



UNIVERZITET U NIŠU  
ELEKTRONSKI FAKULTET



**Aldina R. Avdić**

**Realizacija servisa pametnog zdravstva i  
njihova integracija u koncept pametnih gradova**

DOKTORSKA DISERTACIJA

Niš, 2021.



UNIVERSITY OF NIŠ  
FACULTY OF ELECTRONIC ENGINEERING



**Aldina R. Avdić**

# **Realization of Smart Health Services and Their Integration into the Concept of Smart Cities**

DOCTORAL DISSERTATION

Niš, 2021.

## Podaci o doktorskoj disertaciji

Mentor:	Prof. dr Dragan Janković Univerzitet u Nišu, Elektronski fakultet
Naslov:	Realizacija servisa pametnog zdravstva i njihova integracija u koncept pametnih gradova
Rezime:	<p>Razvoj informacionih tehnologija doprineo je pojavi koncepta pametnog grada čija je jedna od komponenata i pametno zdravstvo. Ova komponenta je značajna jer pametni gradovi direktno ili indirektno utiču na zdravlje stanovnika.</p> <p>Glavni doprinosi ove doktorske disertacije su predlog softverske platforme koja objedinjuje postojeće servise javnog zdravstva kao i nove servise e-zdravstva, predlog uniformnog načina integracije heterogenih servisa pametnog zdravstva u pametni grad, i predlog i implementacija metoda za obeležavanje termina u medicinskim tekstovima, zasnovanih na veštačkoj inteligenciji, bez kojih deo servisa pametnog zdravstva ne bi mogao da se realizuje na efikasan način.</p> <p>Pomenuta softverska platforma objedinjuje servise pametnog zdravstva koji će biti predloženi (servisi za informisanje i prevenciju bolesti, kontrolu epidemija, pretragu medicinskih dokumenata, automatsku obradu upitnika, praćenje zagađenja vazduha, označavanje medicinskih tekstova, organizaciju skrining programa, itd). Ovi servisi kreirani su na osnovu ulaznih podataka različitog tipa (senzorski podaci o lokacijama, zagađenju vazduha, tekst iz dokumenata i medicinskih informacionih sistema i podaci prikupljeni kraudsorsingom, podaci iz relacionih i nerelacionih baza podataka), pa je potrebno da se ovi heterogeni servisi integrišu na uniforman način.</p> <p>Deo predloženih servisa e-zdravstva baziran je na obradi podataka u medicinskim informacionim sistemima kao i medicinskih tekstualnih dokumenata na srpskom jeziku. Metode za normalizaciju,</p>

obeležavanje i klasifikaciju termina na srpskom jeziku su preduslov za uspešnu realizaciju ovih servisa, te su u okviru ove disertacije predložene i implementirane metode čiji je *F1-rezultat* 0.9082, što je odličan rezultat u poređenju sa metodama za ovu namenu na drugim jezicima. Za njihovu implementaciju potrebno je koristiti metode veštačke inteligencije, kao što su procesiranje prirodnog teksta, istraživanje podataka i teksta, mašinsko učenje itd. Pojedini predloženi servisi e-zdravstva su i praktično implementirani i integrisani u koncept pametnog grada.

Naučna  
oblast:

Elektrotehničko i računarsko inženjerstvo (Računarstvo i informatika)

Naučna  
disciplina:

Medicinska informatika, Obrada prirodnog jezika, Istraživanje teksta

Ključne reči:

pametni gradovi, e-zdravstvo, pametno zdravstvo, servisi pametnog zdravstva, istraživanje teksta, srpski jezik, mašinsko učenje

UDK:

(005.21:711.4+614.2):004.8+(004.431.2+811.163.4)

CERIF  
klasifikacija:

T120, Sistemski inženjering, računarska tehnologija

Tip licence  
Kreativni  
zajednice:

**CC BY-NC-ND**

## Data on Doctoral Dissertation

Doctoral Supervisor: PhD Dragan Janković, full professor  
University of Niš, Faculty of Electronic Engineering

Title: Realization of Smart Health Services and Their Integration into the Concept of Smart Cities

Abstract:

The development of information technologies has contributed to the emergence of the concept of a smart city, and one of the components of this concept is smart health. This component is important because smart cities directly or indirectly affect the health of residents.

The main contributions of this doctoral dissertation are a proposal of a software platform that integrates existing public health services as well as new e-health services, a proposal for a uniform way of integrating heterogeneous smart health services into a smart city, and a proposal and implementation of methods for labelling terms in medical texts, based on artificial intelligence, without which part of the smart health service could not be realized in an efficient way.

The above-mentioned software platform combines the smart health services that will be proposed (services for information and prevention of diseases, epidemic control, search of medical documents, automatic questionnaire processing, air pollution monitoring, labelling of medical texts, organization of screening programs, etc.). These services are created based on input data of different types (sensor data on locations, air pollution, text from documents and medical information systems and data collected by crowdsourcing, data from relational and non-relational databases), so it is necessary to integrate these heterogeneous services in uniform way.

Part of the proposed e-health services is based on data processing in medical information systems as well as medical text

documents in the Serbian language. Methods for normalization, labelling and classification of terms in the Serbian language are a prerequisite for the successful implementation of these services, and within this dissertation methods are proposed and implemented whose *F1-score* is 0.9082, which is an excellent result compared to methods for this purpose in other languages. For their implementation, it is necessary to use artificial intelligence methods, such as natural language processing, data and text mining, machine learning, etc. Some of the proposed e-health services are practically implemented and integrated into the smart city concept.

Scientific  
Field:

Electrical and Computer Engineering (Computer Science)

Scientific  
Discipline:

Medical Informatics, Natural Language Processing; Text mining

Key Words:

smart cities, e-health, smart health, smart health services, text mining, Serbian language, machine learning

UDC:

(005.21:711.4+614.2):004.8+(004.431.2+811.163.4)

CERIF  
Classification:

T120, Systems engineering, computer technology

Creative  
Commons  
License Type:

**CC BY-NC-ND**

# ZAHVALNICA

Ova disertacija predstavlja rezultat višegodišnjeg rada i istraživanja, a inspiracija da krenem ovim putem bili su posvećenost, predani rad i trud mojih profesora i asistenata sa Elektronskog fakulteta i uopšte dobrih nastavnika u toku mog školovanja. Najveću zahvalnost za ostvarenje ovog dugogodišnjeg sna dugujem mom mentoru prof. dr Draganu Jankoviću, čiji su nesebična pomoć i uloženo vreme bili od neprocenjivog značaja za doprinos kvalitetu izrade ove doktorske disertacije. Posebno zadovoljstvo i čast je što je on mentor za izradu moje doktorske disertacije, veliki čovek, pre svega, od koga se mnogo može naučiti, a zatim i profesor i naučnik. Veliku zahvalnost dugujem članovima komisije za ocenu i odbranu disertacije, za konstruktivne sugestije koje će mi koristiti i u budućem radu i istraživanju. Posebnu zahvalnost dugujem kolegici Ulfeti Marovac na prijateljskoj pomoći i radu, podeli znanja i iskustva, koji su učinili da se sigurnijim koracima približim cilju. Zaista je saradnja s njom i profesorom Jankovićem jedno neprocenjivo iskustvo, koje će se, nadam se, nastaviti i u buduće.

Rad u jednom predivnom okruženju, punom zaljubljenika u nauku, imao je veliki uticaj na mene da nastavim ovim putem. Zato se zahvaljujem kolegama sa Državnog univerziteta u Novom Pazaru sa studijskog programa Računarska tehnika i onima koji to sada nisu zbog promene radnog mesta ili odlaska u penziju, ali su prijatelji koje sam ovde srela. U ovoj grupi izdvojila bih profesora Ejuba Kajana kao nekoga čiji su jedinstveni pristup, način rada i oblasti istraživanja veoma uticale i na moje izbore i rad u istraživanjima.

Ovaj uspeh dugujem i svom suprugu i ćerkama, koji su mi bili podrška i motivacija da ostvarim ovaj svoj dugogodišnji san. Njihova ljubav i razumevanje bili su izvori energije i kada je sve izgledalo nemoguće i daleko. Hvala mojim roditeljima što su me uvek bodrili i verovali u mene, i što im je obrazovanje dece uvek bio prioritet, kao i ostaloj velikoj porodici za svu pruženu ljubav i podršku.

Ovu disertaciju posvećujem svojoj deci i svojim studentima, kako bi im bila podsetnik da nikad ne odustaju od svojih ciljeva, da ako vredno rade ono što vole, uspeh neće izostati, i da je sve lakše kada smo okruženi inspirativnim ljudima.

# SADRŽAJ

1.	Uvod.....	10
1.1.	Predmet i cilj naučnog istraživanja.....	11
1.2.	Primenjene naučne metode.....	14
1.3.	Pregled disertacije.....	14
2.	Pametni gradovi.....	17
2.1.	Komponente pametnog grada.....	18
2.2.	Infrastruktura pametnog grada.....	20
2.2.1.	Tehnologije na kojima počiva pametni grad.....	20
3.	Pametno zdravstvo.....	22
3.1.	Odnos elektronskog, mobilnog i pametnog zdravstva.....	23
3.2.	Pitanja bezbednosti i privatnosti zdravstvenih usluga.....	25
3.3.	Pravna regulativa elektronskog poslovanja u Republici Srbiji i osvrt na <i>GDPR</i> .....	26
3.4.	Elektronski zdravstveni servisi u svetu i kod nas.....	31
4.	Arhitektura IS-a koji objedinjuje heterogene servise pametnog zdravstva.....	35
4.1.	Integracija heterogenih servisa e-zdravstva u sistem pametnog grada.....	35
4.2.	Struktura uniformnog veb servisa za realizaciju servisa e-zdravstva.....	38
4.3.	Kolaboracija postojećih sistema u pametnom gradu.....	41
5.	Novi servisi pametnog zdravstva.....	43
5.1.	Servisi zasnovani na <i>LMS</i> sistemima i društvenim mrežama.....	44
5.1.1.	Informisanje, komunikacija i prevencija bolesti.....	47
5.2.	Servisi zasnovani na <i>IoT</i> i geografskim informacionim sistemima.....	50
5.2.1.	Pretraga i prikaz najbližih zdravstvenih ustanova.....	51
5.2.2.	Praćenje kvaliteta vazduha.....	53
5.2.3.	Unapređenje pružanja medicinskih usluga.....	61
5.2.4.	Inteligentni transport pacijenata.....	62
5.2.5.	Mobilni <i>SOS</i> sistem.....	71
5.3.	Servisi zasnovani na <i>NLP</i> metodama i kraudsorsingu.....	71
5.3.1.	Automatsko odgovaranje građanima na pitanja.....	73
5.3.2.	Kontrola epidemije.....	79
5.3.3.	Automatska obrada upitnika i psiholoških testova.....	87
5.3.4.	Uvid u raspored i ocenjivanje lekara.....	93
5.4.	Servisi zasnovani na multimediji i veb tehnologijama.....	95
5.4.1.	Komunikacija sa izabranim lekarima i dežurnom ekipom hitne pomoći.....	95



5.4.2.	Podrška zbrinjavanju u slučaju velikih nesreća.....	97
5.4.3.	Efikasnija realizacija skrining programa.....	98
5.5.	Servisi zasnovani na blokčejnu.....	100
5.5.1.	Obezbeđivanje pouzdanosti i trajnosti podataka.....	101
6.	Obrada podataka u MIS-ima kao osnova za realizaciju servisa e-zdravstva .....	105
6.1.	Stanje u oblasti normalizacije medicinskih podataka.....	106
6.2.	Opis korišćenog skupa podataka .....	107
6.3.	Karakteristike medicinskih podataka.....	108
6.4.	Metoda za normalizaciju medicinskog teksta.....	111
6.4.1.	Tokenizacija .....	112
6.4.2.	Brisanje stop reči.....	112
6.4.3.	Obrada negacije.....	113
6.4.4.	Odsecanje na <i>n</i> -grame .....	114
6.4.5.	Klasifikacija <i>n</i> -grama .....	114
6.5.	Rezultati primene metode za normalizaciju medicinskih dokumenata .....	115
7.	Implementacija servisa e-zdravstva korišćenjem algoritama veštačke inteligencije .....	120
7.1.	Pregled postojećih medicinskih resursa.....	122
7.2.	Skup podataka.....	124
7.3.	Korišćeni resursi.....	126
7.3.1.	Nemedicinski resursi .....	126
7.3.2.	Obrada medicinskih leksičkih resursa.....	127
7.3.3.	Medicinski resursi .....	128
7.4.	Predložene metode za obeležavanje medicinskih termina u slobodnom tekstu .....	130
7.4.1.	Metoda zasnovana na jezičkim resursima .....	130
7.4.2.	Metoda zasnovana na nadgledanom učenju .....	131
7.4.3.	Metoda zasnovana na nadgledanom učenju i pravilima.....	132
7.5.	Metrike korišćene za evaluaciju metoda .....	134
7.6.	Evaluacija metoda za obeležavanje medicinskih termina na srpskom jeziku .....	135
8.	Zaključak .....	142
9.	Pravci daljeg razvoja .....	148
	Literatura .....	150
	Spisak korišćenih skraćenica.....	166
	Spisak slika.....	170
	Spisak tabela.....	173
	Biografija autora.....	174

## 1. UVOD

Informacione i komunikacione tehnologije (IKT) omogućile su uslove da se stvori novo okruženje za stanovnike gradova, efikasno u pogledu potrošnje energije, transporta, uprave, učenja, ekonomskog razvoja i drugih sličnih aspekata – takozvani pametni grad [1]. Sa druge strane, IKT već duže vreme nalaze svoju primenu u zdravstvenim sistemima kroz elektronsko zdravstvo i mobilno zdravstvo [2].

Primeri zastupljenosti IKT u sistemima zdravstvene zaštite ogledaju se kroz realizaciju elektronskog kartona pacijenta, telemedicinu, telekonsultacije, telehirurgiju, upravljanje znanjem o zdravlju, mobilno i sveprisutno zdravstvo itd.

Nameće se pitanje: „Da li se u koncept pametnog grada mogu uključiti i rešenja u oblasti zdravstva?“. Oblast koja se bavi korišćenjem IKT u cilju poboljšanja zdravlja stanovnika u konceptu pametnog grada je pametno zdravstvo [3]. Ova komponenta pametnog grada je nedovoljno zastupljena i razvijena, u odnosu na svoj značaj. Otvoreno pitanje je kako povećati prisustvo pametnog zdravstva u konceptu pametnog grada i direktno uticati na zdravlje stanovništva, jer više manje svi standardni servisi pametnog grada direktno ili indirektno utiču na zdravlje.

Povećanje učešća pametnog zdravstva u konceptu pametnog grada obezbedilo bi:

- lakše i uspešnije lečenje pacijenata i smanjenje administrativnih ograničenja prilikom pružanja medicinskih usluga,
- bolje sprovođenje preventivnih aktivnosti u domenu zdravstva,
- zdraviju životnu sredinu, itd.

Moguće kategorije servisa pametnog zdravstva su usluge zasnovane na kolektivnoj inteligenciji grada, kontekstno-orijentisani servisi, usluge zasnovane na *IoT* (engl. *Internet of Things*) uređajima [4], usluge zasnovane na *kraudsorsingu* (engl. *crowdsourcing*) [5] i istraživanju medicinskih podataka itd.

## 1.1. Predmet i cilj naučnog istraživanja

Realizacija novih servisa e-zdravstva i uniformni način njihove realizacije i integracije u cilju unapređenja koncepta pametnog grada i izazovi koji se mogu pojaviti predstavljaju predmet istraživanja ove doktorske disertacije.

Istraživanja obavljena u okviru doktorske disertacije obuhvataju analizu načina za realizaciju servisa pametnog zdravstva. Analizirani su postojeći servisi e-zdravstva u našoj zemlji i svetu, definisani su novi servisi e-zdravstva koji još nisu zastupljeni u Republici Srbiji i neki od njih su realizovani, i dat je način uniformne implementacije i integracije novih, predloženih zdravstvenih servisa u koncept pametnog grada.

Fokus istraživanja u ovoj doktorskoj disertaciji je na razvoju servisa pametnog zdravstva u okviru koncepta pametnog grada koji se ne baziraju isključivo na sprovođenju standardnih zdravstvenih procedura. Na primer, istraživanjem je ispitana mogućnost korišćenja *LMS* (engl. *Learning Management System*) [6] sistema u pametnom zdravstvu, upotreba GIS-a (geografski informacioni sistemi) [7], upotreba društvenih mreža, upotreba blokčejna (engl. *blockchain*) [8], tehnika istraživanja podataka (engl. *data mining*) i teksta (engl. *text mining*) [9] i procesiranja prirodnog jezika (engl. *NLP – Natural Language Processing*) [10], kraudsoursinga i kraudsenzinga (engl. *crowd-sensing*) [11]. Akcenat je na analizi zahteva potrebnih za kreiranje softverske platforme koja bi omogućila centralizaciju pametnih zdravstvenih servisa koji objedinjuju servise kao što su: upotreba *LMS*-a za edukaciju u cilju prevencije bolesti i komunikacije sa osobljem, servisi za prikaz najbližih zdravstvenih ustanova i apoteka, sistem za odgovaranje na pitanja građana u oblasti zdravstva, info-table za vizuelizaciju informacija o epidemijama, vakcinacijama, zagađenju vazduha, automatsko obrađivanje upitnika koje se može primeniti u obradi anketa i psiholoških testova, ocenjivanju lekara, procesiranje medicinskih podataka u medicinskim informacionim sistemima (npr. automatsko označavanje u anamnezama ili u drugim tekstovima iz medicinskog domena), obezbeđivanju pouzdanosti i trajnosti podataka korišćenjem blokčejna (npr. broj zaraženih ili broj preminulih usled zaražavanja) itd.

Pošto je deo prethodno navedenih servisa zasnovan na obradi tekstualnih podataka na srpskom jeziku, i to onih koji pripadaju posebnom medicinskom domenu, analizirani su i

načini za normalizaciju dokumenata na srpskom jeziku i specifičnosti vezane za normalizaciju tekstova u medicinskim zapisima (dokumentima). Tako, na primer, sistem za odgovaranje na pitanja građana ima mogućnost normalizacije pitanja koje građani postavljaju na srpskom jeziku; sistem za analizu simptoma obrađuje anamneze koje su napisane na srpskom jeziku a sadrže medicinske specifične termine, ali i greške i skraćenice svojstvene lekarima koji ih pišu; za automatsku obradu slobodnog teksta u okviru raznih upitnika i anketa primenjene su takođe i tehnike istraživanja teksta. Za servis koji se bavi označavanjem medicinskih termina kreirana je potpuno nova metoda za označavanje na srpskom jeziku.

Jedan od ciljeva doktorske disertacije je i predlog arhitekture i načina za implementaciju potpuno novog klijent-server informacionog sistema koji realizuje servise pametnog zdravstva, prilagođenih za obradu medicinskih podataka na srpskom jeziku (korišćenje podataka iz elektronskih medicinskih izveštaja, kako strukturiranih tako i nestrukturiranih, što uključuje i obradu slobodnog teksta iz anamneza ili tekstova iz medicinskog domena). Za potrebe realizacije ovog informacionog sistema bilo je potrebno kreirati metode za normalizaciju, označavanje i klasifikaciju medicinskih termina na srpskom jeziku, jer one nisu postojale.

Cilj ove doktorske disertacije je predlog načina i metoda za realizaciju e-servisa pametnog zdravstva koji će biti centralizovani i omogućavati prikaz adekvatnih informacija ili sprovođenje odgovarajućih postupaka i procedura na osnovu medicinskih tekstualnih podataka na srpskom jeziku unetih direktno ili preuzetih iz medicinskih informacionih sistema. Svi ovi servisi bi trebalo da se na uniforman način uklupe u koncept pametnog grada i doprinesu njegovom unapređenju prevashodno u domenu očuvanja zdravlja građana. Načini kolaboracije među servisima i sistemima u okviru pametnog grada prikazani u disertaciji omogućavaju jednostavno uključivanje novih servisa, kao i korišćenje postojećih sistema u svrhe kreiranja servisa pametnog grada/zdravstva.

Rezultati naučnog istraživanja sprovedenog u okviru ove doktorske disertacije su sledeći:

- predloženi načini za uniformnu integraciju heterogenih servisa e-zdravstva u sistem pametnog grada,

- predložena struktura uniformnog veb (engl. *Web*) servisa za realizaciju servisa e-zdravstva u cilju njihove integracije u sistem pametnog grada,
- predlog novih servisa e-zdravstva za uključivanje u koncept pametnog grada (engl. *m(obile)-Health* [12] i *h(ome)-Health* [13]):
  - servis za informisanje, komunikaciju i prevenciju bolesti,
  - servis za pretragu i prikaz najbližih zdravstvenih ustanova,
  - servis za praćenje zagađenja vazduha zasnovan na kraudsenzingu,
  - servisi za unapređenje pružanja medicinskih usluga,
  - servis za inteligentni transport pacijenata,
  - servis za mobilni *SOS* sistem,
  - servis za automatsko odgovaranje građanima na pitanja / pretragu medicinskih dokumenata na srpskom jeziku,
  - servis za kontrolu epidemije, vizuelizacija podataka o epidemijama i stanju vakcinisanosti kod građana,
  - servis za automatsku obradu upitnika / psiholoških testova,
  - servis za uvid u raspored i ocenjivanje rada lekara,
  - servisi za komunikaciju sa izabranim lekarima i dežurnom ekipom hitne pomoći,
  - servis za podršku u slučaju velikih nesreća (saobraćajne nesreće, požari, hemijski incidenti, zemljotresi),
  - servis za efikasniju realizaciju skrining programa,
  - servis za obezbeđivanje pouzdanosti i trajnosti podataka baziran na blokčejnu.
- realizacija pojedinih predloženih servisa e-zdravstva (i integracija u pametni grad),
- obrada medicinskih podataka (npr. tekstovi anamneza) u medicinskim informacionim sistemima kao osnova za realizaciju pojedinih servisa e-zdravstva,
- predlog algoritama za implementaciju pojedinih servisa e-zdravstva sa naglaskom na algoritme veštačke inteligencije (tehnikе obrade prirodnog jezika), istraživanja podataka i teksta i mašinskog učenja (nadgledano, nenadgledano učenje i metode bazirane na pravilima),
- resursi medicinskih i nemedicinskih podataka na srpskom jeziku koji se mogu koristiti za označavanje medicinskog teksta,
- obeleženi korpusi elektronskih medicinskih izveštaja na srpskom jeziku.

## 1.2. Primenjene naučne metode

U izradi predložene doktorske disertacije korišćene su različite istraživačke metode. Metod deskripcije i kvalitativne analize primenjen je na postojeću dostupnu literaturu i metode za kreiranje servisa pametnog zdravstva. Za predlog načina za uniformnu integraciju heterogenih servisa e-zdravstva u sistem pametnog grada i predlog strukture uniformnog veb servisa za realizaciju servisa e-zdravstva u cilju njihove integracije u sistem pametnog grada korišćena je metoda modelovanja. Servisi pametnog zdravstva za obradu medicinskih podataka u medicinskim informacionim sistemima (koji su osnova za realizaciju pojedinih servisa e-zdravstva) bazirani su na metodama nadgledanog i nenadgledanog mašinskog učenja i metodama obrade prirodnog jezika. Takođe je primenjena i eksperimentalna metoda, u slučajevima kada je istraživana efikasnost predloženih metoda za kreiranje servisa pametnog zdravstva koji koriste skup elektronskih medicinskih izveštaja iz medicinskog informacionog sistema. Nad skupom medicinskih izveštaja primenjena je metoda analize sadržaja i validacije.

