdenje področij zanimanja

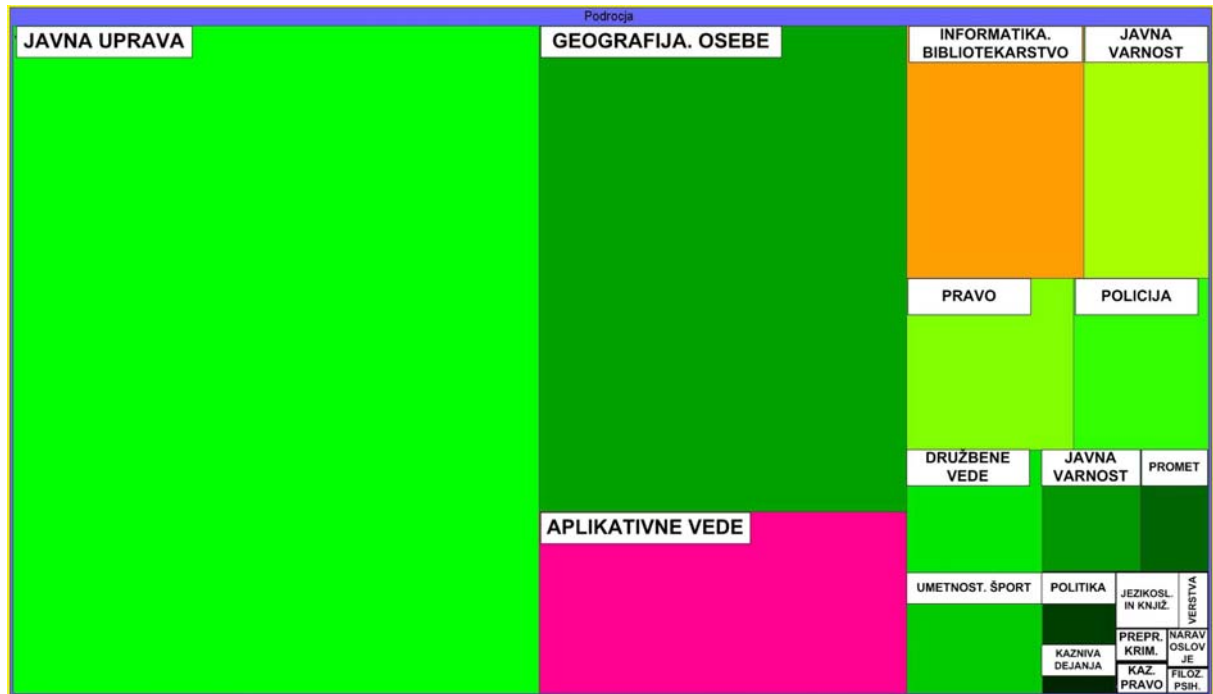
Slika 24 prikazuje hierarhično grozdenje področij zanimanja s pomočjo algoritma evklidske razdalje in Wardove povezanosti (*angl.: Ward's linkage*). Nemudoma spoznamo, da so področja razdeljena v dve glavni skupini (moder in rdeč grozd). Znotraj modre skupine ali grozda so področja, po katerih je bilo s strani uporabnikov največ povpraševanja, medtem ko se znotraj rdeče skupine ali grozda nahajajo področja, po katerih je bilo precej manj zanimanja. Znotraj modrega grozda se nahajajo profili kot so A,B in C (11 predstavnikov; A = 8, B = 2, C = 1), medtem ko se znotraj rdečega grozda nahaja profil D (osem predstavnikov). Še posebej zanimiva je dokaj visoka uvrstitev področja "Informatika, bibliotekarstvo, znanstvena metodologija idr." (gl. zgornji del slike 24), ki posredno kaže na veliko pomembnost Specialne knjižnice znotraj MNZ in OVS. V isto profilno skupino A so se uvrstili tudi področja kot npr. "Javna uprava v širšem smislu", "Geografija, zgodovina, genealogija, osebe ipd.", "Aplikativne vede. Zdravje, tehnika in tehnologija", "Policija, policijsko delo, policijska znanost ipd.". Specialna knjižnica MNZ je v bistvu notranja organizacijska enota, ki že po svoji naravi pokriva različna področja znanosti in stroke. Preseneča nizko število UKB s področja kaznivih dejanj, kriminalitete oziroma kriminologije

nasploh. Ljudje, ki so pri svojih poizvedbah uporabili UKB s področja kriminologije ipd., so mnogokrat obiskali ali spletne strani Specialne knjižnice MNZ ali pa Policije! Morda to dejstvo deloma pojasnjuje majhno stopnjo zanimanja zunanjih uporabnikov po kriminoloških in policijskih vsebinah? Navedene vsebine so dostopne tako na spletnih straneh Fakultete za varnostne vede, Inštitutu za kriminologijo kot tudi Policije. Na osnovi prikazanih besednih grozdov (gl. spodnjo sliko) bom v nadaljevanju kartiral znanstvena in strokovna področja, po katerih so (še zlasti zunanji) uporabniki povpraševali.



4.3 Slika 25: Vizualizacija 368 besednih grozdov

Slika 25 prikazuje vizualizacijo 368 besednih grozdov, ki sem jih s pomočjo UKB izpeljal. Omenjeni besedni grozdi so mi v nadaljevanju pomagali kartirati uporabljena (potrebna) znanstvena in strokovna znanja z vidika uporabnikov spletnih strani MNZ.