Komparativna metoda korišćena je za upoređivanje dobijenih eksperimentalnih rezultata sa rezultatima dobijenim upotrebom drugih metoda za obradu podataka iz elektronskih medicinskih izveštaja, uz metode kvantitativne analize. Za evaluaciju poboljšanja predloženih algoritama za implementaciju pojedinih servisa e-zdravstva sa naglaskom na algoritme veštačke inteligencije u odnosu na postojeće metode upotrebljene su statističke metode. Na kraju, metodom sinteze, pojedinačno dobijeni rezultati su objedinjeni u odgovarajuću metodologiju za realizaciju servisa pametnog zdravstva u cilju njihove integracije u sistem pametnog grada.

## 1.3. Pregled disertacije

Doktorska disertacija je organizovana u devet poglavlja. Prvo poglavlje je uvodno i u njemu su prikazani motivacija, predmet istraživanja i postignuti ciljevi disertacije, i dat je pregled primenjenih naučnih metoda u izradi disertacije.

U drugom poglavlju opisan je koncept pametnog grada, kao i karakteristike, infrastruktura, komponente i tehnologije na kojima on počiva.

U trećem poglavlju je stavljen specijalan akcenat na termine elektronsko, mobilno i pametno zdravstvo, na pitanja privatnosti i bezbednosti i pravnu regulativu zdravstvenih usluga u Republici Srbiji. Nakon toga je opisan pregled postojećih servisa pametnog zdravstva koji su realizovani u svetu i kod nas.

U četvrtom poglavlju prikazani su načini integracije heterogenih servisa u okviru koncepta pametnog grada, predlog uniformnog veb servisa na osnovu kojeg su kreirani predloženi i razvijeni servisi pametnog zdravstva i načini kolaboracije postojećih sistema u okviru pametnog grada.

U petom poglavlju sledi prikaz opisa predloženih i realizovanih servisa pametnog zdravstva u okviru istraživanja obuhvaćenog ovom doktorskom disertacijom. U ovom poglavlju razvijeni su: servis za informisanje, komunikaciju i prevenciju bolesti, servis za pretragu i prikaz najbližih zdravstvenih ustanova, servis za praćenje zagađenja vazduha zasnovan na kraudsenzingu, servisi za unapređenje pružanja medicinskih usluga, servis za inteligentni transport pacijenata, servis za mobilni *SOS* sistem, servis za automatsko odgovaranje građanima na pitanja / pretraga medicinskih dokumenata na srpskom jeziku, servis za kontrolu epidemije, vizuelizacija podataka o epidemijama i stanju vakcinisanosti kod građana, servis za automatsku obradu upitnika / psiholoških testova, servis za uvid u raspored i ocenjivanje lekara, servisi za komunikaciju sa izabranim lekarima i dežurnom ekipom hitne pomoći, servis za podršku u slučaju velikih nesreća (saobraćajne nesreće, požari, hemijski incidenti, zemljotresi), servis za efikasniju realizaciju skrining programa i servis za obezbeđivanje pouzdanosti i trajnosti podataka baziran na blokčejnu. U ovom poglavlju dat je detaljan pregled prethodno pobrojanih servisa pametnog zdravstva.

Šesto poglavlje opisuje obradu podataka u medicinskim informacionim sistemima kao osnovu za realizaciju pojedinih servisa e-zdravstva. Ovde je dat pregled postojećih jezičkih medicinskih resursa dostupnih u svetu i kod nas. Takođe, ovo poglavlje sadrži pregled tehnika za istraživanje podataka. Zatim je opisan način za kreiranje medicinskih leksičkih resursa, i dat je opis korišćenih leksičkih resursa.

Metode za obeležavanje termina u elektronskim medicinskim izveštajima, bazirane na algoritmima veštačke inteligencije do kojih se došlo u procesu istraživanja, a neophodni su za realizaciju predloženih servisa pametnog zdravlja, opisane su u sedmom poglavlju. Ovde su dati motivacija za primenu metode, opis skupa podataka, potrebna normalizacija, struktura i funkcionisanje metoda, i metrike neophodne za evaluaciju rada metoda. Testiranje i rezultati primene metode za označavanje medicinskih termina prikazani su takođe u ovom poglavlju. Testirana su tri pristupa:

- obeležavanje termina metodom koja se zasniva na medicinskim leksičkim resursima,
- obeležavanje termina metodom nadgledanog učenja,
- obeležavanje termina hibridnom metodom, koja je zasnovana na nadgledanom učenju i pravilima.

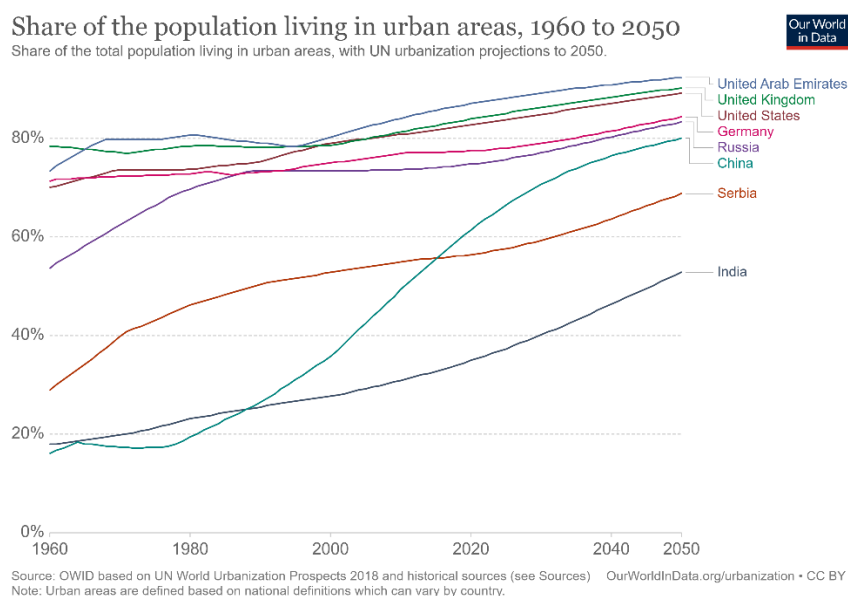
Takođe su dati i rezultati primene postojećih metoda mašinskog učenja i alata za označavanje (engl. tagger) nad opisanim skupom podataka.

Osmo poglavlje je zaključak, a deveto poglavlje daje opis pravaca daljeg istraživanja i razvoja sa predlozima budućeg unapređenja i proširenja metoda.



## 2. PAMETNI GRADOVI

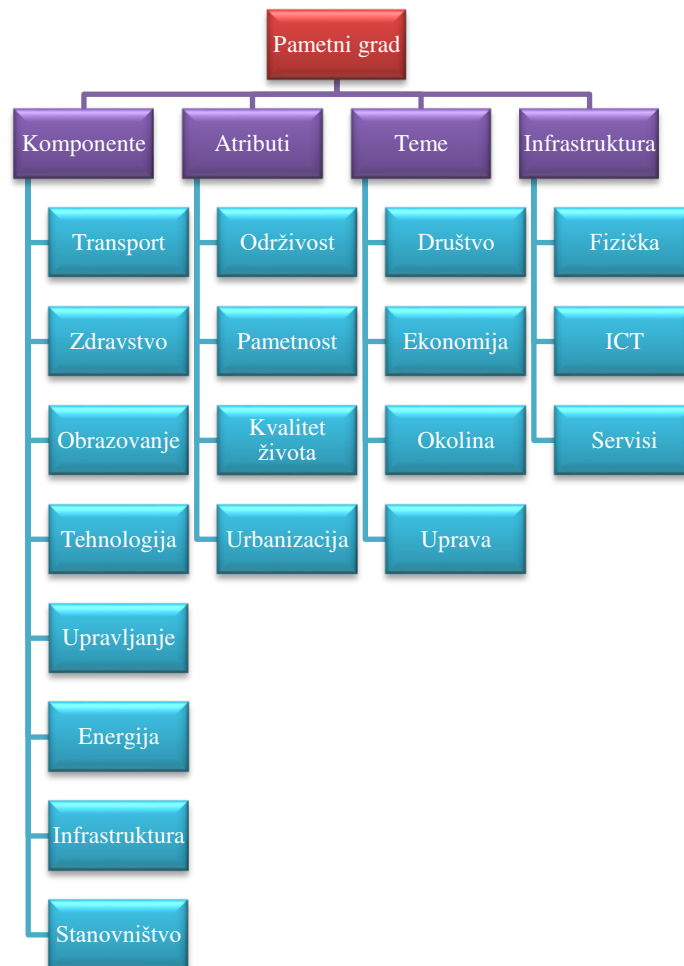
Stalni napredak informaciono komunikacionih tehnologija (IKT) omogućio je njihovu primenu u gotovo svim segmentima života pojedinaca, sa ciljem pružanja što većeg broja kvalitetnih servisa koji korisnicima olakšavaju život. Savremena tehnološka infrastruktura (Internet, pametni telefoni sa ugrađenim sensorima, pametni uređaji u domaćinstvu) dostupna je danas većinskom delu stanovništva. Sve to rezultira mogućnošću prikupljanja velikog broja podataka o kontekstu korisnika [14]. Tehnike za obradu velike količine prikupljenih podataka, korišćenje kolektivne inteligencije i povećanje broja stanovnika u urbanim područjima (Slika 1) omogućile su stvaranje novog okruženja za stanovnike gradova, koje je efikasno u pogledu potrošnje energije, transporta, uprave, učenja, ekonomskog razvoja i drugih sličnih aspekata.



Slika 1. Procenat stanovništva u gradovima (preuzeto iz [15])

Korišćenje IKT u gradovima za različite gradske aktivnosti dovelo je do povećane efikasnosti obavljanja istih, a za označavanje ovakvih gradova koristili su se brojni pojmovi kao što su „digitalni grad“, „elektronski grad“, „umreženi grad“ i „pametani grad“. Poslednji od ovih termina se među njima posebno izdvojio i označava mesto gde tradicionalne usluge

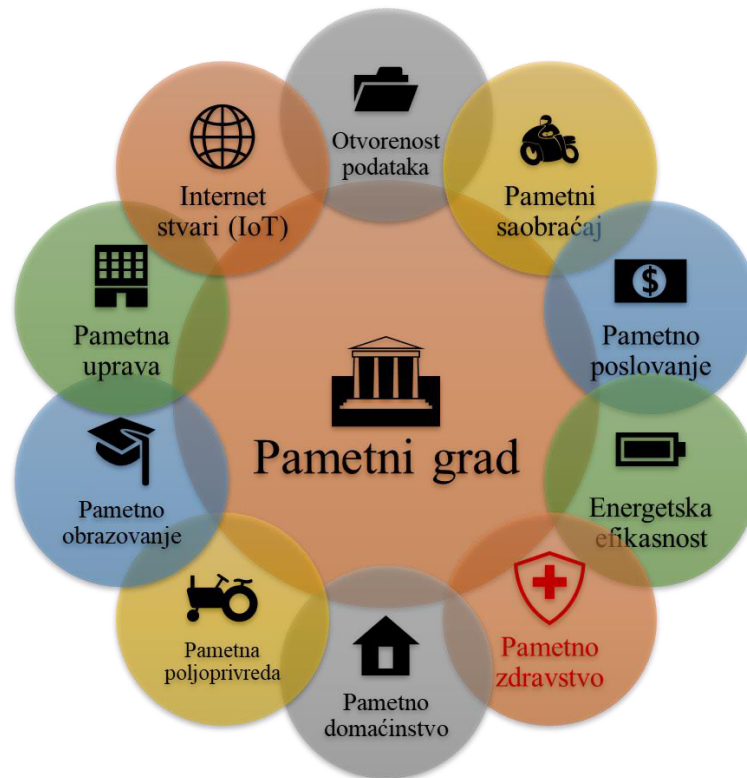
postaju fleksibilnije i efikasnije koristeći pri tome informacione, digitalne i telekomunikacione tehnologije [1]. U pametnom gradu digitalne tehnologije omogućavaju kvalitetnije javne usluge za stanovnike i bolje iskorišćenje resursa. Na slici Slika 2 prikazane su komponente, atributi, teme i infrastruktura pametnih gradova.



Slika 2. Komponente, atributi, teme i infrastruktura pametnih gradova

## 2.1. Komponente pametnog grada

Postoji veliki broj komponenti na kojima su zasnovani pametni gradovi, a neke od njih su pametni transport, energetska efikasnost, pametna tehnologija i infrastruktura, pametno obrazovanje, pametno zdravstvo i pametno upravljanje. Neke od njih prikazane su na slici Slika 3.



Slika 3. Komponente pametnih gradova

Atributi pametnih gradova su održivost, kvalitet života, urbanizacija i „pametnost“. Održivost pametnog grada odnosi se na gradsku infrastrukturu i upravljanje, energiju i klimatske promene, zagađenje, otpad, socijalna pitanja, ekonomiju i zdravlje. Kvalitet života se može meriti u smislu emocionalnog i finansijskog blagostanja građana. Urbanizacija se odnosi na tehnologije, infrastrukturu, upravljanje i ekonomiju dok „pametnost“ podrazumeva pametnu ekonomiju, „pametne ljude“, pametno upravljanje, pametnu mobilnost i pametan život.

Postoje četiri centralne teme u pametnim gradovima: društvo, ekonomija, životna sredina i upravljanje. Tema društva odnosi se na to da grad postoji za dobrobit svojih stanovnika ili građana. Tema privrede pametnog grada odnosi se na uspehe grada kroz stalan rast zaposlenosti i ekonomski rast. Tema životne sredine ukazuje da će grad moći da održi svoju funkciju i da ostaje u funkciji za sadašnje i buduće generacije. Tema upravljanja pametnim gradom ukazuje na to da je grad robusan u svojoj sposobnosti upravljanja politikom i kombinovanja ostalih elemenata [2].

## 2.2. Infrastruktura pametnog grada

Infrastruktura pametnog grada obuhvata fizičke, informacione i komunikacione tehnologije i usluge. Fizička infrastruktura je stvarna fizička ili strukturalna celina pametnog grada koja uključuje objekte, puteve, železničke pruge, sisteme snabdevanja strujom, vodom, gasom itd. IKT infrastruktura je osnovna inteligentna komponenta pametnog grada koja povezuje sve ostale komponente. Servisna infrastruktura bazirana je na fizičkoj infrastrukturi i može imati neke IKT komponente [2].

### 2.2.1. Tehnologije na kojima počiva pametni grad

IKT infrastruktura pametnih gradova bazira se na dvema blisko povezanim i novim tehnologijama, Internet stvari (*IoT*) [16] i „veliki podaci“ (engl. *Big Data*) [17].

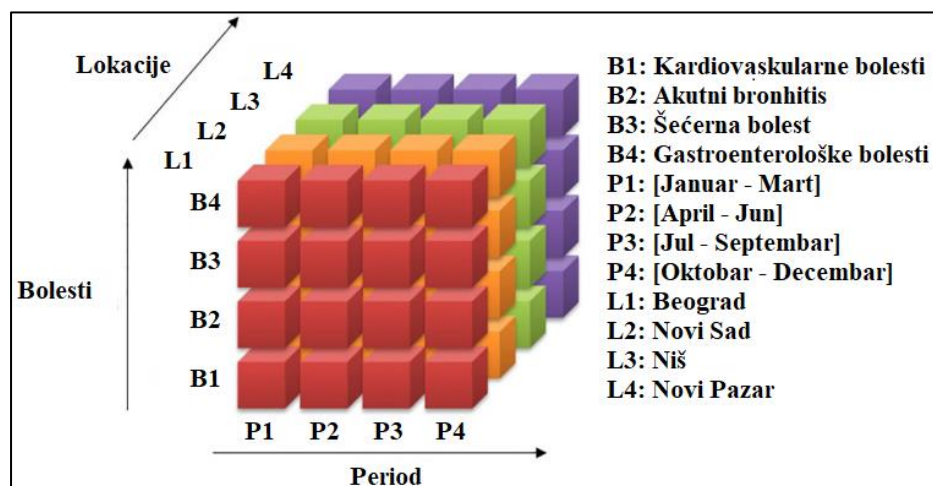
Internet stvari čine stvari (senzor, uređaj povezan na mrežu s mogućnošću interakcije s korisnikom ili upravljanja drugim uređajem), lokalna mreža, internet i klad (engl. *cloud*). Internet stvari predstavlja umrežavanje fizičkih uređaja, vozila i zgrada korišćenjem ugrađenih senzora sa drugim uređajima. Sa razvojem senzorskih tehnologija, značajno je povećan broj konektovanih uređaja sposobnih da prikupljaju informacije i komuniciraju sa okolinom u koju su ugrađeni. Ovakvi uređaji omogućili su razvoj nove vrste aplikacija koje mogu učiniti gradove inteligentnijim, unapređujući gradsku IKT infrastrukturu i samim tim svakodnevni život građana. Kao rezultat interakcije između više aplikacija, formira se veliki broj podataka iz kojih se može izvući korisno znanje [4]. *IoT* uređaji proširuju opseg korišćenja aplikacija, povećavajući mogućnosti za projektovanje inteligentnih gradova, a mogu biti primenjene u domenu zdravstva, pametnog saobraćaja, komunalne infrastrukture, administracije i sl.

Pojmom „veliki podaci“ označavaju se velike količine podataka prikupljene sa senzora, kraudsorsingom i sl. a čije su smeštanje i obrada kompleksniji od tradicionalnih metoda [2][17]. Njih karakteriše raznovrsnost formata, velika brzina pristupa i obrade i veliki obim informacija, pa samim tim projektovanje servisa za pristup ovim podacima, pretragu,

deljenje, njihovu obradu, analizu i vizuelizaciju predstavlja veliki izazov. Obično su ovi podaci smešteni u nerelacione baze podataka [18].

Osnova IKT infrastrukture pametnog grada je zapravo u kolaboraciji različitih sistema, koji su obično realizovani nezavisno, korišćenjem različitih tehnologija i u različita vremena. Potrebno je omogućiti da svi ti postojeći sistemi budu izvori podataka za nove servise pametnog grada. Postavlja se pitanje kako obezbediti komunikaciju ovih sistema.

Rešavanje ovog problema ne zahteva potpunu promenu ili ponovnu izgradnju postojećih sistema, već je samo potrebno obezbediti integraciju i kolaboraciju koja se često odnosi samo na nivo podataka (nekad direktno, a nekad indirektno). Heterogeni sistemi na ovaj način postaju izvori podataka za skladišta podataka, gde se posle mogu npr. kreirati kubovi i primeniti analiza i prikaz podataka (Slika 4), kao što je to slučaj kod sistema poslovne inteligencije, ili *BI* sistema (engl. *business intelligence*) [19] [20]. Ovi sistemi namenjeni su praćenju performansi poslovnih procesa kroz tačnu prezentaciju i analizu višedimenzionalnih podataka (kubovi) preuzetih iz distribuiranih sistema za obradu transakcija. Sastoje se od sedam slojeva: IKT infrastruktura, prikupljanje podataka, integracija podataka, skladištenje podataka, organizovanje podataka, analitika podataka i prezentacija podataka.



Slika 4. Primer čuvanja zdravstvenih podataka u kubovima

Detaljna analiza i predlozi komunikacije različitih podsistema u okviru pametnog grada opisani su u četvrtom poglavlju, kao i tehnologije koje mogu biti osnova za kreiranje servisa pametnog grada i zdravstva.

### 3. PAMETNO ZDRAVSTVO

Grana koja je posvećena zdravlju stanovnika u okviru koncepta pametnog grada je pametno zdravstvo. Nezavisno od koncepta pametnog grada, IKT već duže vreme nalaze svoju primenu u zdravstvenom sistemu, pa se često susrećemo sa terminima e-zdravstvo i m-zdravstvo [3]. Nešto ređe se sreće i termin h-zdravstvo (engl. *home-Health*)[3][21]. E-zdravstvo (engl. *e-Health*) je termin nastao krajem prošlog veka i zasniva se na inovativnom, efikasnijem i kvalitetnijem pružanju zdravstvenih usluga uz pomoć savremenih tehnologija i to uz vrlo visok stepen integracije sa drugim postojećim sistema, i mogućnost za obezbeđivanje mobilnosti lekara i pacijenata. Termin m-zdravstvo obuhvata niz usluga u zdravstvu koje se ostvaruju kombinovanjem medicinskih tehnologija i mobilnog i sveprisutnog računarstva. Kako su kapaciteti zdravstvenih ustanova sve kritičniji, u poslednje vreme se značajan napor ulaže na pružanju odgovarajućih usluga kod kuće bez potrebe za dolaskom u zdravstvenu ustanovu, što se označava kao h-zdravstvo.

Primeri zastupljenosti IKT u sistemima zdravstvene zaštite ogledaju se kroz: kompletno elektronsko vođenje zdravstvene dokumentacije u svim segmentima rada (elektronski karton pacijenta), telemedicinu, telekonsultacije, telehirurgiju, upravljanje znanjem o zdravlju, mobilno i sveprisutno zdravstvo. Ovo omogućava lakše i uspešnije lečenje pacijenata i smanjenje administrativnih ograničenja prilikom pružanja medicinskih usluga [22].

Svrha primene e-zdravstva u sistemima zdravstvene zaštite pacijenata jeste omogućavanje efikasnijeg funkcionisanja klasičnih zdravstvenih sistema kroz korišćenje IKT procesa. Prednosti e-zdravstva su elektronsko praćenje i beleženje zdravstvenog stanja pacijenata, bolja dijagnostika, pristup medicinskim podacima bilo gde i bilo kada, kao i brz prenos informacija korisnicima putem telemedicine i Internet servisa.

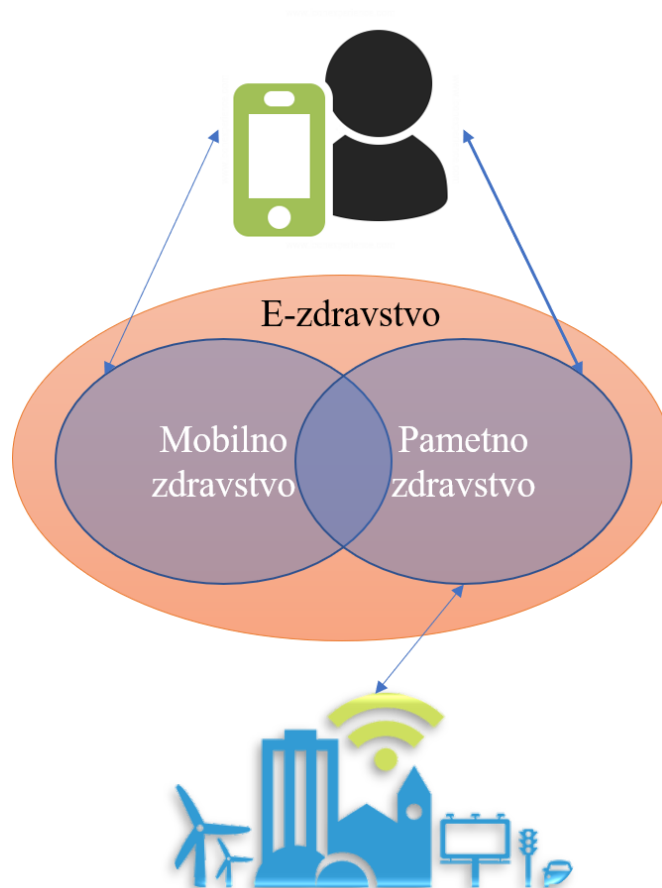
Radovi i istraživanja iz ove oblasti opisuju veći broj servisa pametnog zdravstva koji pripadaju jednoj od sledećih kategorija:

- **Usluge zasnovane na kolektivnoj inteligenciji grada**, koje koriste podatke prikupljene sa senzora koji se nalaze na teritoriji pametnog grada i analiziraju dobijene podatke za predviđanja u različite svrhe. Primer ovakvog servisa je na primer aplikacija za predlaganje puta (rute) sa manjim stepenom zagađenosti vazduha za korisnike koji imaju problem sa respiratornim organima ili naprosto ruta kojom se najbrže stiže na odredište.
- **Kontekstno-orijentisani servisi**, koji analiziraju blisko okruženje korisnika na osnovu slike, zvuka, detekcije pokreta, itd. Primeri ovih servisa su softveri za obradu i analizu slike u cilju otkrivanja bolesti i abnormalnih pojava (npr. analiza slike za otkrivanje bolesti oka, kancera), kao i softveri za detekciju pokreta koji se koriste za određivanje emocija, nivoa stresa, disanja.
- **Usluge zasnovane na IoT uređajima**, koje analiziraju usko korisničko okruženje pomoću nosivih IoT uređaja. Primeri su uređaji za monitoring šećera u krvi, holter krvnog pritiska, EKG-a i temperature.
- **Pametne kuće za pacijente** su pametno okruženje za brigu najčešće o starijim i hroničnim bolesnicima, a podrazumevaju upotrebu ugrađenih senzora, kamera i mobilnih uređaja (koji su fiksno postavljeni u prostoru ili ih pacijenti nose na sebi) i veb dostupnih usluga.
- **Usluge zasnovane na kraudsorsingu i istraživanju medicinskih podataka** koje su zasnovane na analizi velikog broja originalnih informacija iz šireg područja, koje mogu biti i izvan teritorije grada [23].

### 3.1. Odnos elektronskog, mobilnog i pametnog zdravstva

U eri analiziranja podataka o korisnicima računarskih sistema ništa se više ne pretpostavlja, već se čine svi mogući koraci da se dođe do što preciznijih informacija o korisniku i njegovom okruženju. Svaki, nekome naizgled beznačajan podatak, može doprineti nekom novom saznanju i olakšati interakciju sa korisnikom. Ti kontekstni podaci mogu se iskoristiti u svrhu poboljšanja zdravstvenog stanja korisnika aplikacije.

Primena mobilnog i sveprisutnog računarstva učinila je da se prikupljanju podataka iz okruženja korisnika posvećuje sve veća pažnja. Ovi podaci čine kontekst, a aplikacije koje koriste ove podatke nazivaju se kontekstno-svesne aplikacije [14]. Korišćenje pametnih telefona u svrhe poboljšanja zdravstvenog stanja korisnika dovelo je do formiranja nove oblasti u okviru elektronskog zdravlja, a to je mobilno zdravstvo (engl. *m-Health*).



Slika 5. Odnos mobilnog i pametnog zdravstva

Prirodna sinergija mobilnog i elektronskog zdravstva sa konceptom pametnih gradova iznedrila je novi termin – pametno zdravstvo, skraćeno p-zdravstvo (engl. *smart health, s-Health*). Ovaj termin se prvi put pominje u radu [3], gde autori predstavljaju zdravlje i pametne gradove zasebnim ravnima, pri čemu su m-zdravstvo i p-zdravstvo podskupovi e-zdravstva, ali imaju i svoj presek. Definicija p-zdravstva slična je definiciji m-zdravstva, s tim što je akcenat na tome da se ono upotrebljava u okviru pametnih gradova. Pored toga, p-zdravstvo i m-zdravstvo razlikuju se u izvornim informacijama koje koriste, i toku tih informacija. Za m-zdravstvo izvorne informacije potiču od korisnika/pacijenata, dok kod p-



zdravstva, pored tih informacija koriste se i kolektivni podaci dobijeni od infrastrukture pametnog grada. Što se tiče razlike u toku podataka, ona se odnosi na to da se nakon obrade kod m-zdravstva povratni podaci vraćaju nazad korisniku, a kod p-zdravstva i korisniku, ali i utiču na kolektivne podatke pametnog grada (Slika 5). Pametno zdravstvo predstavlja jednu komponentu u okviru pametnih gradova, i u stvari obuhvata elektronsko zdravstvo i njegove aspekte obavljane u svrhu unapređenja delatnosti stanovništva u pametnim gradovima.

### **3.2. Pitanja bezbednosti i privatnosti zdravstvenih usluga**

Na osnovu gore pomenutih pametnih zdravstvenih usluga, može se uočiti da prikupljanje velike količine informacija može biti od velike pomoći pacijentu kroz različite servise zasnovane na ovim podacima. Nasuprot tome, ovi podaci se mogu zloupotrebiti. Što se tiče ličnih podataka, iz elektronskih izveštaja pacijenata moguće je dobiti informacije kao što su navike pacijenta, socijalni status i, naravno, njegovo zdravstveno stanje. Na primer, u kartici zdravstvenog osiguranja postoje podaci o zanimanju i zaposlenju. U malim mestima je čak i na osnovu ovih podataka moguće otkriti o kome se radi.

Ako su korisnički podaci šifrovani, iz njih se ne mogu izvući novi zaključci. Ako ostanu nepromenjeni, oni mogu poslužiti za sticanje novih znanja, ali privatnost podataka je neizvesna. Postoji nekoliko metoda za rešavanje ovog problema, od kojih su najvažnije:

1) Statistička kontrola objavljivanja - Tabelarni podaci prikupljeni na pacijentima nazivaju se i mikropodaci. Ovom metodom se vrši promena mikropodataka, kako bi se obezbedio minimalan gubitak informacija, a istovremeno smanjila mogućnost ponovnog identifikovanja korisnika na osnovu podataka [24].

2) Model privatnosti *W3* omogućava zaštitu privatnosti kada se koriste usluge zasnovane na lokaciji (*LB* – engl. *location based*). Budući da takve usluge dobijaju informacije da neko sa određene lokacije nešto traži, neophodno je onemogućiti praćenje ovih ličnih podataka, već samo dati odgovore na pitanje [25]. Ovaj model se takođe može primeniti na *LB* usluge u svrhu zdravstvene zaštite.

3) Model 5D privatnosti [25] zasnovan je na sledećih pet dimenzija: privatnost identiteta, privatnost upita, privatnost lokacije, privatnost otiska i privatnost vlasnika. Ovaj model je

produžetak prethodnog modela za poslednje dve dimenzije. Privatnost otisaka odnosi se na zaštitu privatnosti kada se koristi odgovarajuća tehnologija za prikupljanje podataka sa senzora, tako da se oni ne daju trećoj osobi i ne zloupotrebljavaju. Privatnost vlasnika odnosi se na privatnost korisničkih podataka koje različite gradske službe imaju o korisniku [26].

Predloženi modeli za obezbeđivanje sigurnosti i privatnosti podataka su aktuelni za servise u okviru pametnog zdravstva u svetu. U narednom poglavlju dat je opis zakonskog okvira koji se bavi ovim problemom koji je na snazi u Republici Srbiji i zemljama u okruženju. Ovde su sadržani još neki opisi i predlozi za obezbeđivanje sigurnosti podataka građana u skladu sa opštom odredbom o zaštiti podataka (u daljem tekstu *GDPR*, engl. *General Data Protection Regulation*) [27].

### **3.3. Pravna regulativa elektronskog poslovanja u Republici Srbiji i osvrt na *GDPR***

Strategija razvoja elektronskog poslovanja u Republici Srbiji sadržana je u Strategiji razvoja informacionog društva u Republici Srbiji do 2020. godine, i u Nacionalnoj strategiji za održivi razvoj i Strategiji razvoja elektronskog poslovanja u Republici Srbiji za period od 2009. do 2013. godine. Strategija informacionog društva Republike Srbije do 2020. godine koja pruža reformu i modernizaciju javne uprave na osnovu široko rasprostranjene upotrebe IKT tehnologija jedan je od ključnih elemenata sveukupne tranzicije Republike Srbije u moderno informaciono društvo. Prema Nacionalnoj strategiji održivog razvoja u Republici Srbiji neophodno je bilo podržati povećanje nivoa digitalne pismenosti među građanima i uvođenje kvalitetnog obrazovanja u smislu informaciono komunikacionih tehnologija na nivou osnovne škole. Trebalo je obezbediti uslove za dalje povećanje broja korisnika interneta i pristupa IKT svima, fizičkim i pravnim licima.

Zakoni Republike Srbije koji su usvojeni proteklih godina i koji su u skladu sa Evropskom unijom (Zakon o slobodnom pristupu informacijama od javnog značaja i Zakon o elektronskom dokumentu, elektronskoj identifikaciji i uslugama od poverenja u elektronskom poslovanju i relevantni podzakonski akti), sastoje se zakonskog regulisanja koncepta osnovnih elemenata e-poslovanja. Primeri ovih koncepata su uvođenje elektronskih potpisa i digitalnih

sertifikata, elektronskog pružanja usluga klijenata putem Interneta, komunikacija korisnika i autoriteta meglom, sankcionisanje nemarnih i zlonamernih dela itd.

Među zakonskim odredbama Republike Srbije najznačajnija je primena Zakona o elektronskim potpisima (usvojena u decembru 2004). Na ovaj način, građanima Republike Srbije omogućeno je da u poslovnicama Pošte Srbije podnesu zahtev za aparat koji proizvodi elektronski potpis. Nakon popunjavanja i potpisivanja obrasca, građanima se u roku od dva ili tri dana na kućnu adresu isporučuju sertifikat, kartica, čitač kartica i softver koji moraju biti instalirani na računaru korisnika. Odvojeno, radi sprečavanja zloupotrebe, šalje se lični identifikacioni broj, nazvan pin, koji se otkucava u softveru kada se potpisivanje dokumenata vrši elektronskim potpisom, kao što su npr. dokumenti kreirani u paketu *Microsoft Office*. Srpska pošta ima široku infrastrukturu od oko 1.200 automatizovanih i povezanih ogranaka u realnom vremenu. Upotreba elektronskih potpisa trebalo bi da ispuní osnovni pravni preduslov za prelazak sa papira na elektronske procese u svim oblastima života, kako bi bio postignut jedan od osnovnih preduslova za dalji razvoj e-poslovanja u Republici Srbiji [18].

Uredba Evropske unije o zaštiti podataka o ličnosti (engl. *GDPR – General Data Protection Regulation*) [28] primenjuje se neposredno u svim zemljama članicama Evropske unije počevši od 25. maja 2018. godine. Ova regulativa definiše propise o zaštiti podataka ličnosti u Evropskoj uniji, ali se ona odnosi i na kompanije sa sedištem u zemljama izvan Evropske unije, ukoliko one obrađuju podatke o ličnosti rezidenata Evropske unije. Po članu 3 *GDPR* regulative [27] potrebno je da se uskladi zaštita osnovnih prava fizičkih lica u vezi sa aktivnostima obrade ličnih podataka. U Srbiji je na snazi Zakon o zaštiti podataka o ličnosti iz 2018. godine koji je usklađen sa *GDPR* [29].

*GDPR* regulativa obezbeđuje pojedincima veću kontrolu nad ličnim podacima i obavezuje kompanije koje prikupljaju i analiziraju lične podatke da promene svoj model poslovanja u skladu s njom. Pod privatnim podacima podrazumevaju se podaci iz kojih se može utvrditi identitet lica, zatim podaci o političkoj, seksualnoj orijentaciji, rasi, imovnom stanju, ali i podaci kao što su istorija pretraga, meta podaci na fajlovima koje je napravilo fizičko lice, a koji mogu da otkriju identitet, podaci o zdravlju, kretanju, navikama. Pod ličnim podacima podrazumeva se bilo koja kombinacija ličnih činjenica koja tačno određuje jednog pojedinca, to su, između ostalog, ime i prezime, JMBG, podaci o lokaciji, fizički, fiziološki, genetski, mentalni, ekonomski, socijalni, kulturni ili bilo koji drugi faktori.

Principi *GDPR* regulative su sledeći:

- potrebna je zakonita i transparentna obrada ličnih podataka građana,
- prikupljanje ličnih podataka u posebne, tačno određene svrhe i čuvanje samo minimalno potrebnih podataka,
- kompanija nema ovlašćenje da od pojedinca traži dodatne informacije koje izlaze van definisane svrhe prikupljanja i obrade podataka,
- kompanija nema ovlašćenje da čuva podatke duže nego što je definisano,
- potrebno je da kompanija implementira određene sigurnosne mere,
- lični podaci moraju biti tačni i ažurirani,
- osiguravanje odgovarajućeg nivoa sigurnosti ličnih podataka uključujući zaštitu od neovlašćene obrade i poverljivost podataka [30].

Kompanija koja prikuplja lične podatke pre obrade mora dobiti saglasnost pojedinca za njihovo korišćenje i neophodno je da fizičko lice koje daje saglasnost bude obavešteno o svrsi prikupljanja ličnih podataka [31].

Pojedinac ima pravo da od kompanije zatraži detaljne informacije o tome kako čuva lične podatke. Prava pojedinca su jasno definisana *GDPR* regulativom i dele se na:

- Pravo pristupa – pojedinci mogu tražiti da znaju u koju svrhu se lični podaci obrađuju, mogu da traže uvid u podatke, i da im se predstave metode koje kompanija koristi za proces čuvanja podataka.
- Pravo na brisanje – pojedinac može da zatraži brisanje ličnih podataka koje kompanija vodi o njemu. Međutim, neki podaci se ne mogu brisati, jer su vezani nekim drugim zakonskim aktima.
- Pravo na ispravku – pojedinac može da traži od kompanije da se njegovi netačni podaci isprave.
- Pravo na prenosivost – pojedinci mogu tražiti da kompanija koja obrađuje podatke prenese lične podatke, u sigurnom obliku i sigurnim putem, trećim licima.
- Povreda privatnosti – ukoliko dođe do “curenja” ličnih podataka, kompanija ima obavezu da u roku od 72 sata od saznanja o incidentu, obavesti fizičko lice.

Pojedinac pre ostvarenja svojih prava, mora potpisati saglasnost koja omogućava da kompanija koristi njegove lične podatke. Međutim u slučaju da postoji zakon koji obavezuje davanje podataka ili ugovor između kompanije i pojedinca, to podrazumeva da je fizičko lice već pristalo na korišćenje određenih podataka i u tim slučajevima kompanija ne mora da pribavlja saglasnost pojedinca. Isto važi i u slučaju da se podaci nalaze negde javno objavljeni.

U tabeli Tabela 1 navedene su aktivnosti nad podacima koje je potrebno uskladiti sa *GDPR*, problemi koje je potrebno u okviru tih aktivnosti rešiti, kao i predlozi rešenja navedenih problema. [30].

Tabela 1. Usklađivanje aktivnosti nad podacima sa *GDPR*

Aktivnost	Problem koji je potrebno rešiti	Predlog rešenja
Pretraga	Strukturirani i nestrukturirani podaci	Mašinsko učenje
	Podaci na serveru	
	Podaci na klauđu	
	Mejlovi	
Upravljanje	Prenos ličnih podataka	Klasifikacija podataka
	Označavanje medija koji su nosioci podataka	
Zaštita ličnih podataka	Zavisno od vrste korišćenja	Šifriranje ili maskiranje podataka
	Zavisno o pretnji zloupotrebe	
Nadzor i izveštavanje	Sistemske zapise i log fajlovi	Istraživanje procesa
	Curenje poverljivih podataka	

Prva i osnovna faza implementacije *GDPR* uredbe je faza pretrage postojećih ličnih podataka. Potrebno je omogućiti pretragu kompletnog opsega ličnih podataka, odrediti identifikatore koji će jasno definisati lični podatak (ime, prezime, email, adresa, telefon, JMBG). Zatim sledi podela podataka na strukturirane (tabela sa označenim ličnim podacima) i nestrukturirane (lični podaci u slobodnom tekstu). Alat za pretraživanje treba da omogući pretraživanje fotografija, podatke sa društvenih mreža, razne onlajn identifikatore: kao što su korisničke prijave, *MAC* podaci, korisnikove geolokacije. Predlaže se korišćenje sistema za

mašinsko učenje (engl. *machine learning*), koji će moći iz velikog broja dokumenata i ugovora prepoznati šta je lični podatak.

Kada se podaci odrede, tada ih je potrebno klasifikovati. Lične podatke moguće je označiti po vrsti (strukturirane/nestrukturirane) i po tajnosti (javni podaci, tajni podaci). Svaki korisnik baze ličnih podataka treba da dobije pristup ovakvom alatu. Podaci (datoteka, dokument, fotografija, video, audio, mejl) se klasifikuju ručno ili pomoću alata. Klasifikacija podataka omogućiti će da se odredi da li je podatak uopšte lični podatak, da li mu je potrebna enkripcija, digitalni potpis i slično.

Tabela 2. Zaštita podataka šifriranjem i maskiranjem [34]

Šifriranje	Maskiranje
Promena ličnih podataka u oblik nerazumljiv neovlašćenim licima	Promena ličnih podataka na način oduzimanja atributa identifikacije određene osobe
Zaštita putem transformacije	Zaštita putem zamene
Reverzibilnost je neophodna	Reverzibilnost može biti potrebna
Matematički algoritam + ključ	Bez matematičkog algoritma
Neupotrebljivi podaci	Podaci upotrebljivi
Promena veličine podataka	Bez promene veličine podataka
Neupotrebljivo pre dešifrovanja	Pogodno za dalju obradu (testiranje, analiza)
Korišćenje: zaštita fajlova, mrežni promet	Korišćenje: kada je potrebna dalja obrada podataka

Za zaštitu podataka uredba koristi pojmove šifriranja i maskiranja podataka. Metode se razlikuju zavisno o tome šta želimo da radimo sa podacima nakon primene jedne od ovih metoda. Karakteristike ovih metoda prikazane su u tabeli Tabela 2. Šifriranje će pretvoriti podatke u nečitljivi oblik i kao takav, podatak se više neće moći koristiti. Maskiranje kao metoda vrši zamenu podatka koji ostaje čitljiv, ali je netačan. I jednom i drugom metodom postiže se gubljenje sposobnosti podatka da omogućiti identifikovanje neke osobe i samim tim podaci više ne podležu *GDPR* regulativi. Maskiranje se koristi kada se želi obavljati dalja obrada nad ličnim podacima. Podaci će biti prepoznatljivi kao lični podaci, ali neće biti tačni.

Za nadzor i izveštavanje koristi se istraživanje procesa (engl. process mining) koji može da se primeni nad zapisima događaja (engl. event logs) iz poslovanja informacionih sistema [32]. Istraživanje procesa podrazumeva automatizovano otkrivanje neusklađenih procesa, proveru odstupanja od uobičajenog ponašanja korisnika, proširivanje i popravku tih modela [33][34].

### 3.4. Elektronski zdravstveni servisi u svetu i kod nas

Implementirane servise u okviru pametnog zdravstva možemo svrstati u nekoliko kategorija u zavisnosti od načina na koji se dolazi do potrebnih informacija za analizu zdravstvenog stanja pacijenta.

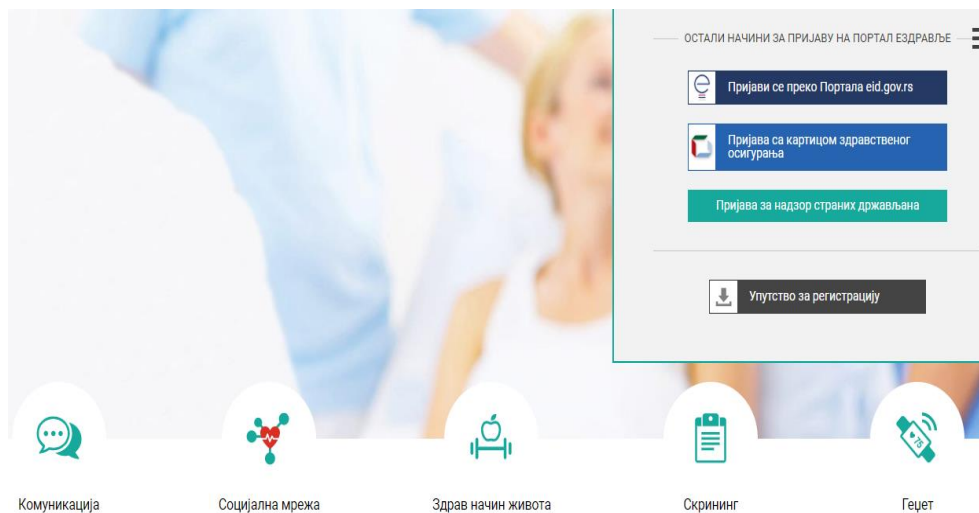
U prvu grupu spadaju mobilne aplikacije koje koriste kolektivne podatke u okviru infrastrukture pametnih gradova. Takvi servisi opisani su u radovima [35][36], a radi se o aplikacijama koje korisnicima koji imaju problem sa respiratornim organima predlažu rute gde je zagađenost vazduha manja.

U sledeću grupu servisa pametnog zdravstva spadaju usluge koje se ostvaruju preko raznih vrsta nosivih uređaja, a korišćenjem *IoT* i bežičnih mreža, neretko i *M2M* komunikacije. Primeri ovih servisa opisani su u radovima u nastavku. U radu [37] opisani su nosivi uređaji u sklopu *RFID-WSN* hibridnog sistema koje omogućuju praćenje lokacije i praćenje pacijenta u realnom vremenu. U radu [38] opisani su servisi za merenje šećera u krvi, elektrokardiograma srca, krvnog pritiska i temperature uz korišćenje *IoT*. Značaj *IoT* u implementaciji zdravstvenih usluga u pametnim gradovima opisan je i u radovima [39], [40]. U radu [41] opisana je integracija *IoT* i klad servisa u cilju određivanja bolesti na osnovu simptoma koji se ogledaju u promenama u glasu pacijenata.

U radu [13] dat je pregled tehničkih, psihološko-etičkih i ekonomskih izazova za formiranje pametnih okruženja (kuća) za brigu o starima i o hroničnim bolesnicima, koje koriste tehnološku infrastrukturu (senzore, kamere, nosive uređaje i veb servise) u cilju da se ispoštuje želja pacijenata da budu kod kuće. O pametnim kućama namenjenih zdravstvenoj nezi pisano je i u radovima [42][43].