4.4 Slika 26: Kartiranje potrebnih znanj z vidika uporabnikov

Slika 26 prikazuje kartiranje potrebnih znanj z vidika uporabnikov, kjer gre za poudarke vsebinskih in velikostnih razmerij. Kot že v prejšnjih poglavjih ugotovljeno so zunanji uporabniki narekovali vsebinske poudarke. Trditev velja še zlasti za področji kot sta "Javna uprava v širšem smislu" in "Geografija, zgodovina, genealogija, osebe idr." Pri področju "Aplikativne vede, zdravje, tehnika, tehnologija idr." in "Informatika. Bibliotekarstvo". Znanstvena metodologija ipd." je vpliv zunanjih uporabnikov velik, vendar manj izrazit. Pri vseh ostalih področjih, po katerih je bilo mnogo manj povpraševanja je bil vpliv notranjih uporabnikov nekoliko močnejši. Vsebine, po katerih so uporabniki spletnih strani MNZ najpogosteje povpraševali pomenijo v bistvu dobro izhodišče za izboljševanje organizacijskih in informacijskih procesov. Prav v tem tiči glavni namen kartiranja znanj, saj preučevalcem in drugim vpletenim nazorno prikaže uporabniške vsebinske poudarke. Dobljena karta znanosti v bistvu daje posredno priporočilo, na katerih področjih se še posebej izplača izboljšati organizacijske, poslovne, informacijske in komunikacijske procese ter aplikacije. Pri spletnem zavihku "MNZ ZA VAS", bi bilo možno spraviti vse spletne podstrani v slikovni model z ustreznimi spletnimi povezavami in opisi. Na področju "Javna uprava v širšem smislu" je možno e-storitve še bolj približati zunanjim uporabnikom, ki bi npr. imeli vse pomembne e-

obrazce znotraj takšnega slikovnega modela s spletnimi povezavami. Predlagano rešitev je smotrno prenesti še na druge vsebinske gradnike kot so registri, evidence, zbirke podatkov o družtvih, političnih strankah, ustanovah idr. Zaradi večje razumljivosti predlagane rešitve naj bralcu ponudim nekaj slikovnih modelov s spletnimi povezavami, ki so bili izdelani za druga problemska področja s strani Specialne knjižnice MNZ:

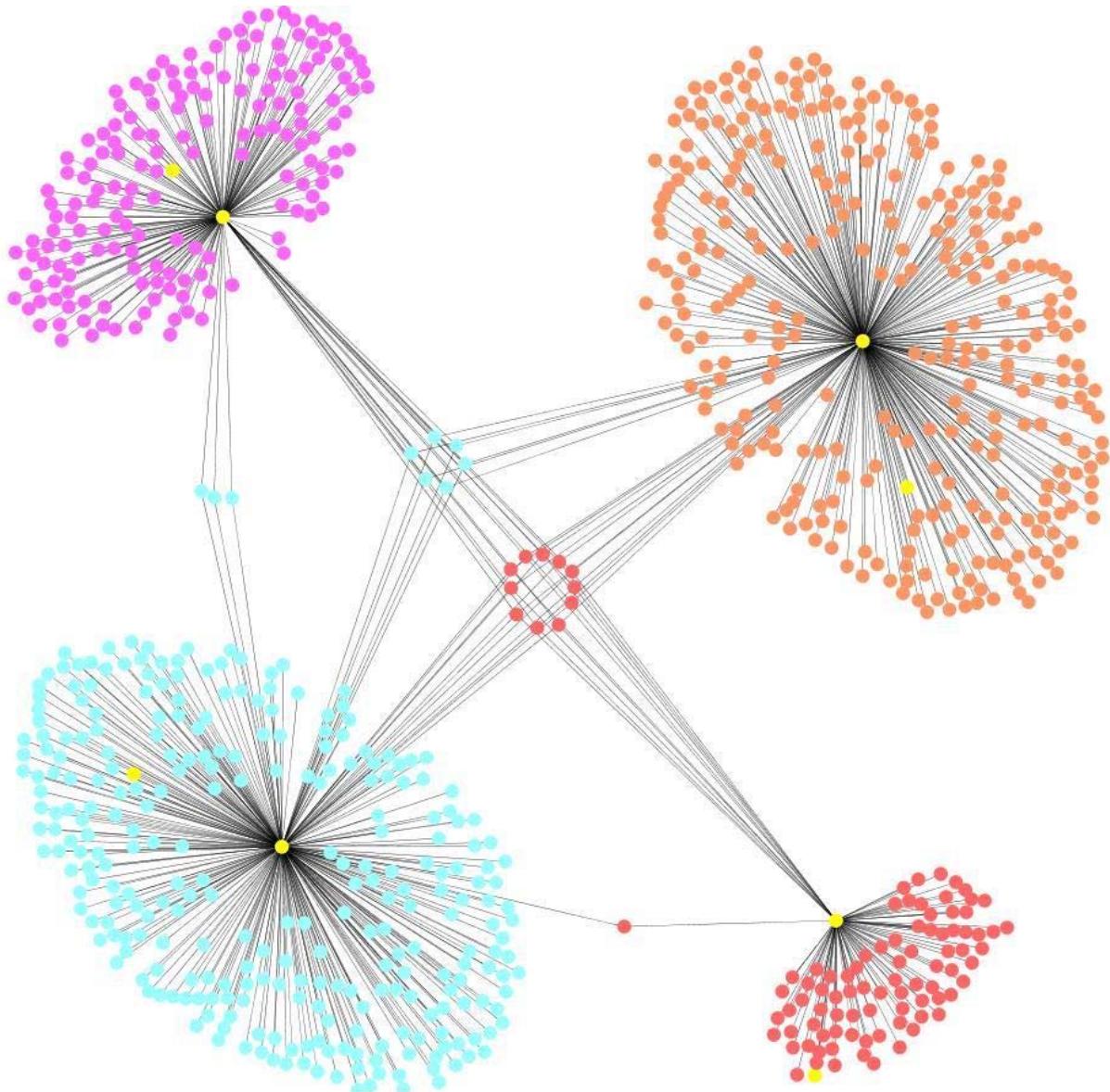
- [ZPOL idr. zakoni kronologija 1991 - 2009](#) (prirejeni model delovnega toka s časovnico),
- [Različne policijske teme](#) (model s pripravljenimi poizvedbami s področja Policije),
- [Gospodarska kriminaliteta](#) (slikovni indeks vsebine v obliki organigrama s področja gospodarske kriminalitete),
- [Družinsko nasilje](#) (model portala za področje družinskega nasilja),
- [Vizualni mikrotezaver](#) (model omrežja s področja kazenskega prava in javne uprave).

Pri področju "Geografija, zgodovina, genealogija, osebe idr." so možne še druge rešitve kot npr. model socialnega omrežja in / ali organizacij. V ta namen je možna uporaba že obstoječega FaceBook-a na glavni spletni strani MNZ. Povpraševanje po nekaterih osebah je stalnica in skoraj vedno aktualno, medtem ko se po nekaterih osebah zgolj začasno povprašuje (npr. navadno gre za aktualne ministre, državne in generalne sekretarje).

Pri področju "Aplikativne vede, zdravje, tehnika, tehnologija idr." je prav tako smiselno združiti pomembne in uporabne vsebine znotraj slikovnega modela s spletnimi povezavami. V ospredju so zdravstvene (npr. zdravniški pregledi, ambulanta dela), tehnične in tehnološke vsebine, ki jih ni bilo smiselno povezovati z inovacijami in drugimi nadgradnjami.

Pri področju "Informatika. Bibliotekarstvo". Znanstvena metodologija ipd." imamo pretežno opraviti z intelektualnimi stvaritvami kot so razna/-e knjižnična gradiva, aplikacije, programska orodja, storitve, informacijske tehnologije itd. V tem vpogledu je možno izpostaviti zamisel o digitalnem tezavru, ki bi vseboval povezave do tovrstnih vsebin.

Nazadnje bi bilo možno združevati različne zamisli v aplikacijski sistem. Razvijanje tovrstne zamisli pa že presega zastavljeno vsebino tega dela. Možnost, da bi uresničili zamisel o določeni uporabni aplikaciji je seveda mnogo večja (npr. javanska spletna aplikacija za primerjavo različnih zakonodaj v Evropski uniji), modelirna in analitična platforma za poslovne procese, ki je dinamične in interaktivne narave. Prav za uresničitev omenjenih rešitev potrebujemo ustrezen kader, ki bi to zmožel znotraj delovne skupine uresničiti.



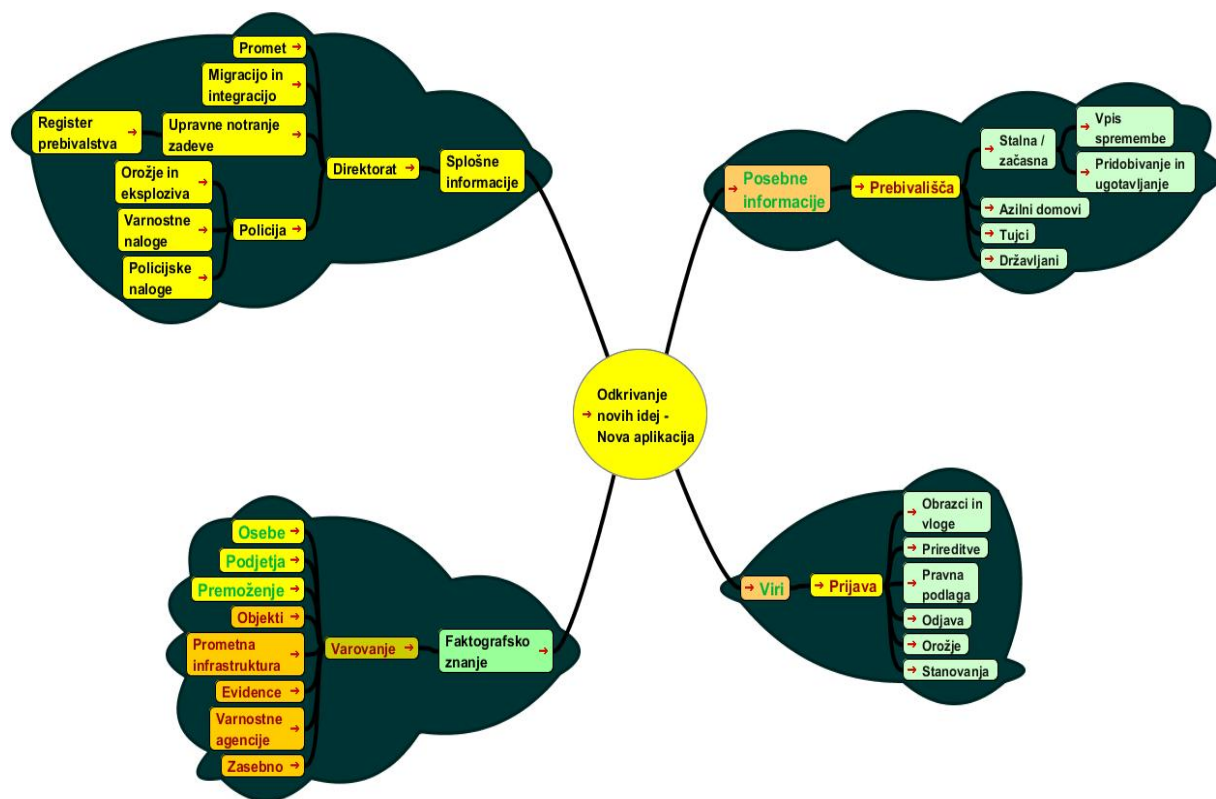
4.5 Slika 27: Ekstrahirano omrežje za UKB, UDK, LK in ND