U svrhe pametnog zdravstva koristi se i obrada i analiza slike za slučaj detektovanja abnormalnih pojava za odgovarajuće bolesti [44]. Detekcija pokreta koristi se u svrhe određivanja emocija, nivoa stresa, praćenje disanja [45]. Istraživanje podataka i kraudsorsing takođe su našle primenu u servisima pametnog zdravstva. U radu [46] prikazan je način postupanja sa donošenjem zaključaka u slučajevima merenja i dobijanja abnormalnih rezultata, dok je u radu [5] opisano detektovanje nivoa depresije kroz kraudsorsing poređenjem sa podacima drugih korisnika i korišćenjem medicinskog istraživanja podataka.

U Republici Srbiji u svojstvu elektronskog zdravstva koje bi moglo biti primenjivo i u realizaciji pametnih gradova, možemo izdvojiti javno dostupne informacije o zdravstvu. Primer takvih servisa jesu veb portali *E-Zdravlje*[47] i *MojDoktor* [47] koji je obezbedilo Ministarstvo zdravlja Republike Srbije. Ovo je deo integrisanog zdravstvenog informacionog sistema Republike Srbije u kome se čuvaju i obrađuju svi medicinski i zdravstveni podaci pacijenata, medicinskih radnika, podaci o ustanovama i intervencijama i uslugama izvršenih od strane tih zdravstvenih ustanova, podaci o receptima, uputima na specijalističke preglede i sl. Na ovaj način mogu se zakazati pregledi elektronskim putem, dobiti uput za specijalistički pregled ili elektronski recept. Prikaz ekrana ovih aplikacija dat je na slikama Slika 6 i Slika 7.

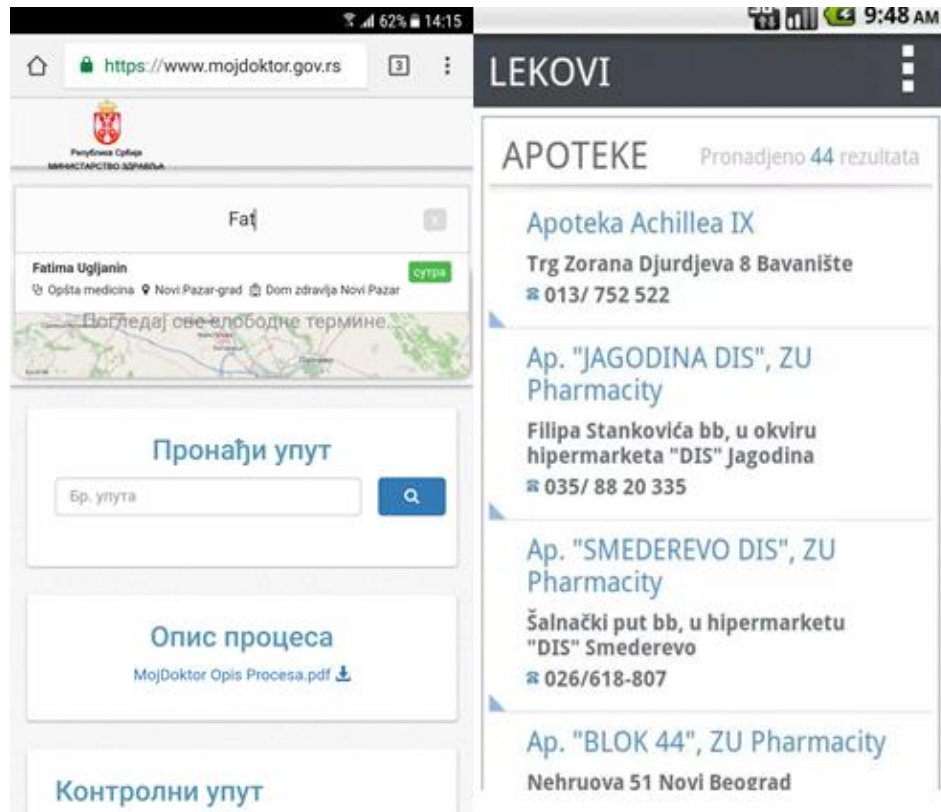


Slika 6. Postojeći servisi na portalu e-Zdravlje

Takođe, postoji i registar dostupnih lekova i medicinskih sredstava po apotekama i mogućnost pretraživanja po imenu leka ili supstanci. Radi se o *Mediatelly* portalu [49] za koji postoje urađene i *Android* i *IOS* aplikacije. On je deo projekta *Portal otvorenih podataka*,



[50], koji pored podataka o zdravlju, sadrži i javne podatke o upravi, obrazovanju, javnoj bezbednosti, životnoj sredini itd. Prikaz ovog servisa dat je na slici Slika 7.



Slika 7. Prikaz ekrana postojećih servisa *MojDoktor.rs* i *Mediatelly*

Što se tiče internih informacionih sistema, koji služe za softversku podršku zdravstvenoj zaštiti, možemo izdvojiti *MEDIS.NET* [51] informacioni sistem koji se uspešno koristi u 25 zdravstvenih ustanova. Od brojnih komponenti sistema za čuvanje administrativnih podataka, anamneze, zdravstvenog kartona i sl., ovaj informacioni sistem ima mogućnost korišćenja od strane eksternih aplikacija, pa je u radu [51] opisan primer mobilne aplikacije koja koristi senzore za merenje pulsa i srčanog pritiska koji interaguje sa ovim informacionim sistemom uz mogućnost poređenja merenih rezultata sa istorijom bolesti pacijenta i dobijanja odgovora od strane lekara. Prikaz ovog informacionog sistema dat je na slici Slika 8.

Pored *MEDIS.NET* informacionog sistema namenjenog zdravstvenim ustanovama, u našoj zemlji u upotrebi su još i rešenja poput *HELIANT Health* [52], *ZipSoft e-Karton* [53] i *Comtrade* [54] informacionih sistema za zdravstvo.



## **4. ARHITEKTURA IS-A KOJI OBJEDINJUJE HETEROGENE SERWISE PAMETNOG ZDRAVSTVA**

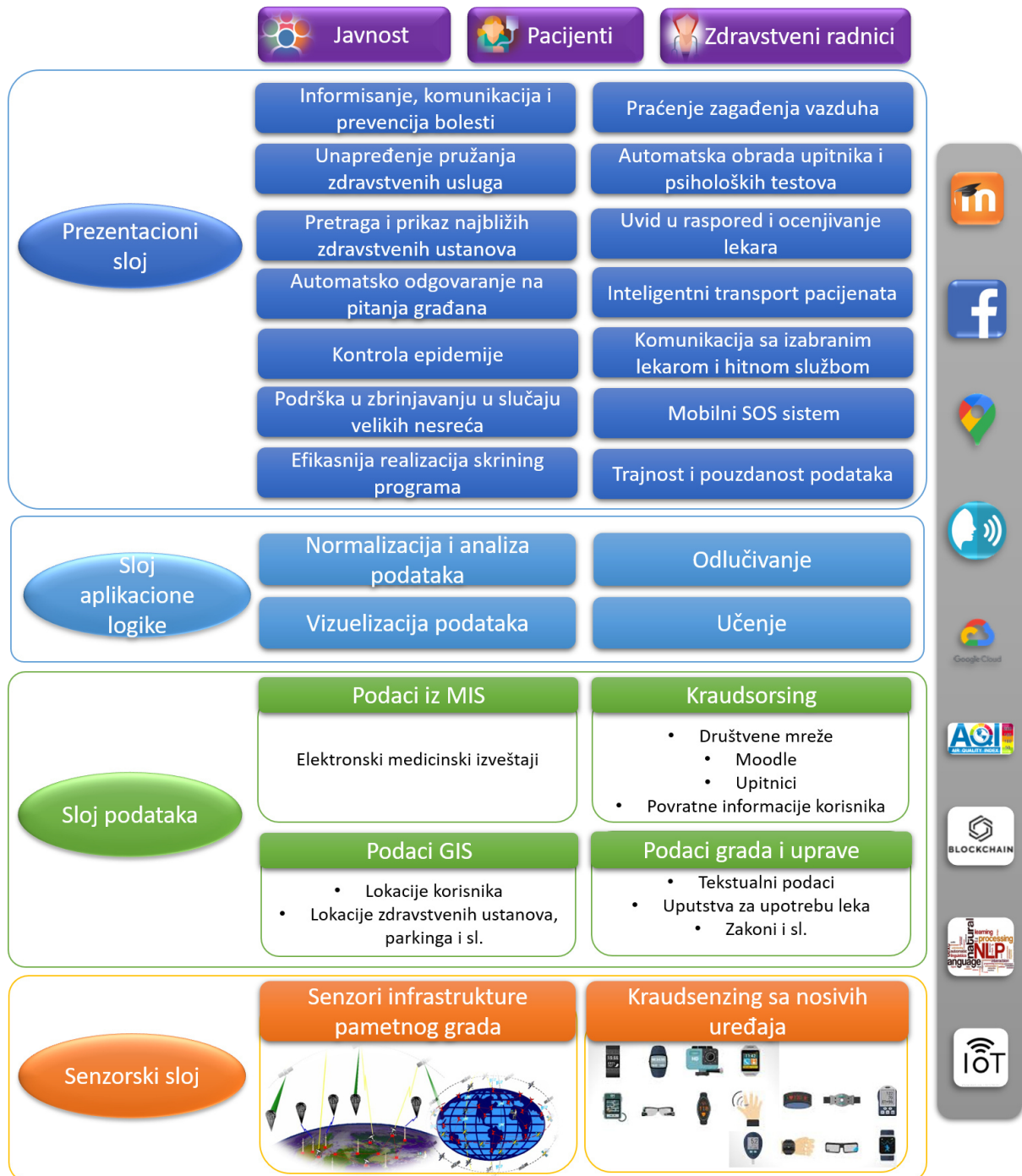
Razvoj informacionih tehnologija doprineo je digitalizaciji servisa u raznim oblastima, pa ni zdravstvo nije odolelo tom trendu. Uveliko se koriste softveri za kreiranje, čuvanje i upravljanje elektronskim medicinskim izveštajima interno u bolnicama, zatim su tu i elektronski recept, portali za zakazivanje pregleda i dijagnostičkih procedura, portali za informisanje o dostupnosti lekara, pretraga lekova u bazama lekova, informacije o zagađenju vazduha i sl. Problem je što svi ovi servisi se ne nalaze na jednom mestu, i ne pružaju mogućnost interakcije s korisnicima. U ovom poglavlju dat je predlog softverske platforme koja bi objedinila postojeće servise javnog zdravstva ali i pružila servise koji do sada nisu realizovani, kao što je servis za kontrolu epidemija, pretraga medicinskih dokumenta, automatska obrada psiholoških testova i mnoštvo drugih. Prikazani su uniformni način za njihovu implementaciju i integraciju u koncept pametnog grada, kao i način za ostvarivanje kolaboracije sa postojećim sistemima. Ovakav način realizacije servisa pametnog zdravstva obezbeđuje jednostavnije uključivanje i dodavanje novih servisa, a može se primeniti i na ostale servise u pametnog grada.

### **4.1. Integracija heterogenih servisa e-zdravstva u sistem pametnog grada**

Motivacija za predlog integracije heterogenih servisa pametnog zdravstva proizašla je iz činjenice da su postojeći servisi javnog zdravlja u Srbiji nedovoljni, nisu svi centralizovani i mogli bi da pruže građanima više od onoga što danas pružaju. Doprinos ove disertacije je i u identifikaciji relevantnih funkcionalnosti koje bi sistem pametnog zdravstva trebao da

obezbedi i predloži, i u razvoju arhitekture IS pametnog zdravstva i njegovih najznačajnijih usluga.

Na slici Slika 9 dat je prikaz informacionog sistema (IS) koji objedinjuje heterogene servise pametnog zdravstva.



Slika 9. Slojevita arhitektura IS koji objedinjuje heterogene servise pametnog zdravstva

Arhitektura ovog IS-a sastoji se od četiri komponente: senzorski sloj, sloj podataka, sloj aplikacione logike i prezentacioni sloj. Zatim su na slici prikazani korisnici ovakvog sistema, i tehnologije i koncepti na kojima on počiva.

Senzorski sloj se sastoji iz dve komponente. Prvu komponentu čine senzori koji pripadaju infrastrukturi pametnog grada, kao što su senzori blizine namenjeni parkiranju, senzori za prikupljanje podataka o zagađenju vazduha i sl. Drugu komponentu čine senzori koji se nalaze na nosivim uređajima korisnika (telefoni, satovi, narukvice, tableti i sl.). U zavisnosti od uređaja i senzora kojima on raspolaže odavde se mogu slati informacije o temperaturi, krvnom pritisku, saturaciji kiseonika, broju otkucaja srca korisnika itd.

Sloj podataka se sastoji se od četiri komponente: podaci iz medicinskih informacionih sistema iz zdravstvenih ustanova, podaci i dokumenta kojima raspolaže pametni grad i uprava, podaci dobijeni iz geografskih informacionih sistema i podaci dobijeni kraudsorsingom. Podaci iz medicinskih informacionih sistema sadrže elektronske medicinske izveštaje, iz kojih se može izvlačiti znanje o broju pacijenata sa odgovarajućom dijagnozom, simptomima i sl. Nad podacima i dokumentima pametnog grada i uprave može se vršiti pretraga, i građani mogu automatski dobiti odgovore na neka pitanja. Geografski informacioni sistemi sadrže informacije o lokacijama zdravstvenih ustanova i drugih entiteta koje mogu biti od koristi prilikom kreiranja servisa pametnog zdravstva. Podaci dobijeni od građana sa društvenih mreža, povratnim informacijama iz servisa, preko *LMS* sistema, popunjavanjem upitnika, takođe mogu biti osnova za dobijanje znanja i kreiranje servisa koji će pozitivno uticati na zdravlje građana.

Sloj aplikacione logike čine takođe četiri komponente. Normalizacija i analiza podataka odnose se na podatke koji su sakupljeni. Oni su obično u različitim formatima, sadrže redundantne informacije koje treba da se prečiste, i treba da budu pogodni za dalju kompjutersku obradu i izvlačenje znanja. Zatim sledi odlučivanje i učenje na osnovu dobijenih podataka, i vizuelizacija podataka je neophodna za određene vrste servisa (prikaz zagađenja, najbližih zdravstvenih ustanova, grafikona sa podacima o epidemiji).

Prezentacioni sloj čini četrnaest servisa pametnog zdravstva koji se mogu realizovati na osnovu opisanih podataka i tehnologija, a to su: servis za informisanje, komunikaciju i prevenciju bolesti zasnovan na *LMS* sistemima i društvenim mrežama, servis za unapređenje

pružanja medicinskih usluga, servis za prikaz najbližih apoteka i zdravstvenih ustanova, servis za automatsko odgovaranje na pitanja građana i pretragu dokumenata iz medicinskog domena, servis za praćenje zagađenja vazduha, servis za automatsku obradu upitnika i psiholoških testova, servis za inteligentni transport pacijenata, servis za kontrolu epidemija i vizuelizaciju podataka o vakcinacijama, servis za komunikaciju sa izabranim lekarom i ekipom hitne pomoći, uvid u raspored i ocenjivanje lekara, servis za mobilni *SOS* sistem, servis za podršku zbrinjavanju u slučajevima velikih nesreća kao što su poplave, požari, zemljotresi i sl., servis za efikasniju realizaciju skrining programa, i servis za obezbeđivanje trajnosti i pouzdanosti podataka, zasnovan na blokčejnu. Ovaj skup servisa je odabran tako da pojedini servisi zapravo reprezentuju podskupove sličnih servisa koji se mogu razvijati za konkretan grad i konkretne potrebe stanovništva i nikako se ne može tretirati kao potpuni skup svih mogućih servisa pametnog grada orijentisanih ka zdravlju.

Korisnici ovakvog IS-a pametnog grada su javnost, pacijenti i zdravstveni radnici. Javnost čine zdravi građani kojima su potrebne razne informacije kao što na primer informacija o stanju epidemije u njihovom mestu, koji žele da obave kontrolne preglede, da učestvuju u skrining programu i sl. Pacijentima predloženi servisi pružaju mnoštvo informacija, kao što su komunikacija sa stručnim osobljem, i efikasnije izvršavanje procedura bez administrativnih poteškoća itd. Zdravstveni radnici koriste pojedine servise za olakšavanje medicinskih procedura, onlajn komunikaciju sa pacijentom, vozači hitne pomoći za izbor najbezbednije rute za prevoz pacijenta itd.

Na kraju su dati i servisi i tehnologije na kojima počiva ovakav IS, a to su *LMS* sistemi, GIS sistemi, društvene mreže, kraudsorsing, kraudsensing, blokčejn, *IoT*, obrada prirodnih jezika i drugi dostupni servisi.

## **4.2. Struktura uniformnog veb servisa za realizaciju servisa e-zdravstva**

Predloženi IS iz prethodne sekcije objedinjuje podatke iz različitih vrsta izvora (senzori, tekstualni dokumenti, geolokacije). Potrebno je da se njihova transformacija i

korišćenje realizuje na uniforman način, kako bi se obezbedila skalabilnost sistema uz jednostavnost dodavanja i integracije novih servisa pametnog zdravstva. Potrebno je da se podaci čuvaju u jedinstvenom formatu, i da im se pristupa na uniforman način. Stoga se arhitektura medicinskog informacionog sistema koja omogućava integraciju heterogenih računskih celina oslanja na principe *REST* (engl. *REpresentational State Transfer*) [55], standardne protokole komunikacije i formate predstavljanja resursa. *REST* omogućava jednostavnost u smislu protokola interakcije i fleksibilnost u pogledu podržanih formata predstavljanja i to ga čini pogodnim arhitekturnim stilom za integraciju u sveprisutnim okruženjima [56].

*REST* je stil softverske arhitekture orijentisan na resurse, kreiran od strane *R. Fielding-a*, a služi za izgradnju distribuiranih aplikacija na Internetu [55]. Obično *REST* definiše principe enkodiranja (tipovi sadržaja), adresiranja (imenice) i pristupa (glagoli) kolekcijama resursa pomoću Internet standarda.

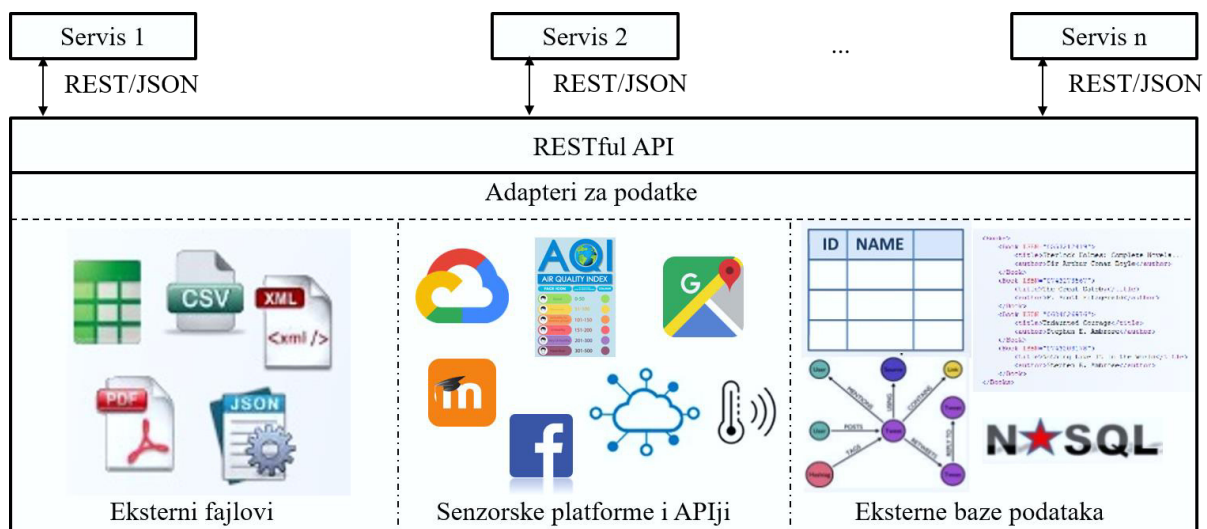
Resursi koji su centralni za *REST* mogu se jedinstveno adresirati koristeći univerzalnu sintaksu (npr. *URL* u *HTTP-u*) i dele jedinstveni interfejs za prenos stanja aplikacija između klijenta i servera (npr. *GET / POST / PUT / DELETE* u *HTTP-u*). *REST* resursi se obično mogu prikazivati u više reprezentacija - na primer - *XML*, *JSON*, *YAML* ili tekstualnim dokumentima. *RESTful* sistemi su labavo povezani sistemi koji slede ove principe kako bi se razmena stanja aplikacija obavila kao reprezentacija resursa. Ova vrsta interakcije nezavisne od prethodnog stanja (engl. *stateless*, tretiranje svakog zahteva kao nezavisne transakcije koja nije povezana sa bilo kojim prethodnim zahtevom) efikasnija je u konzumiranju resursa i pogodna za skalabilnost sistema.

Prema radu [55], „*REST*-ovo razdvajanje klijent-server aktivnosti pojednostavljuje implementaciju komponenti, smanjuje složenost semantike konektora, poboljšava efikasnost podešavanja performansi i povećava skalabilnost čistih serverskih komponenti. Ograničenja slojevitog sistema omogućavaju uvođenje posrednika, kao što su *proxy*, *gateway* i *firewall*, na različitim tačkama komunikacije bez promene interfejsa između komponenti, omogućavajući im tako pomoć u prevođenju komunikacije ili poboljšanje performansi putem visoko-skalabilnog deljenog keširanja. *REST* omogućava međusobnu obradu ograničavanjem poruka da budu samoopisne: interakcija je bez pamćenja prethodnog stanja između zahteva, koriste se

standardne metode i tipovi medija semantike i razmene informacija, a odgovori izričito ukazuju na mogućnost keširanja“.

*REST* na efikasan način omogućava razmenu resursa, koji se mogu enkodirati u različite prezentacije. Stoga, *REST* pruža odgovarajući okvir za posredovanje i obradu heterogenih informacija na efikasan i skalabilan način.

Na slici Slika 10 prikazana je struktura uniformnog veb servisa za realizaciju servisa pametnog zdravstva.



Slika 10. Struktura uniformnog veb servisa za realizaciju servisa pametnog zdravstva

```

{
  "MIS": "DZ NIŠ",
  "EMRS": [
    {
      "datum_rodjenja": "12-05-87",
      "datum_pruzene_usluga": "29-03-18",
      "naziv_usluga": "KRATKA POSETA LEKARU",
      "anamneza": "Laborat GGT 56 AST 62 ALT 134 tg 4.37 crp 7.0",
      "dijagnoza_naziv": "Morbili-male boginje",
      "dijagnoza_sifra": "B05",
      "orgj_pruzanja_usluga": "Opšta medicina",
      "lokacija_pruzanja_usluga": "Centralna zgrada"
    },
    {
      "datum_rodjenja": "02-10-48",
      "datum_pruzene_usluga": "12-01-18",
      "naziv_usluga": "PRVI PREGLED ODRASLIH",
      "anamneza": "",
      "dijagnoza_naziv": "Morbili-male boginje",
      "dijagnoza_sifra": "B05",
      "orgj_pruzanja_usluga": "Opšta medicina",
      "lokacija_pruzanja_usluga": "Centralna zgrada"
    }
  ]
}

```

Slika 11. Primer dela *JSON*-a koji sadrži deo elektronskih izveštaja iz MIS



Ulazni podaci u ovaj servis svrstani su u tri velike grupe: eksterni fajlovi, senzori i *API*-ji i podaci iz baza podataka. Eksterni fajlovi mogu se naći u više različitih formata (*csv*, *txt*, *doc*, *docx*, *pdf* itd.). Senzorski podaci i podaci iz *API*-ja obično se nalaze u *JSON* formatu, a eksterne baze podataka mogu biti relacione, nerelacione, objektno-orijentisane, *XML*, grafovi itd. Potrebno je da se za svaku vrstu izvornih podataka obezbedi adapter podataka, i da se zatim njima pristupa na jedinstven način kroz *RESTful API*. Za dalji prikaz ovih informacija koristi se *REST/JSON* reprezentacija, ali, *REST* je pogodan i za dodavanje dodatnih prikaza podataka, kao što je već rečeno ranije.