Slika 27 prikazuje ekstrahirano omrežje za UKB, UDK, LK in ND. Ekstrakcijo velikega omrežja sem izvedel z vtičnikom CytoHubba, s čimer sem določil stopnjo vpliva vozlišča. Po procesiranju podatkov sem kot izid dobil najpomembnejša vozlišča, najkrajše poti in najugodnejše sosede. Nastalo okrnjeno omrežje sem nato obdelal s pomočjo vtičnika Glay, pri čemer sem kot algoritem določil vodilni eigenvektor (*angl.: Leading Eigenvector*), ki s pomočjo matrice izračuna najbolj pozitivne vrednosti znotraj večjih ali manjših skupin. V nadaljevanju sem okrnjeno omrežje obdelal s pomočjo algoritma APC, ki v bistvu išče grozde z najvišjimi vrednostmi in to na principu podobnosti določenega omrežja. V našem primeru je algoritem znotraj štirih glavnih vozlišč določil štiri UKB. Povrhu tega sem vsem glavnim vozliščem izračunal stopnjo vpliva.

Izidi so bili naslednji:

1. Glavni predstavnik vozlišča ND Splošne informacije (stopnja vpliva = 320) je UKB "direktorat",
2. Glavni predstavnik vozlišča ND Faktografsko znanje (stopnja vpliva = 299) je UKB "varovanje",
3. Glavni predstavnik vozlišča ND Posebne informacije (stopnja vpliva = 187) je UKB "prebivališča",
4. Glavni predstavnik vozlišča ND Viri (stopnja vpliva = 96) je UKB "prijava".

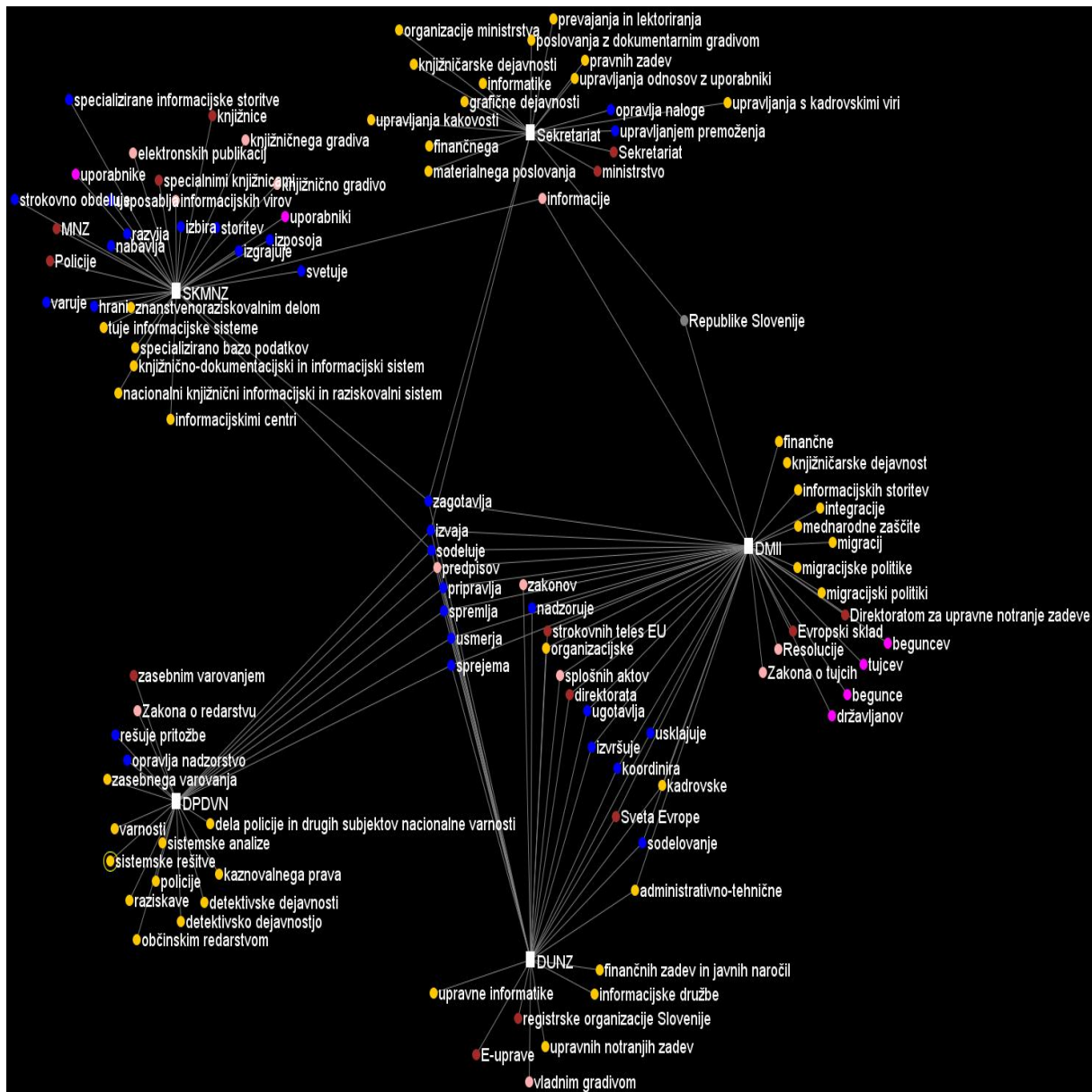
Na osnovi teh izidov sem dobljene glavne predstavnike uporabil kot izhodišče za ustvarjanje miselnega vzorca in v naslednji stopnji za razvijanje novih idej o novi aplikaciji.