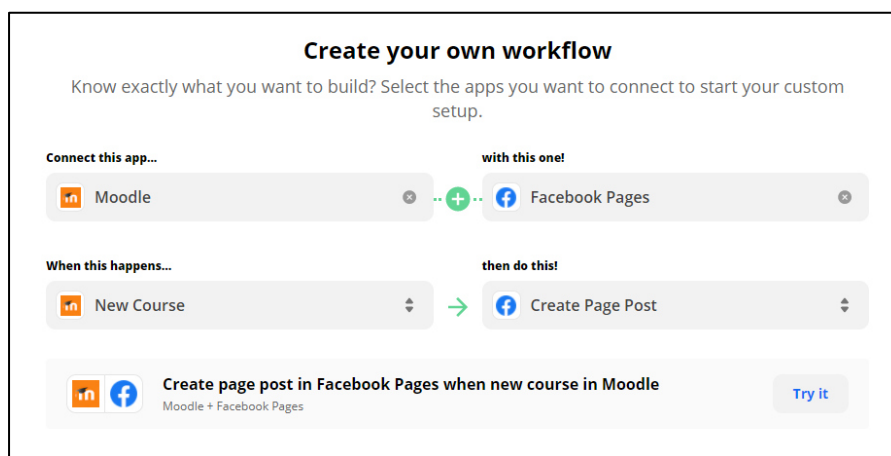
### 4.3. Kolaboracija postojećih sistema u pametnom gradu

Način komunikacije među sistemima i servisima u pametnom gradu mora da obezbedi korišćenje postojećih sistema bez velikih intervencija, ali i laku nadogradnju i uključivanje novih, relativno nezavisnih sistema. Pored toga što servisi koriste *REST* aplikacioni interfejs, i podatke daju u uniformnom *JSON* formatu, potrebno je obezbediti da se ažuriranje podataka obavi na što efikasniji način. Kako servisi koriste podatke iz različitih sistema i obrađuju ih u realnom vremenu, potrebno je da ovi servisi budu obavešteni o promenama podataka koje koriste. Ako jedan ili više servisa ili sistema koriste informacije iz jednog sistema koji je izvor podataka, obaveštavanje o izmeni izvornih podataka može da se realizuje na sledeće načine:

- a) da servis koji koristi podatke stalno periodično proverava da li se desila promena podataka u sistemu koji je izvor (što zahteva veliki saobraćaj),
- b) da se realizuje sistem notifikacije gde se jedan sistem "pretplati" na podatke iz drugog sistema i kada se desi promena podatka koji je od interesa onda sistem izvor šalje notifikaciju pretplaćenom sistemu da se nešto promenilo.

Upravo je drugi slučaj, zbog svoje efikasnosti u pogledu smanjenja saobraćaja, način koji se u okviru ove disertacije predlaže za komunikaciju sistema u pametnom gradu. Na primer, u servisu za praćenje kvaliteta vazduha može se desiti da određeni parametri pređu granicu, i tada treba obavestiti odgovarajuće sisteme (npr. vrtiće ili staračke domove, pretplaćene korisnike) o stepenu zagađenja vazduha. Za obaveštavanje se upravo mogu iskoristiti predloženi sistemi notifikacije.

Postoji više softverskih rešenja koja se mogu iskoristiti u ove svrhe, i oni se zovu sistemi sa mehanizmom notifikacije (engl. *webhook*) [57], a ovaj mehanizam se može i jednostavno implementirati u sistemu koji je izvor podataka. Sistemi notifikacije često se nazivaju inverznim *API*-jem, ali to nije u potpunosti tačno. Oni zapravo ne moraju da dobiju zahtev od sistema koji od njih koristi podatke, već podatke šalju kada god dođe do njihove izmene. Prednost korišćenja sistema notifikacije je jednostavna integracija podataka, zbog razmene podataka u obliku odgovora. Podaci su obrađeni i sačuvani u promenljivoj koja se može koristiti u drugoj aplikaciji. Njihova glavna svrha je približavanje podacima u realnom vremenu.



Slika 12. Primer povezivanja dva različita sistema korišćenjem alata *Zapier*

Na slici je prikazan primer jednog rešenja *Zapier* [58] kojim se mogu povezati dva sistema, od onih koji su ponuđeni. U ovom slučaju izabrani sistemi su *Moodle LMS* [59] i *Facebook* društvena mreža. Događaj kreiranja novog kursa na *LMS*-u izazvaće dodavanje novog posta na *Facebook* stranici. U servisima pametnog zdravstva, u petom poglavlju, pokazano je kako se *LMS* sistemi i društvene mreže mogu iskoristiti u svrhe informisanja, prevencije bolesti i komunikacije sa stručnim osobljem. Upravo kreirani kurs može biti namenjen informisanju i komunikaciji ljudi koje povezuje ista dijagnoza.

## 5. NOVI SERVISI PAMETNOG ZDRAVSTVA

U ovom poglavlju opisani su i isprojektovani novi servisi pametnog zdravstva, a deo njih je realizovan i dati su detalji njihove implementacije. Opisane servise možemo svrstati u različite tipove servisa, zavisno od porekla podataka koje obrađuju i u zavisnosti od tehnologije na kojoj su zasnovani.

Prema načinu porekla podataka, servisi pametnog zdravstva mogu biti zasnovani na podacima dobijenim:

- sa senzora, odnosno kraudsensingom,
- direktnim učešćem građanina tako što daje svoj opis neke situacije, odgovor, obaveštenje, odnosno kraudsorsingom,
- iz medicinskih informacionih sistema,
- iz dokumenata iz IKT infrastrukture pametnog grada.

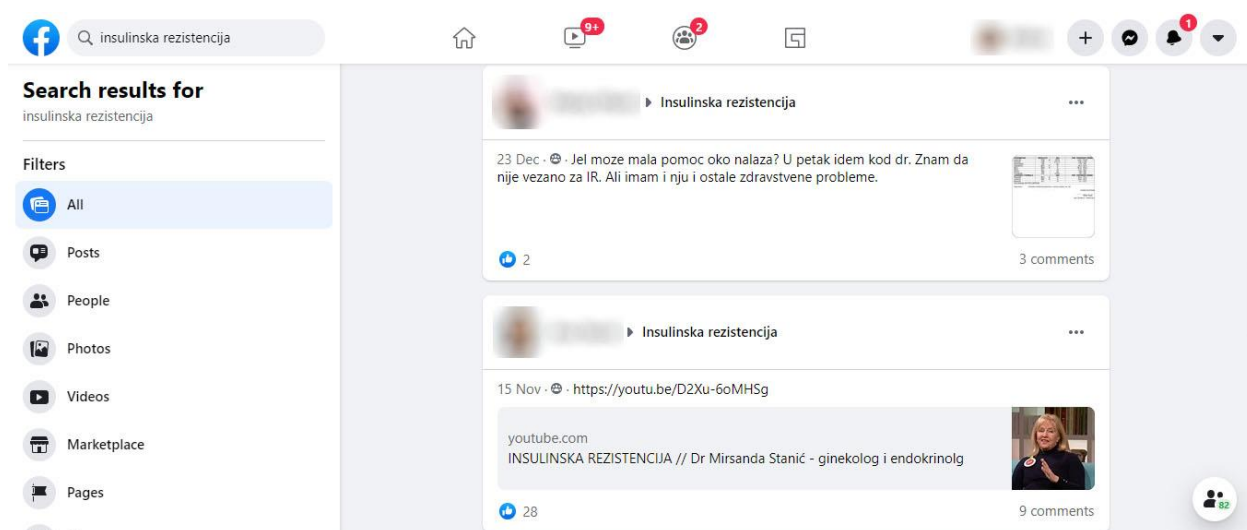
Prema tehnologijama na kojima su zasnovani, servisi pametnog zdravstva prikazani u ovoj doktorskoj disertaciji, svrstani su u pet kategorija:

- servisi zasnovani na *LMS* sistemima i društvenim mrežama,
- servisi zasnovani na *IoT* i geografskim informacionim sistemima,
- servisi koji koriste metode veštačke inteligencije i procesiranja prirodnih jezika,
- servisi zasnovani na multimediji i veb tehnologijama,
- servisi zasnovani na blokčejnu.

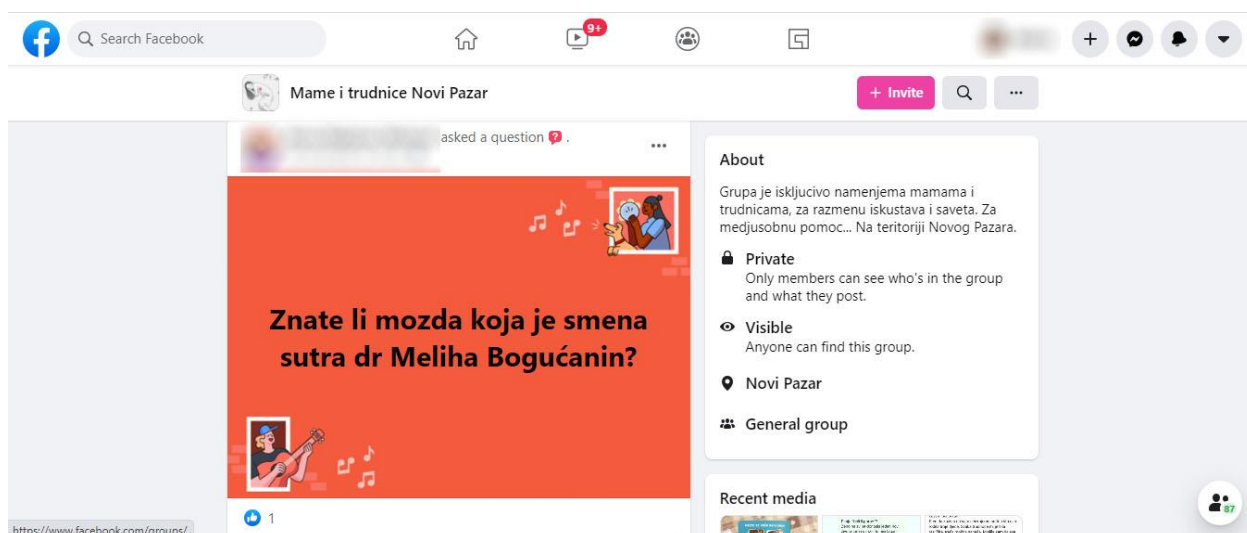
Ovi servisi pametnog zdravstva namenjeni su komunikaciji, prevenciji, informisanju, vizuelizaciji podataka, pružanju podrške, čuvanju trajnosti i pouzdanosti podataka i efikasnijem pružanju usluga. Opisani servisi prikazani su na osnovu primarne tehnologije na kojoj su zasnovani, dok mogu imati više namena, i integrisati podatke iz više ulaza.

## 5.1. Servisi zasnovani na *LMS* sistemima i društvenim mrežama

Sve više ljudi traži informacije vezane za zdravlje na Internetu. Postoje brojni forumi za odgovaranje na pitanja vezana za zdravlje korisnika (*Doktor.rs* [60], *Stetoskop.info* [61]), zatim grupe na društvenim mrežama koje okupljaju korisnike sa sličnim zdravstvenim problemima (Slika 13, Slika 14, Slika 15).



Slika 13. Primer razgovora pacijenata o bolesti na *FB* grupi



Slika 14. Primer informisanja o radu doktora na *FB* grupi

Pogledaj neodgovorene postove | Pogledaj aktivne teme

Zdravlje forum » Medicina » Endokrinologija Sva vremena su u UTC + 2 sata

## Zlatibor (štitna žlezda)

[nova tema](#) [odgovori](#) Stranica 1 od 2 [ 18 Posta ]

**Idi na stranicu 1, 2 Sledeća**

Prethodna tema   Sledeća tema	
Autor	Poruka
Tema posta: Zlatibor (štitna žlezda) <span style="float: right;">Poslato: Pet Maj 16, 2008 4:25 pm</span>	
<b>ana22</b> <span>offline</span> Stalni član	<b>Pridružio se:</b> Pet Maj 09, 2008 4:06 pm <b>Postovi:</b> 64
Na forumu sam pronašla da se u vezi sa problemima štitne žlezde pominje Zlatibor. Kako tek od skoro znam da imam hipotireozu interesuje me da li je odlazak na Zlatibor koristan i u slučaju smanjene funkcije, kao i koja je korist odlaska na Zlatibor (da li treba biti na lečenju u institutu ili je jednostavno tamo dobar vazduh, klima ili šta li već).	

Slika 15. Post na forumu *Doktor.rs*

Važno je ovim grupama korisnika obezbediti odgovore iz pouzdanih izvora, i na taj način podstaći informisanje o bolesti, prevenciju bolesti i komunikaciju sa stručnim licem. Kombinovanjem mogućnosti koje nude sistemi za elektronsko učenje i društvene mreže, moguće je formirati servise pametnog zdravstva koji bi centralizovali podatke o bolestima i omogućili komunikaciju između pacijenata sa istom dijagnozom i konsultacije sa stručnim licem. Primer takvog servisa namenjenog informisanju, prevenciji i komunikaciji sa stručnim osobljem prikazan je u ovom poglavlju.

U nastavku su definisani osnovni pojmovi neophodni za razumevanje narednog dela teksta.

**Elektronsko učenje** - učenje korišćenjem informacione i komunikacione tehnologije [59].

**LMS (engl. Learning Management System)** - softver za upravljanje procesom učenja koji omogućava da se studenti upisuju na kurseve, da se studentima distribuiraju materijali, da se prati njihovo napredovanje u elektronskom učenju i da se beleže rezultati rada i učenja [6].

**Moodle (Modular Object-Oriented Dynamic Learning Environment)** - besplatna softverska platforma, koja spada u grupu LMS aplikacija [59].

**Socijalna mreža** - društvena struktura koja je određena društvenim interakcijama čiju osnovu čine društveni mediji, koji predstavljaju „skupinu Internet aplikacija koje su izgrađene

na ideološkim i tehnološkim osnovama Veb 2.0 tehnologije koje omogućavaju kreiranje i razmenu korisnički generisanog sadržaja“ [63].

**Facebook** - društvena mreža koja omogućava postavljanje multimedijalnog sadržaja na profil, stranicu ili grupu, koji je vidljiv određenoj skupini korisnika, sadrži alat za komunikaciju, i mogućnost integracije sa standardnim veb aplikacijama i brojne druge mogućnosti [64][65].

Učenje, već dugi niz godina, uz tradicionalne metode koristi i sve povoljnosti Interneta, od bržeg pristupa resursima znanja do sistema koji nude pored literature u vidu dokumenata, i multimedijalne sadržaje, kvizove, komunikaciju sa učesnicima kurseva i nastavnicima kroz forum ili privatne poruke, i sl. Takav *LMS* je i *Moodle*. Može se reći da je bez nekog *LMS*-a danas skoro nemoguće zamisliti podelu literature i onlajn komunikaciju između profesora i nastavnika na kvalitetan način. Postavlja se pitanje da li je moguće iskoristiti ovakav *LMS* sistem za informisanje i komunikaciju između pacijenata i stručnih lica, i na taj način obezbediti da ljudi koji sve češće traže informacije o bolestima na Internetu, budu upućeni na pouzdane izvore informacija.

S druge strane, veliki broj korisnika računara provodi puno vremena na društvenim mrežama, od kojih je *Facebook* jedna od najpopularnijih, i sa najviše korisnika. Na *Facebook*-u korisnici mogu postavljati objave multimedijalnog tipa, ali daleko od toga da se ova društvena mreža koristi samo za zabavu. Takođe, većina sajtova omogućava prijavu korišćenjem *Facebook* profila, pa se ne moraju pamtiti sva moguća korisnička imena i šifre za različite sajtove na koje smo prijavljeni. Pored toga, postoje i aplikacije, koje obavestavaju svoje učesnike o trenutnim zbivanjima, pokušavajući na taj način da nagovore korisnika da pokrene aplikaciju. Postoje i otvorene i zatvorene grupe korisnika koje polemišu o nekoj temi. Neretko su to grupe ljudi koji imaju neku specifičnu dijagnozu, i žele da su u kontaktu sa ljudima koji prolaze kroz isto, i na taj način vide tuđa iskustva. Neretko se na ovaj način ljudi informišu gde je moguće kupiti neki lek, ako je nestašica u gradu korisnika, zatim ljudi neretko postavljaju rezultate analiza i pitaju za mišljenje. Ovo pokazuje da postoji potreba da se obezbedi komunikacija sa stručnim licima. Zašto onda i sistem za učenje namenjen informisanju o bolesti, prevenciji i komunikaciji sa stručnim licima ne bi bio na raspolaganju svome korisniku i omogućio mu da koristi *Facebook* profil za prijavu, i dobija notifikacije o

važnim dešavanjima na kursevima na koje je korisnik upisan a distribuiraju mu se pomoću *LMS*-a.

Ovaj pristup ima opravdanja i u činjenici da neuporedivo više vremena korisnici provode na nekoj društvenoj mreži u odnosu na vreme provedeno u korišćenju *LMS*-a odnosno češće prate dešavanja na društvenim mrežama, a ređe na *LMS*-u. Ovu činjenicu je moguće iskoristiti tako da se obezbedi aktivnije praćenje aktivnosti na *LMS*-u putem pojavljivanja različitih informacija o *LMS* aktivnostima na društvenim mrežama. Ovim servisom se opisuje jedan način integracije socijalne mreže *Facebook* sa *Moodle LMS*-sistemom, u cilju kreiranja servisa pametnog zdravstva namenjenog prevenciji bolesti, informisanju i komunikaciji sa stručnim medicinskim osobljem [65].

Postoje radovi u kojima je opisana upotrebljivost socijalnih mreža u edukativne svrhe, kao što su [68][69][70]. Ono što je zanimljivo jeste istraživanje na temu šta studenti u Srbiji više koriste, *Facebook* ili *Moodle*, što može biti pokazatelj koliko su težnje njihove integracije ispravne i upotrebljive [71]. Da se *Moodle* može koristiti u svrhu informisanja pacijenata pokazano je u radu [72].

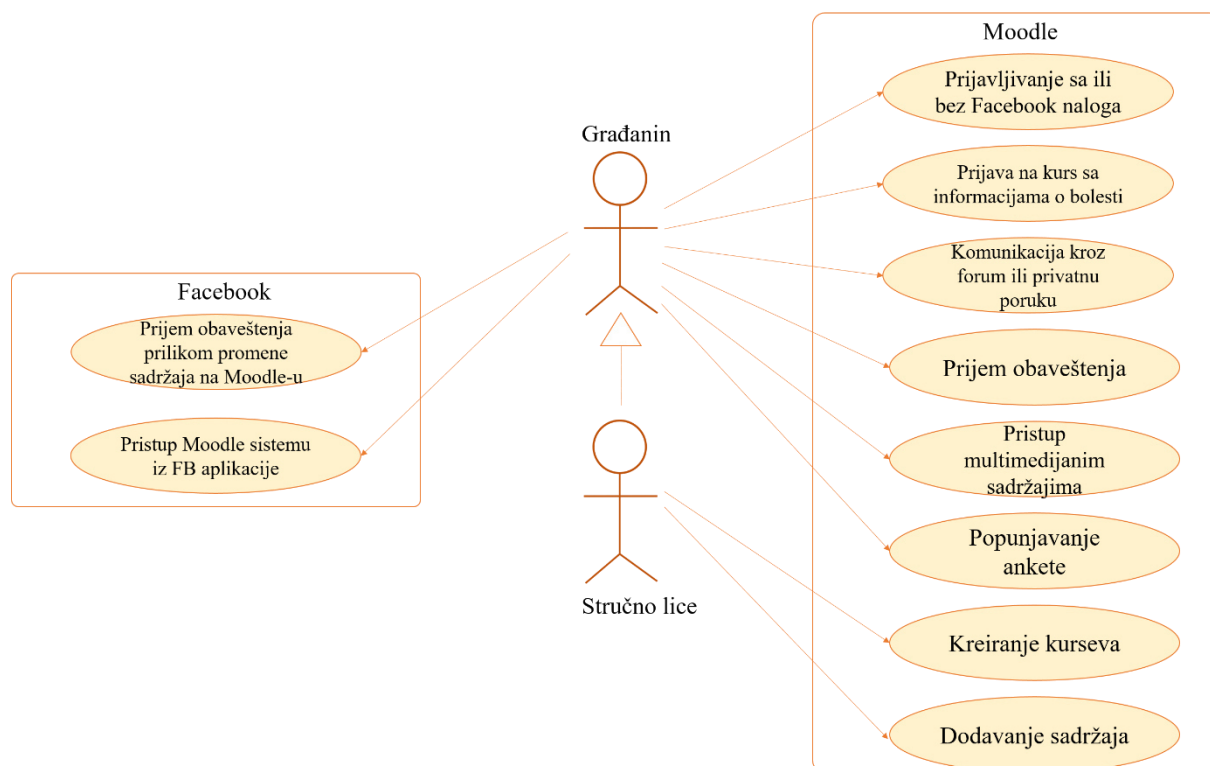
### **5.1.1. Informisanje, komunikacija i prevencija bolesti**

Promovisanje zdravlja i prevencija postaju sve važniji u zdravstvenoj zaštiti pacijenata i šire javnosti. Promovisanje zdravlja kroz Zakon o javnom zdravlju, kao i Zakon o zdravstvenoj zaštiti postali su obavezujući [66][67]. Faktori okruženja, životna sredina, kao i životni stil predstavljaju ključne polazne tačke za očuvanje i promovisanje i individualnog i socijalnog zdravlja.

Unapređenje zdravlja je usmereno na proces koji omogućava svim ljudima da imaju veće samoopredeljenje o svom zdravlju, i na taj način osnažuju svoje zdravlje. Prevencija ima za cilj da spreči, učini manje verovatnim ili da usporava određena oštećenja zdravlja ili bolesti putem ciljanih aktivnosti. Razlikuje se primarna prevencija, sekundarna prevencija i tercijarna prevencija u zavisnosti od statusa bolesti u preventivnoj meri.

Promocija zdravlja i prevencija bolesti se u praksi mnogo preklapaju i dopunjuju, budući da prevencija podiže svest o tome koje rizike treba uzeti u obzir u pojedinačnim slučajevima. Promocija zdravlja takođe pomaže da se mobilišu resursi građana kao što su pacijenti, rođaci i okruženje u tu svrhu. Aplikacije e-zdravstva, posebno aplikacije m-zdravstva, potencijalno se smatraju korisnim u podržavanju implementacije određenih aktivnosti za promociju zdravlja i prevenciju.