4.6 Slika 28: Miselni vzorec za razvijanje nove ideje o novi aplikaciji

Slika 28 prikazuje miselni vzorec za razvijanje nove ideje o novi aplikaciji. Glavni povezovalni pojem predstavlja "Direktorat", saj lahko poveže ostale predstavnike kot so "Varovanje" (gl. Direktorat za policijske in druge varnostne naloge), "Prebivališča" (gl. Direktorat za migracije in integracijo) in "Prijava" (gl. Direktorat za migracije in integracijo, Direktorat za upravne notranje zadeve, Direktorat za policijske in druge varnostne naloge, Policija). Na osnovi prejšnjega spoznanja sem lahko v naslednjem koraku s pomočjo kvalitativne metode vsebinsko analiziral kratka besedila s spletne strani "O MNZ", ki

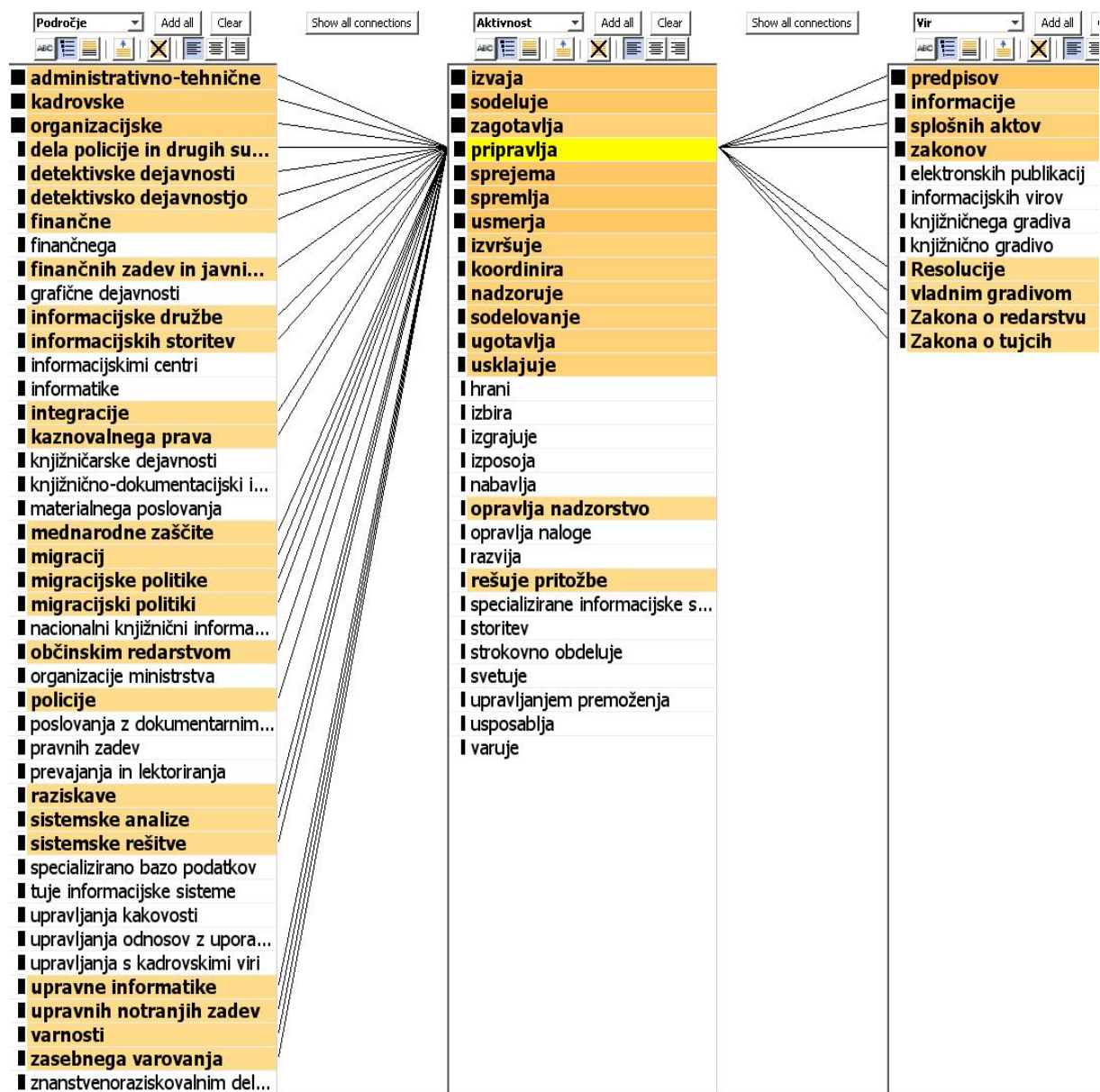
poročajo o delu omenjenih notranjih organizacijskih enot. Analizo besedil sem izvedel s programskim orodjem JigSaw. Predstavil bom zgolj izide v obliki mrežnega grafa.



4.7 Slika 29: Mrežni graf analiziranih predstavitev besedil

Slika 29 prikazuje mrežni graf analiziranih predstavitev besedil s spletnih strani MNZ za že omenjene (gl. opis miselnega vzorca na prejšnji strani) notranje organizacijske enote. K temu sem še dodal Sekretariat in Specialno knjižnico MNZ. S pomočjo mrežnega grafa dobimo nazornejši vpogled v vsebinske gradnike posameznih notranjih organizacijskih enot. Upošteval sem vsebinske gradnike kot so "Področje", "Aktivnost" in "Vir". Na podlagi prikazanega omrežja bi v bistvu že lahko razmišljali o novi (spletni) aplikaciji, ki bi prispevala k dodani vrednosti tako za državljane kot tudi za javne uslužbence. Po eni strani bi

lahko na nekaterih področjih delovanja zblíževala med sabo obravnavane notranje organizacijske enote, a po drugi strani bi lahko bila prijaznejša do državljanov idr. Na sliki 11 in njenem opisu smo že spoznali, da je pomen Specialne knjižnice (SKMNZ) za delovanje MNZ dokaj velik, saj podpira različne in številne informacijske potrebe javnih uslužbencev notranjih organizacijskih enot. SKMNZ je tudi predstavljena na spletu in je s spletnega vidika odprta slehernemu zunanjemu uporabniku. Sorazmerno visok izmerjen vpliv SKMNZ (prav tako sem še vključil Sekretariat, ki je nadrejena SKMZ) je bil tudi glavni vzrok, da sem njene naloge vključil v pričujočo razmišljanje o novi aplikaciji. Naj še prikažem isto snov s pomočjo slikovnega pregledničnega seznama s povezavami in frekvencami.



4.8 Slika 30: Slikovni preglednični seznam s povezavami

Slika 30 prikazuje slikovni preglednični seznam s povezavami in frekvencami v obliki večjih ali manjših stolpcev. Primerjajo se med sabo entitete kot so "Področje" (gl. stolp 1), "Aktivnost" (gl. stolp 2) in "Vir" (gl. stolp 3). Lastnosti entitet dajejo odgovor na naslednja ključna vprašanja:

- s katerimi zadevami se notranje organizacijske enote ukvarjajo?
- kakšne aktivnosti izvajajo?
- katere vire uporabljajo?

Sicer lahko odgovorijo še na naslednja vprašanja, kar pa s slike 30 ni razvidno:

- na katerih geografskih lokacijah potekajo navedene aktivnosti?
- katere osebe so vpletene oziroma kdo so udeleženci?
- katera delovna organizacija izvaja navedene aktivnosti in kje?

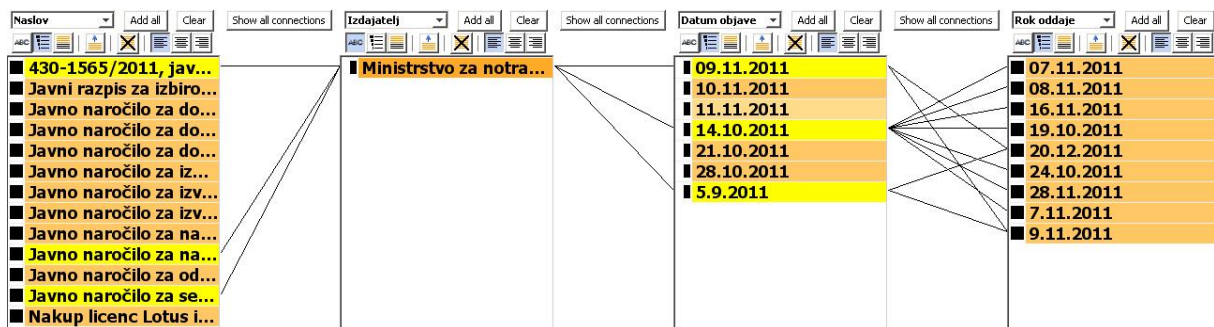
Predstavljeni Vizualni analitični prikaz ni zgolj primeren za kvalitativno analizo različnih besednih gradnikov oziroma entitet, ampak je v tem delu mišljen kot že ekstrahirana zamisel o novi (spletni) aplikaciji, ki bi npr. lahko delovala na načelu vleci in porini (*angl.: push and pull application*). Povrhu tega bi takšna zamišljena (spletna) aplikacija omogočala dodajanje velikega števila stolpcev (npr. do osem ali pa celo še več entitetnih stolpcev). V tem vpogledu bi dodal še zamisel, da bi bila aplikacija relacijska, kar pomeni zmožnost povezovanja različnih entitet (obrazec za prijavo javne prireditve povežemo z obstoječo zakonodajo, obrazec za prijavo prebivališča povežemo z odjavo idr.).

Prav v tem je čar in smisel vizualnih analitičnih modelov, ker analitiku ne olajša zgolj interpretacijo izidov, ampak mu tudi omogoča vizijo, na kakšen način bi lahko bile informacije pregledno in racionalno predstavljene? V primeru, da zraven še prištejemo različne funkcionalnosti, ki jih je možno včasih znotraj takšnih modelov uporabiti, je s tem včasih tudi dana možnost, da se razvije možen model ali vsaj videnje prihodnje aplikacije ipd. Na osnovi skupinske razprave o predstavljenih izidih je bilo možno izpostaviti še naslednje zanimive ideje:

- v aplikacije vključiti razne koristne vsebine za javne uslužbence kot so npr. strategije objavljanja kriminalitete, programi dela, službena poročila, osnutki zakonov. Omenjeno bi bilo bolj primerno za ustrezno intranetno aplikacijo.
- izboljšati vizualno predstavitev in sistematičnost pripravljenih podatkov na spletnih straneh MNZ zaradi lažjega iskanja informacij (npr. signalna pasica za pomoč pri prijavi javne prireditve).

- aplikacija, ki bi pritegnila širšo javnost (npr. v zvezi s prometno varnostjo in preprečevanjem prometnih nesreč). Na glavni spletni strani MNZ se sicer nahaja spletna aplikacija Facebook, kamor državljani in javni uslužbenci lahko odpirajo določene teme, vendar je zanimanja sorazmerno malo. Izpostavljeno dejstvo naj ne bi zaviralo zamisel o spodbujanju dejavnejšega sodelovanja državljanov v smeri preprečevanja prometnih nesreč, zlorab otrok, družinskega nasilja itd. Spletno aplikacijo Facebook je v bistvu možno povezati z drugimi aplikacijami npr. za upravljanje z vsebinami.
- aplikacija, ki bi še dodatno spodbudila učenje slovenskega jezika (npr. z avdio- video predstavitvami, s spletnimi didaktičnimi pripomočki).