Na slici Slika 16 dat je *UML* dijagram slučajeva korišćenja za servis za informaciju, komunikaciju i prevenciju bolesti. Ovaj servis baziran je na postojećem *LMS* sistemu *Moodle* i društvenoj mreži *Facebook*, jer se njihove funkcionalnosti mogu iskoristiti za poboljšanje informisanosti o bolestima, podizanje svesti i prevenciju od bolesti i komunikaciju sa stručnim licima. Na ovaj način objedinile bi se potrebe građana za informisanjem o bolestima, ali bi mogli da se informišu na jednom mestu, da pristupaju sadržajima koji su centralizovani, a odgovor na pitanja i postavljanje sadržaja bilo bi kontrolisano od strane stručnog lica.



Slika 16. Dijagram slučajeva korišćenja servisa za informisanje i prevenciju bolesti

Mogućnosti korisnika na *Moodle*-u bile bi: prijavljivanje sa ili bez *Facebook* naloga, prijavljivanje na kurs koji sadrži informacije o bolesti, a koji kreira i vodi stručno lice iz

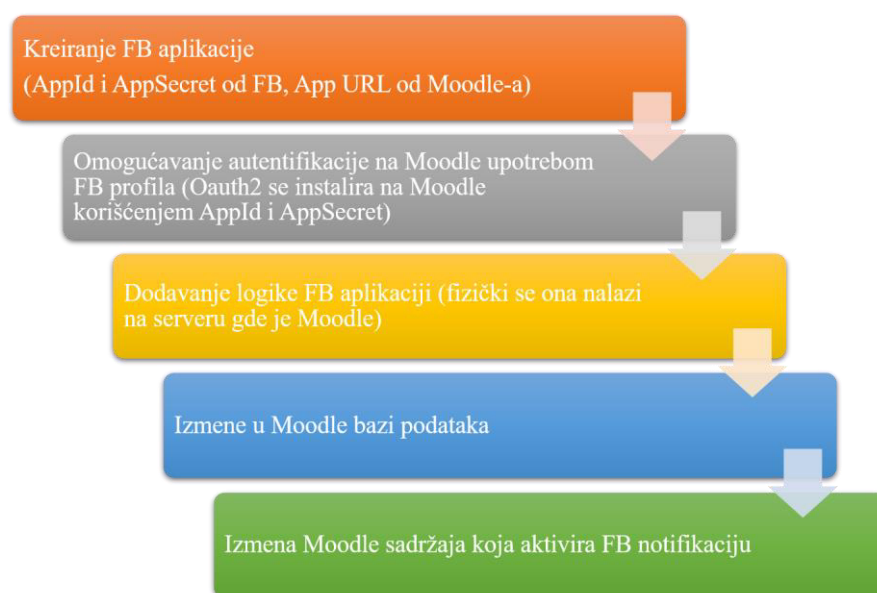


oblasti, komunikacija sa kreatorom i članovima kursa preko foruma ili direktno preko privatne poruke, pristup multimedijalnim sadržajima (tekstualnim, prezentacijama, video snimcima, slikama, linkovima itd.), popunjavanje upitnika, praćenje novosti i obaveštenja na kursu.

Mogućnosti korisnika na FB bile bi prijem notifikacije pri promeni sadržaja na *Moodle*-u i uvid u promene i pristup aplikaciju, korišćenjem FB aplikacije.

*Stručno lice* je korisnik izveden iz korisnika *građanin*, pa bi ono imalo sve mogućnosti kao i građanin, uz dodatne mogućnosti na *Moodle*-u, kao što su postavljanje sadržaja i kreiranje kurseva.

Na slici Slika 14 su dati koraci kolaboracije *Moodle LMS* aplikacije sa *Facebook* aplikacijom, a to su: kreiranje *FB* aplikacije, omogućavanje prijave na *Moodle* upotrebom *FB* profila korišćenjem *OAUTH2* alata, dodavanje logike *FB* aplikaciji kako bi pri promeni sadržaja na *Moodle*-u stizale notifikacije, izmene u *Moodle* bazi podataka - obezbeđivanje ID korisnika kao uslov za mogućnost slanja notifikacije određenom korisniku i promena sadržaja na *Moodle*-u. Tehnički detalji implementacije i njeni rezultati detaljno su prikazani u radu [48].



Slika 17. Koraci u kolaboraciji *Moodle* sistema i *FB* aplikacije

Prikazanim postupkom povećava se opseg funkcionalnosti *Moodle*-a i *Facebook*-a, i omogućava centralizovan pristup informacijama o bolestima, prevenciji i komunikaciji korišćenjem postojećih besplatnih rešenja.

## 5.2. Servisi zasnovani na *IoT* i geografskim informacionim sistemima

**GIS (geografski informacioni sistem)** predstavlja okvir za prikupljanje, upravljanje i analizu prostornih podataka i njima pridruženim osobinama. GIS integriše mnoge vrste podataka, analizira prostorne lokacije i organizuje slojeve informacija u vizuelizacije pomoću mapa i 3D scena. Sa ovom jedinstvenom sposobnošću, GIS otkriva dublji uvid u prostorne podatke, poput obrazaca, odnosa i situacija, pomažući na taj način korisnicima da donose pametnije odluke. GIS se sastoji od četiri interaktivne komponente: podsistem za unos koji vrši konverziju mapa i druge prostorne podatke u digitalnom obliku, podsistem za skladištenje i pozivanje podataka, podsistem za analizu i izlazni podsistem za prikaz mapa i za pružanje odgovora na postavljene upite [7].

**Google Maps API** predstavlja popularni Internet servis za vizuelizaciju geografskih informacija koju je Google pokrenuo početkom 2005. godine kako bi pružio vizuelni interfejs koristeći *AJAX* tehnologije. Ubrzo nakon toga, *Google* je pokrenuo *Google Maps API*, *JavaScript API*, i on omogućava ugradnju stranice *Google Maps* u eksternu veb lokaciju i dodavanje sloja određenih podataka na mapu [73].

**GeoJSON** je standard za geoprostorne podatke na Internetu koje koristi i *Google Maps API*. Klasa podataka prati strukturu *GeoJSON*-a u njegovom predstavljanju podataka i olakšava prikaz *GeoJSON* podataka (tačaka, nizova linija i poligona)[73][74] Na slici Slika 18 dat je primer geografskih podataka prikazanih u *GeoJSON* formatu [75].

Pojam Internet stvari (*IoT*) pomenut je u poglavlju koje opisuje tehnologije na kojima počiva pametni grad, a predstavlja je umrežavanje fizičkih uređaja, vozila i zgrada korišćenjem ugrađenih senzora sa proizvođačem ili drugim uređajima.

```
{
  "type": "Feature",
  "geometry": {
    "type": "Point",
    "coordinates": [43.3158482925053, 21.913748798826735]
  },
  "properties": {
    "name": "Klinički centar Niš",
    "description": "Državna bolnica",
    "telephone": "+38118506906"
  }
}
```

Slika 18. Primer geografskih podataka u *GeoJSON* formatu

U narednim poglavljima opisani su servisi zasnovani na *IoT* i geografskim informacionim sistemima. Oni se koriste u servisima za pretragu i prikaz najbližih zdravstvenih ustanova, u unapređenju pružanja medicinskih usluga, inteligentnom transportu pacijenata i mobilnom *SOS* sistemu.

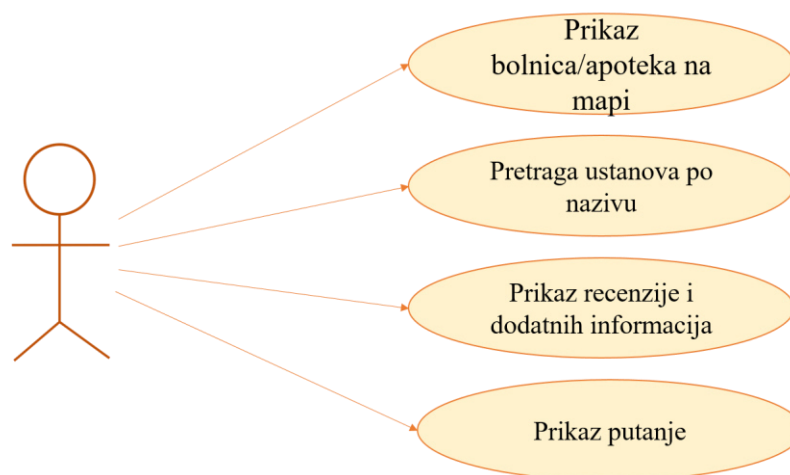
### 5.2.1. Pretraga i prikaz najbližih zdravstvenih ustanova

Centralizovani IS za pametno zdravstvo trebalo bi da iskoristi mogućnosti dostupnih GIS sistema koji sadrže informacije o zdravstvenim ustanovama i apotekama. Korisniku treba obezbediti da na jednom mestu može izvršiti pretragu i prikaz ambulanti domova zdravlja, bolnica, apoteka, privatnih ordinacija, sa pridruženim informacijama o njima, kao što su ocene ustanova i link ka dodatnim informacijama na veb sajtu i prikaz putanje od njegove lokacije do izabrane ustanove.

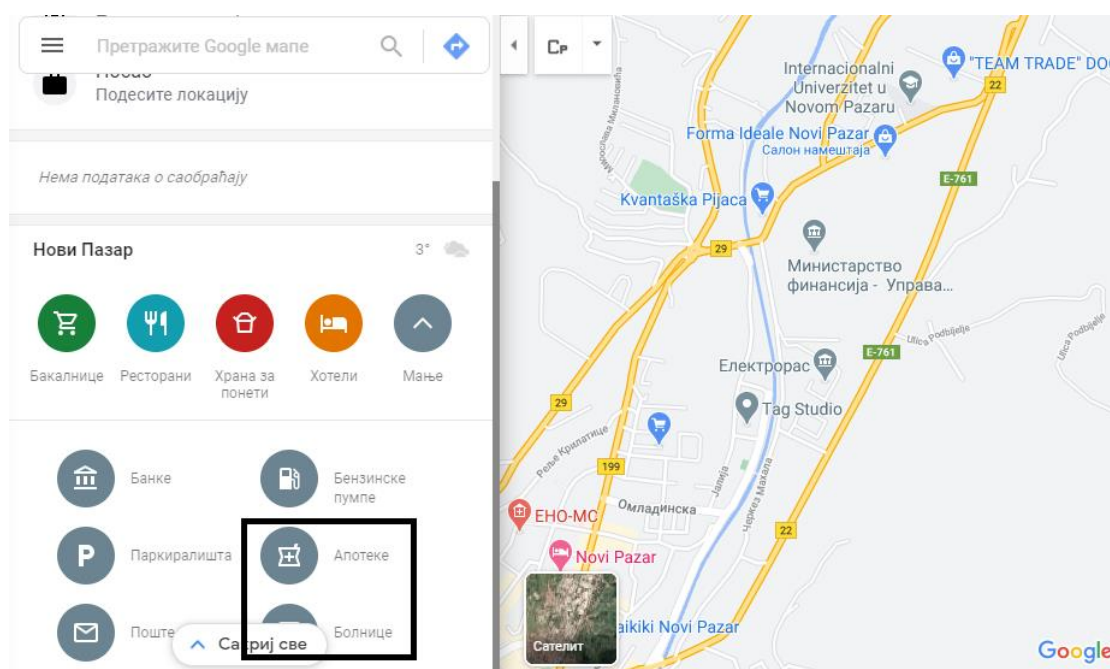
Na slici Slika 19 prikazan je *UML* dijagram slučajeva korišćenja za servis pametnog zdravstva za pretragu i prikaz najbližih zdravstvenih ustanova i apoteka.

Ove mogućnosti već pruža *Google Maps API*, i pogodnosti integrisanja u druge veb sajtove sa izborom podskupa potrebnih geopodataka, pa bi se mogao integrisati i u predloženi IS pametnog zdravstva. Na slici Slika 20 je dat prikaz *Google* mapa sa izborom za prikaz

apoteka i bolnica, a na slici Slika 21 je dat prikaz detaljnih informacija kada se odabere objekat na mapi. Dodatne informacije nude prikaz slike pridružene objektu, ocenu i komentare, link ka veb sajtu ako ga objekat ima, mogućnost deljenja i mogućnost prikaza putanje od lokacije korisnika do izabranog objekta.



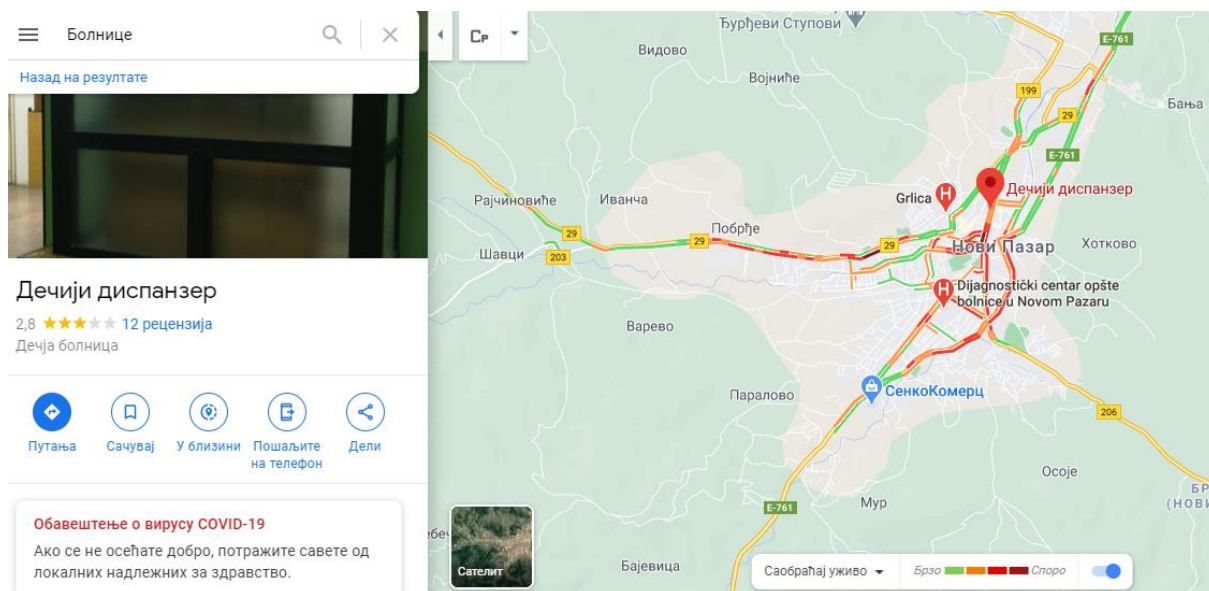
Slika 19. Dijagram slučajeva korišćenja za servis za pretragu i prikaz najbližih zdravstvenih ustanova i apoteka



Slika 20. Izbor apoteka i bolnica na Google mapama

Ovo nije jedini servis pametnog zdravstva u kom se mogu iskoristiti pogodnosti geografskih informacionih sistema. Prikaz mapa i lokacija može se koristiti u servisu za praćenje zagađenja vazduha, u pametnom transportu za prikaz najudobnijih ruta i najbližih

parkirališta, kao i u kontroli epidemije, za vizuelizaciju boja obolelih po gradovima ili manjim oblastima. Značaj GIS-a za javno zdravlje detaljno je opisan u knjizi [76].



Slika 21. Dodatne informacije o izabranoj bolnici

## 5.2.2. Praćenje kvaliteta vazduha

U ovom poglavlju prikazan je opis servisa pametnog zdravstva, zasnovanog na *IoT*, koji bi stanovnicima pametnih gradova pružio podršku u vidu vizuelizacije podataka o kvalitetu vazduha u njihovom okruženju, slanja obaveštenja ukoliko su u opasnosti i pružanja medicinske pomoći. Servis nudi mogućnost interakcije sa korisnicima i obrade informacija koje oni šalju, kao i mogućnost slanja i obrade zahteva za pomoć. Za realizaciju ovog servisa potrebni su senzori za prikupljanje informacija o lokacijama i o kvalitetu vazduha na ovim lokacijama. Pored GIS-a i *IoT* uređaja, ovaj servis zasnovan je na kraudsorsingu i kraudsenzingu, pa slede opisi ovih termina.

**Kraudsorsing (engl. crowdsourcing)** je postupak sakupljanja informacija od strane javnosti putem Interneta kako bi se rešili problemi iz različitih oblasti [5], [77].

**Kraudsensing (engl. crowd-sensing)** je tehnika sakupljanja i deljenja podataka sa mobilnih uređaja koji imaju ugrađene senzore u cilju njihove analize radi zajedničke koristi [78].

**Indeks kvaliteta vazduha (engl. AQI - Air Quality Index – u daljem tekstu IKV)** je vrednost kojom se izveštava kvalitet vazduha na dnevnom nivou [79]. On govori o tome koliko je vazduh čist ili zagađen i kako je to povezano sa zdravstvenim stanjem ljudi podneblja na kome se meri.

Tabela 3. Uticaj indeksa kvaliteta vazduha na zdravlje

Nivo indeksa kvaliteta vazduha	Značenje	Brojna vrednost
Dobro	Zadovoljavajući nivo kvaliteta vazduha; uticaj na zdravlje mali i bez rizika	0-50
Umereno	Prihvatljiv nivo kvaliteta vazduha sa umerenim uticajem na zdravlje malog broja ljudi	50-100
Nezdravo za osetljive grupe	Nivo kvaliteta vazduha koji može da utiče na zdravlje osetljivih grupa	101-150
Nezdravo	Svi mogu da budu izloženi negativnim efektima zagađenja na zdravstvo, osetljive grupe mogu imati ozbiljnije zdravstvene probleme	151-200
Veoma nezdravo	Alarmantno stanje: svi mogu osetiti negativne efekte zagađenja na zdravlje	201-300
Opasno	Kritična situacija, zdravlje cele populacije je pogođeno zagađenjem.	301-500

IKV govori o efektima na zdravlje koji se mogu pojaviti u periodu od nekoliko časova ili dana nakon udisanja zagađenog vazduha. EPA [80] izračunava IKV za pet najvećih zagađivača vazduha regulisanih od strane *Clean Air Act* [81]: ozon, parcijalno zagađenje, ugljen monoksid, sumpor dioksid i azot dioksid. Za svaki od ovih zagađivača, EPA uspostavlja nacionalne standarde kvaliteta vazduha u cilju zaštite zdravlja stanovništva. Prema SEPA [82], najveće zagađenje u Srbiji izaziva PM10.

Vrednost IKV može da varira između 0 i 500. Veća vrednost IKV-a ukazuje na veći stepen zagađenja i veći uticaj na zdravstveno stanje. Na primer, vrednost IKV od 50 predstavlja dobar kvalitet vazduha sa malom mogućnošću odražavanja na javno zdravlje, dok vrednosti iznad 300 prikazuju kvalitet vazduha opasan po zdravlje. Svakoju kategoriji

odgovara različit nivo zabrinutosti za zdravstveno stanje i svakome je pridružena odgovarajuća boja. Cilj indeksa kvaliteta vazduha pomaže da razumemo kako lokalni kvalitet vazduha utiče na naše zdravlje. Njegove vrednosti podeljene su u šest kategorija prikazanih u tabeli Tabela 3.

### **5.2.2.1. Motivacija**

U velikim gradovima javljaju se različiti problemi koji loše utiču na kvalitet života njihovih stanovnika. Cilj pametnih gradova je da se reše problemi poput neefikasnog parkiranja, da se omogući kontrola epidemija, i da se prati i poboljšava kvalitet vazduha i sl. S tim u skladu, javno dostupne informacije o kvalitetu vazduha na teritoriji grada bile bi od koristi svima, a pogotovo plućnim bolesnicima [77]. U tom kontekstu, istražuju se i analiziraju podaci obezbeđeni od strane pametnih uređaja, stanovništva i senzora ugrađenih u infrastrukturu pametnog grada. Kraudsorsing je metoda za prikupljanje podataka od strane korisnika koji dobrovoljno u tome učestvuju. Servisi pametnog grada zasnovani na ovoj metodi omogućavaju povratnu informaciju od stanovnika pametnog grada.

Motivacija za kreiranje ovakvog servisa pametnog zdravstva je poboljšanje uslova života stanovnika pametnog grada, pogotovu onih koji imaju plućna oboljenja. Predložena je arhitektura sposobna da prikuplja podatke sa različitih tipova *IoT* uređaja i od stanovnika i manipuliše njima koristeći tehnike analiziranja podataka u cilju obaveštavanja građana o kvalitetu vazduha u njihovoj neposrednoj blizini, i pruža mogućnost eventualnog pružanja pomoći ukoliko se građanin nađe u kriznoj situaciji.

### **5.2.2.2. Stanje u oblasti**

Praćenje zagađenja vazduha u gradovima zanimljivo je sa raznih stanovišta. Zagađenje utiče na zdravlje ljudi, pa je deo publikacija okrenut ka tim analizama, a i da bi se pratilo potrebni su senzori, pa jedan deo publikacija predlaže hardversku implementaciju sistema za monitoring kvaliteta vazduha. Kako je predloženi servis zasnovan na prikupljanju, analizi i vizuelizaciji podataka o kvalitetu vazduha i interakciji sa građanima pametnog grada, u nastavku su navedeni radovi koji su imali najveći uticaj na formiranje rešenja ovog problema.

Autori u radu [84] opisali su zašto je značajno da postoji sistem za praćenje kvaliteta vazduha, kako određeni zagađivači utiču na zdravstveno stanje i kako senzorski podržati takav sistem. U radu [85] predložena je platforma koja bi vršila sakupljanje i vizuelizaciju podataka o zagađenju grada, pri čemu bi se senzori za merenje zagađenja nalazili na pametnim biciklima. U radu [86] opisan je prototip mobilne aplikacije koja služi za obaveštavanje građana pametnog grada o kvalitetu vazduha, na primeru Novog Zelanda. U radu [87] predstavljen je sistem za praćenje zagađenja zasnovan na vebu, ali ima podršku za Android telefone i notifikacije u slučaju prekoračenja praga zagađenja.

U radu [88] opisan je sistem za praćenje zagađenja zasnovan na *IoT* koji ima mogućnost slanja notifikacije ljudima i okolini i odgovarajućoj centrali u slučaju prekoračenja praga zagađenja. Semantički slojevi podataka u sistemima za praćenje kvaliteta vazduha za pametne gradove, kao i studije koje dovode u vezu zdravlje i kvalitet vazduha prikazani su u radu [89]. Predlog za vizuelizaciju podataka o kvalitetu vazduha dobijenih sa senzora tokom vožnje automobilom u kom su instalirani, opisan je u radu [12].

Arhitektura sistema koja je bazirana na sensorima i na sakupljanju podatka od korisnika i pruža vizuelizaciju mapa zagađenja data je u radovima [90] [91]. Rad [36] opisuje sistem za praćenje kvaliteta vazduha i predikciju, namenjen pametnom zdravstvu. Značaj *BigData* u sistemima za praćenje kvaliteta vazduha prikazan je u radu [92].

U Srbiji postoji vizuelizacija podataka o zagađenju vazduha, ali je broj lokacija i gradova pokrivenih vizuelizacijom veoma mali, pri čemu nikakva interakcija sa građanima nije obezbeđena [82].