Ugotavljam, da je veliko število spletnih strani s področja javnih naročil. Ocenjujem, da bi bilo možno te strani organizirati znotraj .HTML podstrani ali pa v obliki preglednične aplikacije časovnice. Zelo pripravna in pregledna bi lahko bila javanska aplikacija, ki sem jo obravnaval na prejšnji strani.



4.9 Slika 31: Možen model aplikacije za javna naročila na MNZ

Slika 31 prikazuje možen model aplikacije za javna naročila na MNZ. S klikom desne tipke miške bi lahko uporabnik odpiral polna besedila idr. S pomočjo tovrstne aplikacije bi si lahko uporabnik podatke organiziral v različnih pozicijah (npr. po času objave, glede na rok oddaje, po abecedi naslovov javnih naročil, po naročnikih).

Na spletnih straneh MNZ se nahaja veliko število nedelujočih spletnih povezav, ki so deloma nastale ob prehodu iz starega v nov TYPO3 sistem in deloma vsebine niso posodobljene. Predlagam večkratno preverjanje teh spletnih strani s spletnim pajkom ali robotom. Na spletnih straneh MNZ se nahaja tudi mnogo statističnih podatkov, ki bi jih bilo smiselno reorganizirati in morda poenotiti v isto datotečno obliko (npr. samo .XLS)? Trenutno se na teh straneh nahaja veliko statističnih podatkov v stisnjeni obliki (.ZIP).

Vsebinsko in oblikovno podobo spletnih strani MNZ je možno izboljšati, tako da vsebine, po katerih je največ povpraševanja premestimo na glavno spletno stran (uporabniki spletnih strani so najpogosteje obiskali spletni zavihek »MNZ ZA VAS«). Sedanjo glavno spletno stran pa premestimo pod spletni zavihek »Medijsko središče«. Zunanji uporabnik bi tako prej prišel do zaželenih informacij. Sicer je potrebno z vsebinskega vidika spletne strani MNZ zelo pohvaliti. Uporabniki precej lažje najdejo zaželene vsebine, kajti v obdobju pred obnovitvijo spletnih strani MNZ, so mnogi uporabniki vsebine kot so prijava shoda, dovoljenje idr. iskali na spletnih straneh Specialne knjižnice MNZ. Danes je to zelo redek pojav.

Uporabniki, še zlasti zunanji na spletnih straneh MNZ rešujejo predvsem probleme iz vsakdanjega življenja (npr. prijava prebivališča, vozniško dovoljenje, potni list) in zaradi tega je težnja, da bi bilo čim več uporabnih informacij na enem mestu povsem na mestu.

5 Zaključek

V veliki množici informacijskih potreb številnih uporabnikov spletnih strani MNZ sem lahko s pomočjo nove metodologije in odprtokodnih metodoloških orodij izkristaliziral glavne vsebinske tokove. V nadaljevanju sem se še podrobneje omejil na glavni vsebinski tok, ki sem ga klasificiral znotraj UDK 35 (javna uprava v širšem smislu). Nastal je nekakšen model okolja z znanim scenarijem, kjer se je odvijala predvsem zgodba zunanjih uporabnikov s svojimi vsakdanjimi življenjskimi potrebami, da bi rešili vsakdanje probleme. Ob odkrivanju novih idej, znanj in skupinskega pogovora sem spoznal, da je v bistvu možno obstoječo informacijsko ponudbo na spletnih straneh MNZ še izboljšati tako z enostavnimi posegi v spletno stran MNZ kot tudi z izgrajevanjem posebnih aplikacij. Bistvo vseh dejavnosti je biti zadovoljen, t.j. sprejeti in dajati zadovoljstvo. S pomočjo opazovanja množice je možno odkriti nek kolektivni jezik, ki nam v dokaj kratkem časovnem obdobju sporoča več programskih vzgibov kot bi to sleherni posameznik lahko storil. Ta množica, navkljub temu, da v večini primerov proizvaja nekakšne programirane poizvedbe, analitiku le pove nekaj novega. Skratka, na podlagi številnih in intenzivnih ponavljanj je možno rojstvo nove ideje ter v nadaljevanju znanja.

6 Viri in uporabljena programska orodja

Agosti, M., Crivellari, F. & Di Nunzio, G. M. (2011). Web log analysis : a review of a decade of studies about information acquisition, inspection and interpretation of user interaction. *Data Mining and Knowledge Discovery*, (Published online: 06. september), 1-34

Alai, M. (2004). A.I., Scientific Discovery and Realism. *Minds and Machines*, 14(1), 21-42

Alguliev, R. M., Alyguliev, R. M. & Yusifov, F. F. (2007). Automatic identification of the interests of web users. *Automatic Control and Computer Sciences*, 41(6), 320–331

Aula, A. (2005). User study on older adults' use of the Web and search engines. *Universal Access in the Information Society*, 4(1), 67-81

BOLLEN, J. & Van de Sompel, H. (2006). Mapping the structure of science through usage. *Scientometrics*, 89(2), 227-258

Broder, A. (2002). A taxonomy of web search. *ACM SIGIR Forum*, 36(2).
Available at: <http://www.sigir.org/forum/F2002/broder.pdf>

Bystryakova, A. Y. & Mizintseva, V. V. (2011). Innovation informatization for state financial control bodies. *Scientific and Technical Information Processing*, 38(2), 10–17

Cantadora, I., Konstasb, I. & Jose, J. M. (2011). Categorising social tags to improve folksonomy-based recommendations. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(1) 1–15

Cappelin, R. & Post, R. (2009). International knowledge and innovation networks: knowledge creation and innovation and medium-technology clusters. Cheltenham, Northampton: Edward Elgar. VI, p. 275. (New horizons and regional science)

Carley, K. M. et al. (2011). "ORA User's Guide 2011," Carnegie Mellon University, School of Computer Science, Institute for Software Research, Technical Report, CMU-ISR-11-107.