### **5.2.2.3. Opis i arhitektura servisa**

Predloženo rešenje objedinjuje prednosti predloženih rešenja opisanih u radovima iz oblasti. Podaci se sakupljaju sa senzora koji imaju fiksnu lokaciju u infrastrukturi pametnog grada. Sistem pruža podatke i na veb stranici i na mobilnom telefonu ili drugom pametnom uređaju. Ali je i interaktivan, tj. iako prima informacije od korisnika, takođe šalje notifikacije korisnicima. Takođe, nudi mogućnost poziva u pomoć sa date lokacije, pri čemu se aktivira medicinska služba za hitne situacije. Znajući lokaciju korisnika i eventualno njegove privatne podatke, omogućava se efikasnije pružanje pomoći u hitnim slučajevima.



U slučaju kada ne postoji informaciono-komunikaciona infrastruktura pametnog grada namenjena praćenju zagađenja i povezana sa zdravstvom, mogućnosti građana su sledeće:

- 1) građanin ima ili uopšte nema onlajn informaciju o kvalitetu vazduha u svom gradu;
- 2) ukoliko ima informacije, mora sam zaključiti o tome da li je kvalitet vazduha loš u njegovoj neposrednoj blizini;
- 3) ukoliko građanin uoči neki problem ili kritičnu situaciju, ne zna kome tačno da se obrati, a da informacija stigne do svih kojima je neophodna;
- 4) ako građanin dođe u lošu zdravstvenu situaciju, sam mora zvati hitnu pomoć (ili neko iz njegove okoline), a zatim ih obavestiti o svojoj lokaciji, i ličnim podacima i trenutnom stanju.



Slika 22. Slojevita arhitektura servisa za praćenje zagađenja vazduha

Predloženi servis za praćenje kvaliteta vazduha počiva na pametnoj infrastrukturi grada koja prikuplja podatke sa senzora o kvalitetu vazduha sa odgovarajućih lokacija u pametnom gradu. Arhitektura se sastoji od tri sloja kako je prikazano na slici Slika 22.

Sloj senzora obuhvata senzore pametnog grada, u ovom konkretnom slučaju senzore za merenje parametara koji utiču na kvalitet vazduha i pametni uređaji kod građana koji aktivno učestvuju u inicijativi pametnog grada. Ovi pametni uređaji su istovremeno i davaoci geolokacije sa koje građanin izveštava o aero incidentu ili traži pomoć, npr. sloj senzora obuhvata i humane senzore, građane kao najpametnije entitete u pametnom gradu. Ovi humani senzori su značajni izvori informacija koji se slivaju u aplikacije pametnog grada, u konkretnom slučaju pametno zdravstvo.

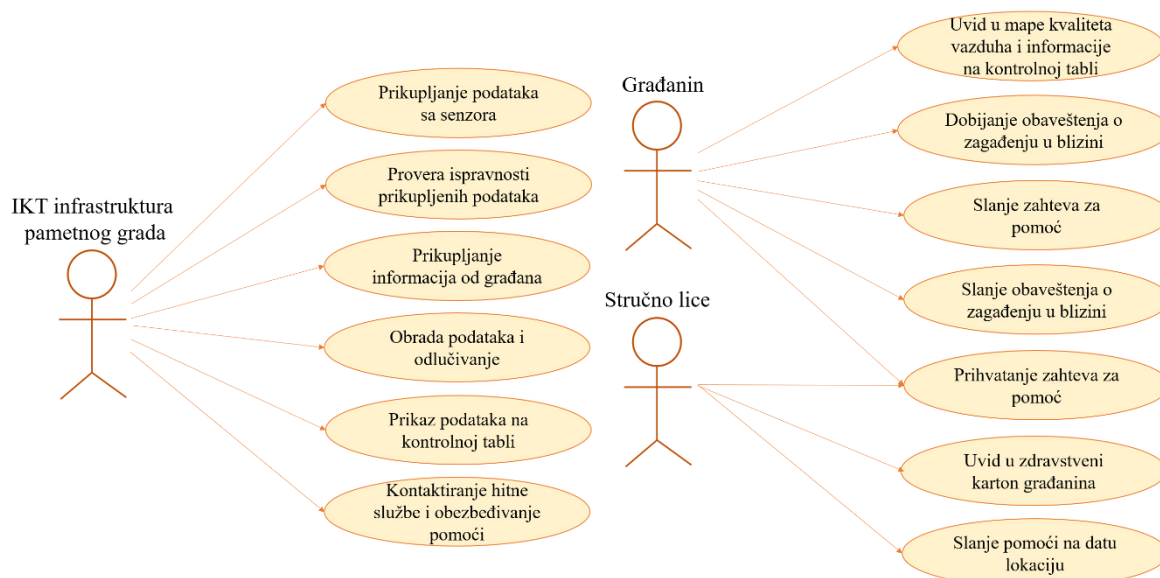
Sloj socijalnih mašina služi za prihvatanje podataka od senzorskog sloja, njihovo organizovanje u klastere po određenim pravilima i prosleđivanje odgovarajućoj aplikaciji pametnog grada. Socijalna mašina je konvergencija tri vizije: socijalnog softvera, ljudi kao jedinica za obradu i softvera koji se može socijalizirati [83]. Ovaj sloj odvaja heterogenost senzora i aplikacija (njihovih poslovnih procesa) čije je funkcionisanje direktno zavisno od podataka sa senzorskog sloja, i uspostavlja povezivanje socijalnih i poslovnih artifakata [93]. Poslovni artifakti su definisani kao konkretno prepoznatljivi i samoopisujući deo informacije koji ljudi upotrebljavaju da odrade dodeljeni posao [94][95]. Npr. slog iz medicinskog kartona pacijenta pomaže lekaru da propiše adekvatnu terapiju.

Sloj klad (engl. *cloud*) infrastrukture sastoji se od adekvatnih aplikacija koje podržavaju inicijative pametnog zdravstva u nekom gradu. Ovaj sloj bavi se analizom podataka sa senzora koji stižu iz senzorskog sloja na osnovu interakcije socijalnih mašina sa odgovarajućim sensorima i građanima koji koriste servis pametnog grada. Ovaj sloj sistema na osnovu ulaznih podataka vrši i odlučivanje, npr. da li nekome treba poslati notifikaciju ili pomoć. Ovim slojem se obezbeđuje i klasifikacija u koju grupu trenutne vrednosti zagađenja spadaju i njihova vizuelizacija. Mape zagađenosti dostupne su svima na info-tabli veb portala ili mobilnoj aplikaciji kod građana. U slučaju opasnosti građanima se upućuje upozorenje, a u slučaju zahteva za pomoć adekvatna medicinska podrška.

Scenario korišćenja predložene platforme je opisan *UML* dijagramom prikazanom na slici Slika 23, koji opisuje sledeće scenarije:

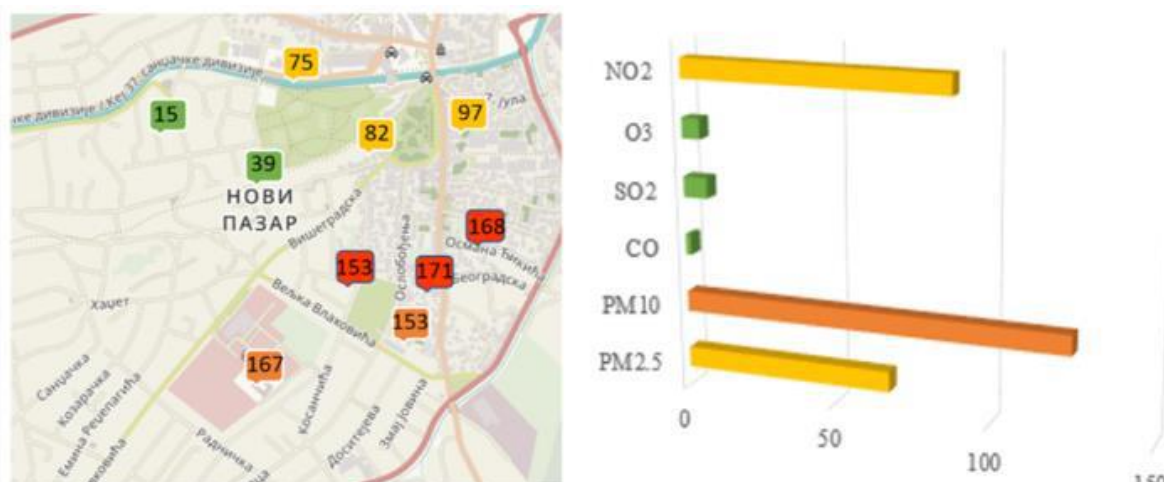
- Slučaj # 1 Obaveštenja: Zagađenje vazduha u nekoj oblasti poraslo je preko praga. Aplikacija sadrži geolokacije svojih pretplatnika i obaveštava sve pretplatnike u zagađenom području.

- Slučaj # 2 Poziv za zdravstvenu pomoć: Ako građanin, pretplatnik usluge, ima problema sa disanjem, može zatražiti pomoć pritiskom na posebno dugme dostupno na klijentskoj strani aplikacije i taj poziv se automatski prosleđuje zdravstvenoj službi. Koristeći lične podatke koje je građanin ostavio tokom pretplate, geolokacije i druge podatke na kontrolnoj tabli, zdravstvena služba može da pošalje pomoć hitne pomoći.
- Slučaj # 3 Izveštavanje o iznenadnim zagađenjima vazduha: Ako je građanin primetio kritičnu situaciju (npr. požar), tu situaciju je moguće prijaviti. Nakon obrade ovog obaveštenja, ove informacije se objavljuju na kontrolnoj tabli i šalju upozorenja svim pretplatnicima u području incidenta.

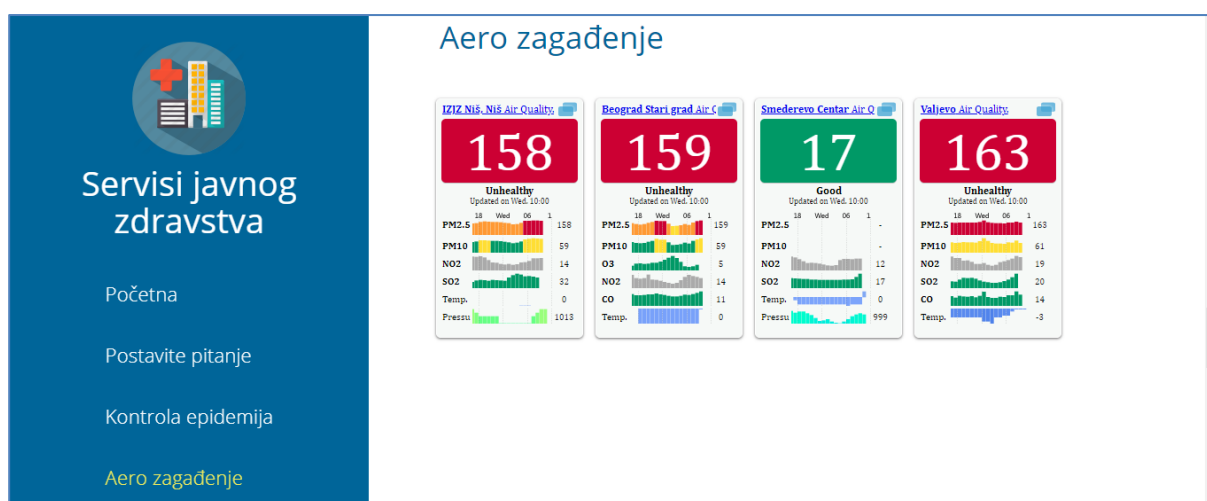


Slika 23. Dijagram slučajeva korišćenja za servis za praćenje kvaliteta vazduha

Na slici Slika 24 prikazan je predlog za vizuelizaciju mape neposrednog okruženja korisnika i vrednosti IKV-a koje pokazuju senzori koji se nalaze na označenim lokacijama u blizini. Boja ukazuje na stepen zagađenja i njegovu vezu sa uticajem na zdravlje stanovnika. Na slici Slika 25 prikazana je realizacija vizuelizacije zagađenja korišćenjem dostupnog *AQP API*-ja [127].



Slika 24. Predlog izgleda mape zagađenja i vizuelizacija vrednosti zagađivača



Slika 25. Vizuelizacija rešenja korišćenjem dostupnog AQP API-ja

Postojeća rešenja koja se bave praćenjem kvaliteta vazduha ne nude sveobuhvatni sistem koji prikuplja podatke i sa senzora i od građana, koji analizira, klasifikuje, vizualizuje, šalje notifikacije građanima, ili pomoć, tj. povezivanje građana sa hitnom medicinskom službom. U odnosu na postojeća rešenja u našoj zemlji, ovaj sistem je upotrebljiv, ali je za njegovu realizaciju potrebno da se senzori postave na što većem broju lokacija, kao i edukacija i dobrovoljnost građana da koriste i učestvuju u funkcionisanju samog sistema. Dalje unapređenje ovog servisa vodiće se u pravcu personalizacije servisa za različite profile korisnika (zdrav čovek, dete, plućni bolesnik, težak plućni bolesnik itd.). U zavisnosti od profila korisnika, servis će obaveštavati korisnike za različite nivoe vrednosti parametara kvaliteta vazduha.

### 5.2.3. Unapređenje pružanja medicinskih usluga

Svedoci smo aktuelne pandemije bolesti *Covid-19* izazvane novim korona virusom, koja je uticala na brojne delatnosti i povećala prisustvo korišćenja onlajn servisa, kako u zdravstvu, tako i van njega. Razlog za ovo je bio i savet izbegavanja fizičkih kontakata. Prognoze su da se slične pandemije mogu dešavati i u budućnosti, ali i u uslovima bez pandemija, u brzom životu u urbanim sredinama, pogodno je iskoristiti mogućnosti koje nude IKT i odraditi elektronskim putem poslove za koje bi nam bili potrebni sati čekanja u redovima. Problemi sa kojima se sreću pacijenti godinama unazad, danas u slučaju preopterećenosti zdravstvenog sistema, samo su kulminirali.

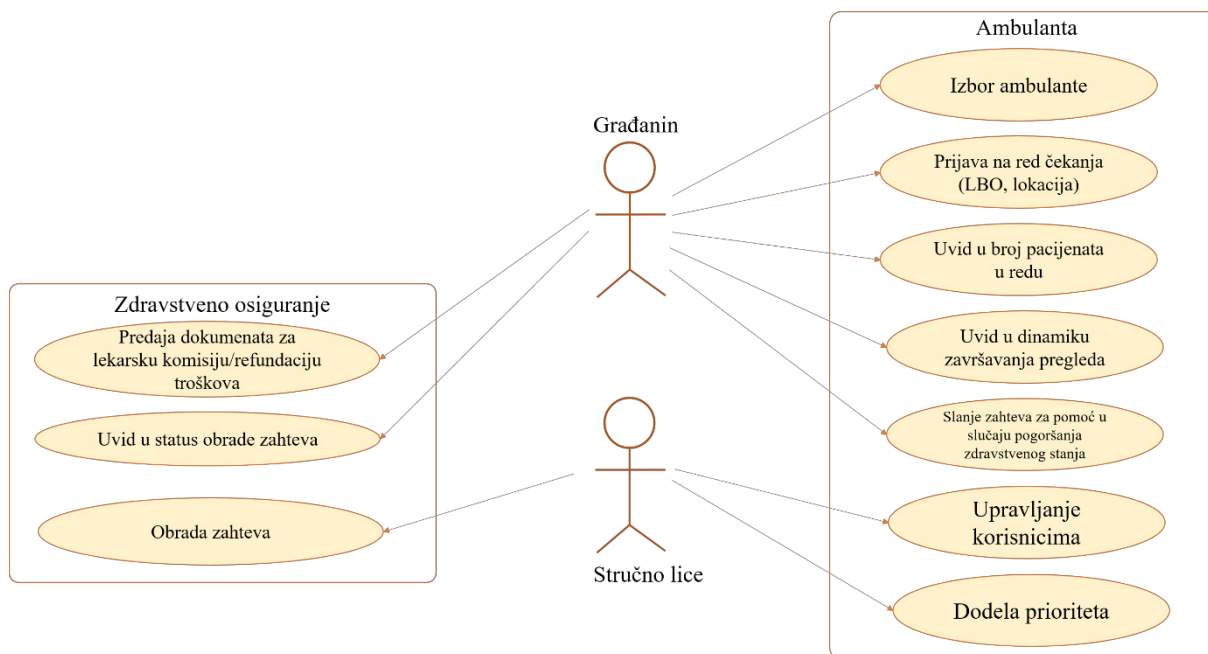
U manjim gradovima često lekari upućuju pacijente da u zdravstvenim centrima u većim gradovima obave specijalističke preglede ili određene intervencije, za čije obavljanje ne postoje uslovi u manjim zdravstvenim centrima. Stoga se daju uputi koje mora odobriti lekarska komisija, često putne troškove pokriva RFZO, a predaja i uvid u status zahteva se trenutno obavlja bez upotrebe elektronskih servisa.

Postavlja se pitanje šta se još može uraditi kako bi se smanjile bespotrebne administrativne stavke, dugotrajna čekanja u redovima, bespotrebne posete lekaru kako bi se pokazali rezultati i dobio savet, predala dokumenta lekarskoj komisiji i sl.

Na slici Slika 26 prikazani su slučajevi korišćenja servisa za unapređenje pružanja medicinskih usluga. Ovde su slučajevi korišćenja svrstani u dva entiteta, prvi se odnosi na ambulante, trijaže, ordinacije, laboratorije i slične ustanove gde se formiraju redovi čekanja uprkos zakazivanju pregleda, a nekada je zakazivanje nemoguće (kao što je bio slučaj tokom aktuelne COVID-19 pandemije). Ovaj servis ima dva tipa korisnika, a to su građanin i stručno lice.

Građanin ima sledeće mogućnosti u okviru entiteta označenog kao ambulanta: izbor ambulante za čiji se red čekanja prijavljuje, elektronsko prijavljivanje čekanja u redu korišćenjem LBO i lokacije kako bi se sprečile zloupotrebe, uvid u broj pacijenata u redu, uvid u dinamiku završavanja pregleda, alarmiranje stručnog osoblja u slučaju pogoršanja stanja. Stručno lice u ovom entitetu ima mogućnost upravljanja prijavljenim korisnicima, i promenu prioriteta (u slučaju alarmiranja od strane nekog građanina kome se pogoršalo

zdravstveno stanje). Stručno lice ima LBO broj pacijenta koji čeka u redu, pa samim tim i uvid u to koje bolesti pacijent ima i može mu promeniti prioritet ukoliko je u pitanju trudnica, stariji građanin ili hronični bolesnik.



Slika 26. Dijagram slučajeva korišćenja servisa za unapređenje medicinskih usluga

U entitetu označenom kao zdravstveno osiguranje, korisniku se ovim servisom obezbeđuje onlajn predavanje zahteva za upućivanje u drugu zdravstvenu ustanovu (koji treba biti odobren od strane lekarske komisije), refundaciju troškova, i onlajn uvid u odobrenje zahteva za lekarsku komisiju ili status refundacije troškova. Stručno lice u ovom entitetu vrši obradu zahteva, odnosno promenu njegovog statusa.

## 5.2.4. Inteligentni transport pacijenata

Pametni uređaji sa mogućnošću sakupljanja podataka konteksta vozila i putnika pružaju mogućnost njihove obrade u svrhu povećanja kvaliteta usluge prevoza. Ukoliko su ove informacije javno dostupne i koriste se za formiranje i korišćenje kolektivne inteligencije u cilju poboljšanja uslova transporta za stanovnike gradova, onda one doprinose oblasti pametnog saobraćaja u konceptu pametnih gradova. U okviru ovog servisa pruža se mogućnost praćenja vibracija u vozilu, kako bi se unapredio nivo udobnosti prilikom vožnje.

Servis sakuplja informacije o vibracijama, a rezultat njihove obrade jesu javno dostupne ocene ruta prema kriterijumu udobnosti prilikom vožnje. Ove informacije mogu biti od posebnog značaja za prevoz pacijenata i specifičnih supstanci i predmeta u zdravstvu ali isto tako i u prevozu dece, prevozu opasnih materija (lakozapaljive, toksične, eksplozivne). Drugi deo ovog servisa sadrži informacije o parkinzima i nudi informacije o najbližem parkingu, olakšavajući na taj način parkiranje vozila hitne pomoći.

Sistemi inteligentnog transporta odnose se na sve tipove saobraćaja (drumski, vodeni, vazdušni) i podrazumevaju postojanje raznih vrsta navigacionih i komunikacionih uređaja za razmenu podataka među vozilima ili između vozila i neke fiksne lokacije. Oni omogućavaju dobijanje informacija o najbržim, najjeftinijim i najsigurnijim rutama.

Ovaj servis pruža mogućnost beleženja nivoa vibracije u vozilima koja utiče na udobnost putnika, i čuvanje, obradu i prikaz podataka o potencijalno kritičnim lokacijama na putu. Sakupljanje podataka vrši se korišćenjem mobilne aplikacije za Android operativni sistem, a njihova obrada korišćenjem veb sistema, i kao rezultat se dobijaju mape udobnosti odgovarajućih relacija puta. Ovaj servis rezultat je sinergije pametnog transporta i pametnog zdravstva, a njegov cilj je pružanje boljeg kvaliteta, sigurnijeg i efikasnijeg prevoza pacijenta.

#### **5.2.4.1. Stanje u oblasti**

U ovom delu dat je pregled radova koji se bave poboljšanjem transporta pacijenata korišćenjem IKT. Autori rada datog u [97] bave se problemom izbora optimalnog izbora vozila za pacijente, uzimajući u obzir različite parametre, sa naglaskom na pronalaženju ravnoteže između efikasnog izbora i zadovoljstva pacijenta.

U prethodnom radu, kao i u radovima [98][99][100], govori se o prevozu pacijenata u hitnim slučajevima, i o takozvanim *EMS* - hitnim medicinskim sistemima. U radu [101] takođe se predlaže algoritam čiji je cilj efikasno upravljanje prevozom pacijenata, uzimajući u obzir dužinu prevoza, dužinu rute, smanjenje troškova ako ima više zahteva itd. U [102] naglasak je na poboljšanju kvaliteta usluga prevoza pacijenta kada su u pitanju vanredne situacije, ali u ovom slučaju udobnost pacijenta nije uzeta u obzir. Rad [103] se takođe bavi prevozom pacijenata sa posebnim potrebama zbog osetljivosti na bolesti.

Udobnost vozila nije uzeta u obzir u sistemima opisanim u prethodnim radovima, a uzimanjem u obzir ovog parametra prilikom kreiranja servisa kvalitet prevoza pacijenta bio bi značajno poboljšan.

#### **5.2.4.2. Sistem za beleženje i obradu podatka o vibracijama**

U okviru inteligentnih transportnih sistema, sistemi koji pružaju efikasne rute mogu pozitivno uticati na negu pacijenta. Naime, ako imamo pacijenta sa povredama koje zahtevaju posebnu negu i osetljivost tokom transporta, takvi sistemi mogu da nam preporuče rutu do odredišta da izaberemo, uzimajući u obzir udobnost u vozilu tokom vožnje ovom rutom.

Da bismo dobili informacije o udobnosti vožnje na ruti, moramo da merimo vibracije vozila tokom vožnje, a to možemo da uradimo pomoću pametnog uređaja sa akcelerometrom, kako je opisano u radovima [104][105]. Prikupljene informacije se šalju na server, gde se obrađuju i tada se beleži nivo komfora za odgovarajuću lokaciju (udoban, srednje udoban, neudoban).