Carmagnola, F. et al. (2011). Supporting content discovery and organization in networks of contents and users. *Multimedia Systems*, 17(3), 199-218

DiGiacomo, J. (2003). Implementing Knowledge Management as a Strategic Initiative: thesis. Monterey: [J. DiGiacomo]. P. 99.

Eggers, W. D. & Singh, S. K. (2009). The public innovator's playbook : Nurturing bold ideas in government. Massachusetts: Harvard Kennedy School, p. 164

Available at:

http://www.deloitte.com/assets/Dcom-Global/Local%20Assets/Documents/dtt_ps_innovatorsplaybook_100409.pdf

Glennisson, P., Glänzel, W. & Persson, O. (2005). Combining full-text analysis and bibliometric indicators : a pilot study. *Scientometrics*, 63(1), 163-180

- Goh, O. S., Fung, C. C. & Wong, W. (2008). Query Based Intelligent Web Interaction with Real World Knowledge. *New Generation Computing*, 26(1), 3-22
- Hafner, T. (2008). Commercial use of Google Analytics study. Kranj: [T. Hafner], 88 f., [3] f. approach.
- Jurisica, I., Mylopoulos, J. & Yu, E. (2004). Ontologies for knowledge management : an information systems perspective. *Knowledge and Information Systems*, 6(4), 380-401
- Keßler, C. (2011). What is the difference? : a cognitive dissimilarity measure for information retrieval result sets. *Knowledge and Information Systems*, (Published online: 12 februar), 1-22
- Landrin-Schweitzer, Y., Collet, P. & Lutton, E. (2006). Introducing lateral thinking in search engines. *Genetic Programming and Evolvable Machines*, 7(1), 9-31
- Miao, Q., Li, Q. & Zeng, D. (2010). Fine-grained Opinion Mining by Integrating Multiple Sources review. *Journal of the American Society for Information Science and Technology*, 61(11), 2288 - 2299
- Moens, M. F. (2001). Innovative techniques for legal text retrieval. *Artificial Intelligence and Law*, 9(1), 29-57
- Ovchenkova, E. A. (2010). The internet as a global search system for scientific articles on information and communication. *Scientific and Technical Information Processing*, 37(3), 178-186
- Papadimitriou, A., Symeonidis, P. & Manolopoulos, Y. (2011). A generalized taxonomy of explanations styles for traditional and social recommender systems. *Data Mining and Knowledge Discovery*, published online 27. march 2011.
Available at:
<http://delab.csd.auth.gr/papers/PSM2011DAMI.pdf>
- Pearson, M. & Somekh, B. (2003). Concept-Mapping as a research Tool : a study of primary children's representations of information and communication technologies (ICT). *Education and Information Technologies*, 8(1), 5-22
- Pennington, D. D. (2011). Bridging the disciplinary divide : co-Creating research ideas in eScience teams. *Computer Supported Cooperative Work*, 20(3), 165-196
- Polanco, X., Roche, I. & BESAGNI, D. (2006). User science indicators in the Web context and co-usage analysis. *Scientometrics*, 66(1), 171-182
- Smith, G. (2004). Folksonomy: social classification.
Accessible online at:
http://atomiq.org/archives/2004/08/folksonomy_social_classification.html
- Soller, A. (2004). Computational modeling and analysis of knowledge sharing in collaborative distance learning. *User Modeling and User-Adapted Interaction*, 14(4), 351-381
- TEXT mining: classification, clustering, and Applications. (2009). Boca Raton, London, New

York: CRC Press. XXX, p. 290., [8] p.. approach. (Chapman & Hall / CRC data mining and knowledge discovery series)

THELWALL, M. A., Wilkinson, D. & Uppal, S. (2010). Data mining Emotion and social network communication: gender differences and MySpace. *Journal of the American Society for Information Science and Technology*, 61(1), 190-199

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4), 189-208

Wei, W. et al. (2011). "Handling Weighted, Asymmetric, Self-Looped, and Disconnected Networks in ORA," Carnegie Mellon University, School of Computer Science, Institute for Software Research, Technical Report, CMU-ISR-11-113

Westerski, A., Iglesias, C. A. & Rico, T. (2010). A model for integration and interlinking of idea management systems. *Metadata and Semantic Research 4th International Conference, MTSR 2010, Alcalá de Henares, Spain, October 20-22. Proceedings, CCIS 108*, 183–194

Wissmann, J. and Bahr, G. S. (2007). Bilingual Mapping Visualizations as Tools for Chinese Language Acquisition. *Human-Computer Interaction, Part II, HCII 2007, LNCS 4551*, 171-180

Yang, B., Song, W. & Xu, Z. (2007). New construction for expert system based on innovative knowledge discovery technology. *Science in China Series F: Information Sciences*, 50(1), 29-40

Zhao, L. & Zhang, Q. (2011). Mapping knowledge domains of Chinese digital library research output, 1994–2010. *Scientometrics*, 89(1), 51-87

Uporabljena programska orodja

AntConc

Cytoscape

Excel 2007

Excellent analytics

Free Plane

Google analytics

JigSaw

NetDraw

Ontogen 2.0

Ora Casos

OrangeCanvas

SocSciBot 4

Tree Mapping