Teško je udobnost kvantitativno predstaviti, jer na nju utiču faktori koji su subjektivnog tipa, kada se radi o transportu putnika. Neadekvatni uslovi tokom transporta mogu nepovoljno psiho-fizički uticati na putnike i na taj način izazvati loše raspoloženje, nervozu, bolove u leđima i zglobovima, mučnine i sl. Kod putnika koji već imaju neke zdravstvene probleme zdravstveno stanje se može pogoršati usled neudobnog transporta. Ovo je posebno važno kod transporta pacijenata. Međutim, i neke tipove materijala je potrebno prevoziti sa većim nivoom udobnosti (lomljivost, opasne/zapaljive materije).

Na udobnost putnika tokom vožnje utiču vibracije sedišta, vibracije ruku, vibracije stopala, buka, dizajn sedišta, temperatura, vlažnost, vazdušni pritisak, razmak između sedišta i drugi. Jedan od najznačajnijih faktora udobnosti je nivo vibracija koje putnik oseća prilikom transporta. Vibroudobnost je direktno srazmerna ubrzanjima koja deluju na telo putnika tokom vožnje. Na udobnost transporta utiču tri faktora: vozač, tip i stanje vozila i stanje puta.

Procena nivoa udobnosti bazira se na merenju vibroudobnosti po standardu *ISO 2631* sa dodatnim proračunima nad izmerenim podacima. U *Android* aplikaciji se može korišćenjem akcelerometra i *GPS* senzora snimiti i izračunati vrednosti *RMS*-a (efektivna vrednost ubrzanja) za sve tri ose, zatim pored navedenih proračuna u svakom intervalu



vremena računa se maksimalna vrednost modula ubrzanja. Za detektovani maksimalni moduo ubrzanja pamte se detektovane vrednosti ubrzanja za sve tri ose, kao geografska širina, dužina i brzina. Takođe u podešenom intervalu vremena računaju se i maksimalne vrednosti ubrzanja po sve tri ose.

#### **5.2.4.3. Predlog i realizacija servisa**

Postoji nekoliko scenarija kako se ovi podaci mogu koristiti za poboljšanje transportnih uslova pacijenata.

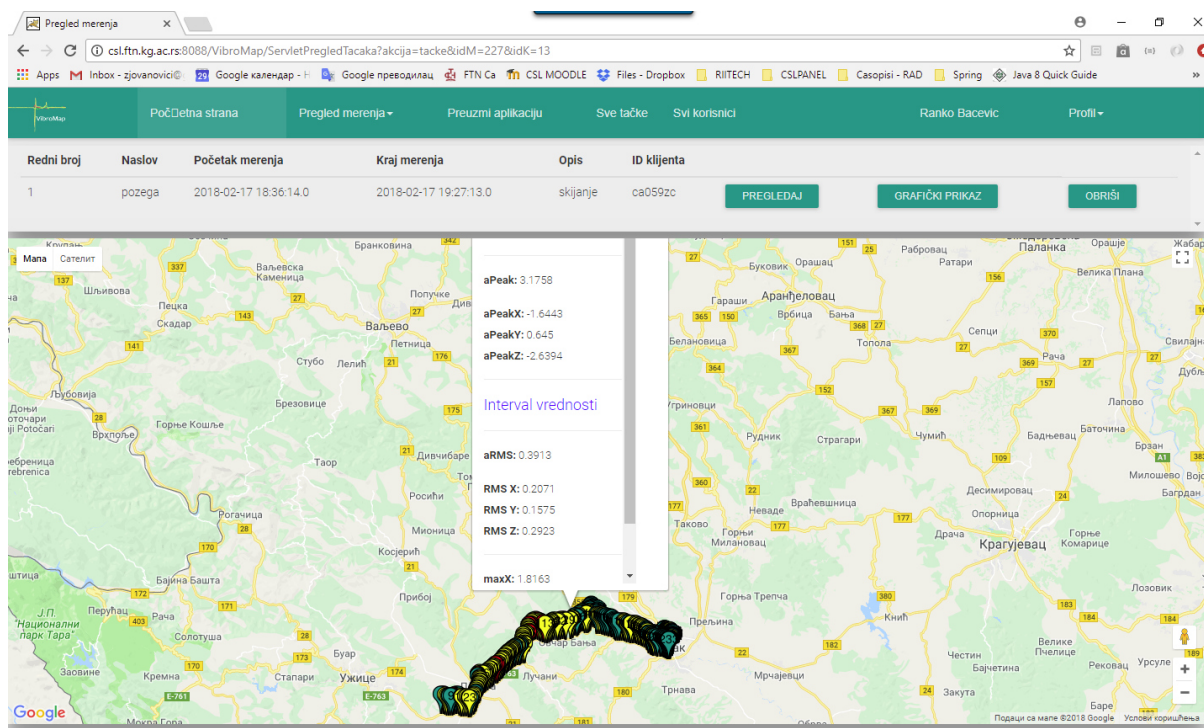
Prvi scenario zasnovan je na podacima o nivou vibracija na datoj lokaciji. Nivo udobnosti za datu lokaciju i ruta na kojoj se vrši merenje može se oceniti na osnovu udobnosti. Ovaj scenario je detaljno opisan u radovima [106][107][108].

Drugi scenario uključuje, pored gore navedenih podataka, i beleženje podataka vozila u kome se vrši merenje. Na ovaj način je moguće uporediti zabeležene udobne rute kada se koriste različita vozila za hitne slučajeve. Ako se zabeleženi podaci u jednom od vozila značajno razlikuju od ostalih, uzrok nelagodnosti može biti samo vozilo, a ne ruta. U ovom slučaju, sistem predlaže najudobnije vozilo i signalizira da postoje vozila kojima će možda biti potrebna popravka.

Treći scenario je snimanje podataka o vozaču, jer na isti način, kao i u gornjem slučaju, to može biti uzrok vibracija (zbog grubog načina vožnje), a ne vozila ili rute. Ovom prilikom sistem preporučuje najodgovornijeg dostupnog vozača na osnovu njegove procene od strane sistema za merenje vibracija.

U svrhu obrade prikupljenih informacija o vibracijama tokom vožnje, koje se koriste za stvaranje javno dostupnih mapa udobnosti, razvijena je klijent server aplikacija, detaljno opisana u radu [109]. Klijentski deo je mobilna Android aplikacija koja se koristi za snimanje i slanje informacija o snimljenim udobnim, manje udobnim ili neudobnim tačkama tokom vožnje. Serverski deo obrađuje ove podatke i kao izlaz generiše mape udobnosti za lokacije na kojima su izvršena merenja (učestalost merenja može da se podešava). Realizacija ovog dela urađena je korišćenjem *Java* veb tehnologija i *API*-ja *Google Maps* za prikaz GIS podataka.



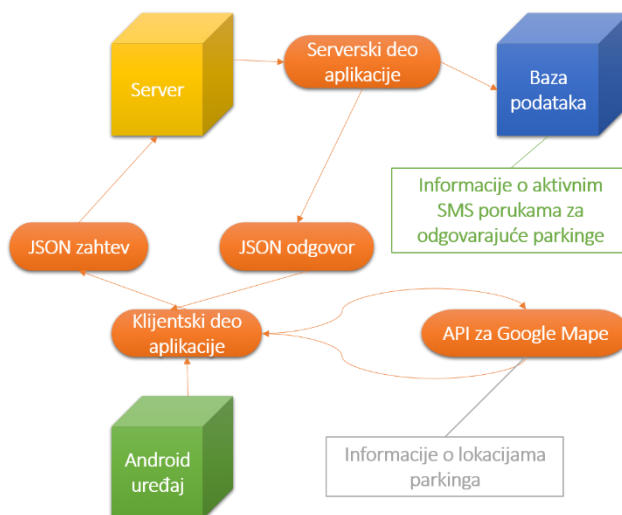


Slika 28. Prikaz jedne kreirane mape udobnosti

#### 5.2.4.4. Servis za pronalaženje najbližeg parkinga

Prilikom rešavanja problema pošlo se od trenutne situacije i problema parkiranja u Novom Pazaru. Naime, u ovom gradu parkinzi su označeni po zonama. Prilikom dolaska na parking, vozač šalje SMS poruku sa brojem registracije vozila i ona aktivira važenje parkinga za dato mesto na sat vremena. Uzimajući u obzir ove podatke, razvijena je aplikacija za pametno parkiranje sa test podacima, koje bi trebalo da budu omogućeni od parkinga servis grada, odnosno IKT infrastrukture pametnog grada. To je ukupan broj parking mesta na svakom od parkinga, i skup poslatih SMS poruka za odgovarajući parking, pri čemu se iz podataka o vremenu pristizanja poruke može izračunati da li se radi o aktivnom parking mestu ili ne [111].

Arhitektura rešenja prikazana je na slici Slika 29. Radi se o klijent-server aplikaciji. Klijentski deo urađen je za pametne telefone sa Android operativnim sistemom.



Slika 29. Arhitektura rešenja aplikacije za pametno parkiranje

Kao što je prikazano na slici Slika 29, serverski deo aplikacije vrši pristup podacima parking servisa za odgovarajuće parkinge, i na osnovu vremena pristiglih *SMS* poruka, računa se koliko je parking mesta u datom trenutku na parkinzima slobodno. Informacije o ukupnom broju parking mesta i pristiglim *SMS* porukama za odgovarajući parking, obrađuju se na serverskom delu aplikacije. U ovom rešenju nema informacija o konkretnom parking mestu koje je zauzeto ili slobodno, ali se i ova informacija može dobiti ukoliko bi te informacije bile javno dostupne, a unete od strane zaposlenog na parkingu ili u slučajevima da se ovi podaci šalju sa senzora namenjenih za ove svrhe.

Klijentski deo aplikacije komunicira sa *Google API*-jem za mape, kako bi se dobili podaci o parkinzima u blizini korisnika koji aktivira zahtev. Da bi ovo bilo omogućeno, potrebno je da se omogući pristup lokaciji korisnika. Podaci koji se razmenjuju u ovom delu aplikacije dati su u *JSON* formatu, i u njima se nalaze koordinate geografske lokacije parkinga, kao i naziv parkinga (Slika 30).

Kombinujući podatke sa klijentskog i serverskog dela aplikacije, korisnik dobija informaciju o udaljenosti parkinga, vremenu potrebnog da se do njega dođe i broju slobodnih mesta za taj parking.

```
{
  "geometry": {
    "location": {
      "lat": 43.1649896,
      "lng": 20.4927057
    },
    "viewport": {
      "northeast": {
        "lat": 43.1663385802915,
        "lng": 20.4940546802915
      },
      "southwest": {
        "lat": 43.1636406197085,
        "lng": 20.4913567197085
      }
    }
  },
  "icon": "https://maps.gstatic.com/mapfiles/place_api/icons/ge...",
  "id": "016a3708e3996156069c7cc5fcdfaa8a24b4c10",
  "name": "Parking Kod Đurđevih Stupova",
  "opening_hours": {
    "open_now": true,
    "weekday_text": []
  },
  "place_id": "ChIJDU0JEKopVkcRgDJHBoi_GAo",
  "rating": 1.7,
  "reference": "CmRRAAAATm1A_Hi4aPoLE5ypsJTm8DJft8trvEbjbPx1ZpG...",
  "scope": "GOOGLE",
  "types": [ "parking", "point_of_interest", "establishment" ],
  "vicinity": "Đurđevi Stupovi, Raška"
},
```

Slika 30. Prikaz dela *JSON* podataka o parkinzima

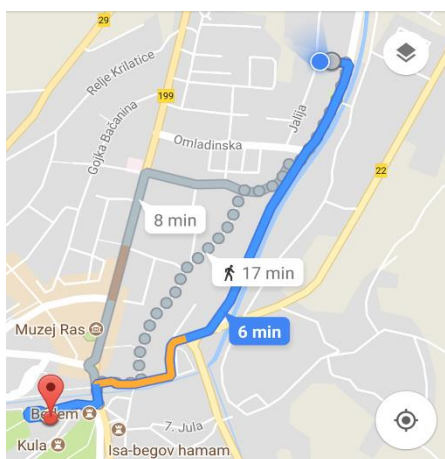
Na slici Slika 31 prikazan je izgled početnog ekrana aplikacije, gde se na osnovu pozicije korisnika određuju najbliži parkinzi.



Slika 31. Prikaz ekrana aplikacije

*Google Maps API* obezbedio je da se filtriraju podaci tipa parking, i elemente njihove lokacije vraća u *JSON* formatu. Pomoću ovih podataka izračunata je udaljenost od geolokacije

pošiljaoca zahteva do odgovarajućih parkinga. U aplikaciji se oni mogu sortirati u odgovarajući redosled. Za svaki parking na serverskoj strani obrađuju se slogovi pristiglih SMS poruka, i računa se na osnovu vremena pristizanja da li su aktivne. Na taj način dolazimo do informacije koliko je zauzetih/slobodnih parking mesta. Takođe je iskorišćena i mogućnost koju daje *Google Maps API*, pa se u datom trenutku može učitati prozor sa prikazom mogućih ruta do izabranog parkinga, kako bi korisnik izabrao najbolje rešenje, kao što je prikazano na slici Slika 32.



Slika 32. Prikaz ekrana aplikacije – moguće rute

Realizacijom ovog servisa obezbeđuje se predikcija parkiranja i simbioza dveju grana pametnih gradova, a to su pametni transport i pametno zdravstvo. Iako se ovaj servis može koristiti nezavisno od servisa pametnog zdravstva, u slučajevima kada se radi o transportu pacijenata ovaj servis može biti od značaja. U scenariju kada neko želi da doveze pacijenta do neke zdravstvene ustanove, pogotovo ako se on teško kreće, a potrebna je hitna intervencija, onda je vrlo bitno da se omogući ostavljanje vozila što bliže zdravstvenoj ustanovi. Rešenje je moguće uz dostupnost relevantnih podataka za unapređenje kvaliteta života stanovništva. Tema daje prostor za dalje istraživanje, pre svega, moglo bi se ići na obeležavanje pojedinačnih parking mesta, zatim u zavisnosti od gužve u saobraćaju vršiti predikcija najoptimalnijeg parkinga, i sl. Primer proširenja sistema pametnog transporta uključivao bi i analizu sentimenta pristiglih poruka sa odgovarajućih geolokacija u cilju detektovanja gužve ili nekog drugog interesantnog događaja u saobraćaju.

### 5.2.5. Mobilni SOS sistem

Servis pametnog zdravstva za mobilni *SOS* sistem u stvari predstavlja deo servisa za komunikaciju, ali implementiran tako da se bez dodatnih ograničenja i na što brži način može poslati zahtev za pomoć na samo jedan klik. Taj klik automatski poziva i alarmira hitnu službu. Na ovaj način hitna služba odmah ima uvid u informacije o pacijentu koje ima i prilikom standardnog servisa komunikacije. Postojanje ovog servisa značajno je i kao komponenta servisa za praćenje zagađenja vazduha, opisanog u poglavlju 5.2.2. Zamišljeno je da na nosivom uređaju, kao što je sat ili narukvica (Slika 33), postoji *SOS* dugme kojim se aktivira mobilni *SOS* sistem. Određuje se lokacija korisnika koji zove u pomoć, šalje se najbliže vozilo hitne pomoći, zdravstvenim radnicima iz vozila sugeriraju se stanje iz istorije pacijenta.



Slika 33. Primer nosivog uređaja za svrhe pametnog zdravstva [145]

### 5.3. Servisi zasnovani na *NLP* metodama i kraudsorsingu

U ovom poglavlju opisani su servisi koji su zasnovani na metodama veštačke inteligencije, konkretno na istraživanju teksta i procesiranju prirodnih jezika (engl. *Natural Language Processing, NLP*). Izvor podataka za realizaciju ovih servisa su dokumenta napisana na prirodnim jezicima, odgovori građana dobijeni kraudsorsingom i elektronski medicinski izveštaji iz medicinskih informacionih sistema.

Iako se na istraživanje podataka (engl. *data mining*) i istraživanje teksta (engl. *text mining*) često gleda kao na komplementarne analitičke procese koji analizom podataka rešavaju poslovne probleme, oni se razlikuju u pogledu vrste podataka koje obrađuju. Dok istraživanje podataka obrađuje strukturirane podatke - visoko formatirane podatke (kao što su npr. baze podataka), istraživanje teksta se bavi nestrukturiranim tekstualnim podacima - tekstom koji nije unapred definisan ili organizovan na bilo koji način (npr. postovi društvenih medija) [112][113][114][115].

Druga razlika je u tome kako istraživanje podataka i istraživanje teksta pristupaju analizi. Nijedno od njih nije jedinstvena tehnologija, već koriste širok spektar funkcija za pretvaranje dostupnih podataka u znanje. Istraživanje podataka kombinuje discipline koje uključuju statistiku, veštačku inteligenciju i mašinsko učenje nad strukturiranim podacima. Neke od funkcija modeliranja podataka su navedene u nastavku:

- Asocijacija - određuje koliko je verovatno da će se jedna pojava dogoditi u odnosu na drugu pojavu tokom vremena. Na primer, u prodajnim transakcijama funkcija asocijacije može otkriti obrasce kupovine kupaca koji kupuju mleko prilikom kupovine žitarica.
- Klasifikacija - otkriva obrasce koji se koriste za predviđanje klase u koju će podaci spadati. Na primer, vremenska predviđanja o tome da li će biti sunčano ili oblačno u zavisnosti od vremenskih uslova.
- Klasterovanje - organizuje podatke identifikujući sličnosti i grupišući ih u klustere da bi se identifikovale nove činjenice o tim podacima. Na primer, segmentacija tržišta je jedna od njegovih primena.
- Regresija - predviđa numeričku vrednost u zavisnosti od promenljivih na datom skupu podataka. Na primer, cene polovnog automobila s obzirom na njegovu kilometražu i druge promenljive uslove.

Platforme za analitiku i poslovnu inteligenciju mogu brzo da identifikuju i preuzmu informacije iz velikih skupova strukturiranih podataka i primene ove funkcije istraživanja podataka da bi kreirali modele koji omogućavaju njihovu deskriptivnu, prediktivnu i preskriptivnu analizu. S druge strane, istraživanje teksta zahteva dodatni korak uz zadržavanje istog analitičkog cilja kao i kod istraživanja podataka. Istraživanje teksta se bavi



nestrukturiranim podacima, tako da pre nego što se primeni bilo koja funkcija modeliranja podataka ili prepoznavanja obrazaca, nestrukturirani podaci moraju biti organizovani i strukturirani na način koji omogućava njihovo modeliranje i analizu. To zahteva sofisticirane statističke i lingvističke tehnike kako bi se mogla analizirati široka paleta nestrukturiranih formata tekstualnih podataka i kako bi se svaki dokument obogatio metapodacima, poput autora, datuma, sažetka sadržaja itd. Ovaj proces je obično povezan sa tehnikom veštačke inteligencije (engl. *artificial intelligence, AI*) koja se naziva procesiranje prirodnog jezika i omogućava sistemu da razume značenje podataka na ljudskom jeziku. Krajnji cilj *NLP*-a je čitanje, dešifrovanje, razumevanje i nalaženje smislenosti u prirodnom jeziku. Većina *NLP* tehnika oslanja se na mašinsko učenje da bi se zaključilo o značenju podataka na prirodnim jezicima .

Kao jedan od izvora podataka u servisima koji će biti opisani u ovom poglavlju, navode se elektronski medicinski izveštaji. Elektronski medicinski zapis ili izveštaj (engl. *EMR* ili *EHR - electronic medical (health) report*) [116] je izveštaj pacijenta koji čuva podatke o zdravstvenom stanju, dijagnozama i terapijama. *EMR*-ovi se kreiraju pomoću softvera koji se nalazi u zdravstvenim ustanovama. *EMR*-ovi sadrže privatne podatke o pacijentu (ime, prezime, lični broj, datum rođenja, broj kartice, broj osiguranja, adresu itd.), beleške lekara, dijagnoze, laboratorijske izveštaje, terapije itd. I ovi izveštaji mogu sadržati strukturirane i nestrukturirane podatke, napisane na prirodnom jeziku, pa se njihovom obradom može dobiti znanje koje je, u ovom slučaju, primenjeno u vizuelizaciji podataka o epidemiji.

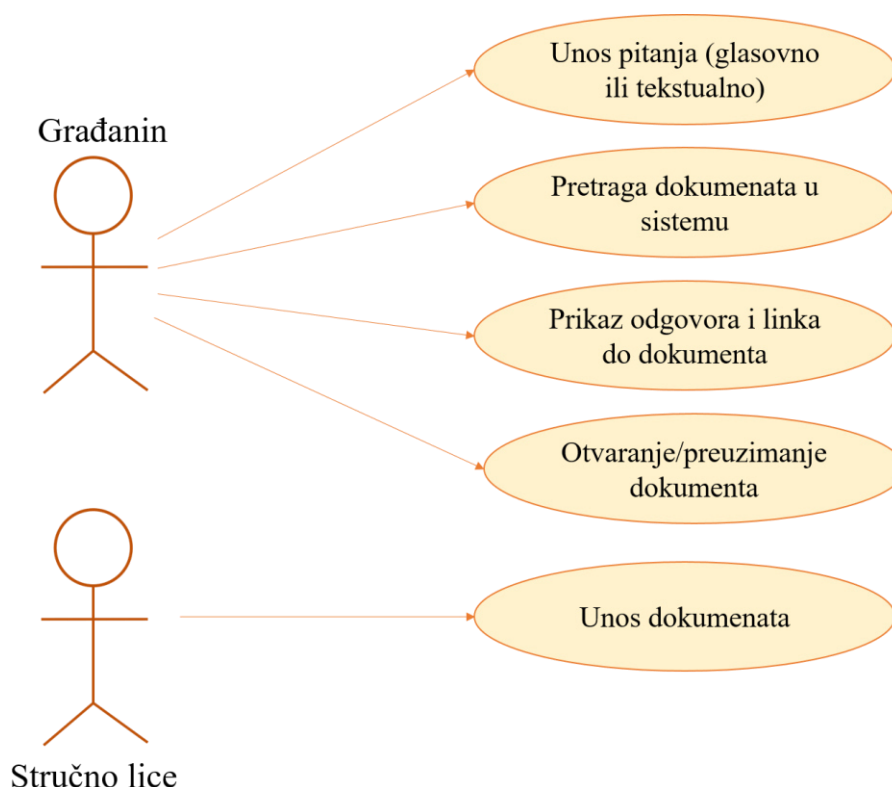
Poslednji u nizu opisanih servisa u ovom poglavlju (uvid u raspored i ocenjivanje lekara) nije direktno baziran na *NLP* metodama, već spada u grupu servisa koji je zasnovan na podacima dobijenim kraudsorsingom.

### **5.3.1. Automatsko odgovaranje građanima na pitanja**

Servis pametnog zdravstva za automatsko odgovaranje građanima na pitanja / pretragu medicinskih dokumenata trebalo bi da obezbedi korisniku da unese pitanje i da na to pitanje

dobije odgovor na osnovu sličnosti sa rečenicama u medicinskim dokumentima koje su unete u sistem od strane stručnih lica.

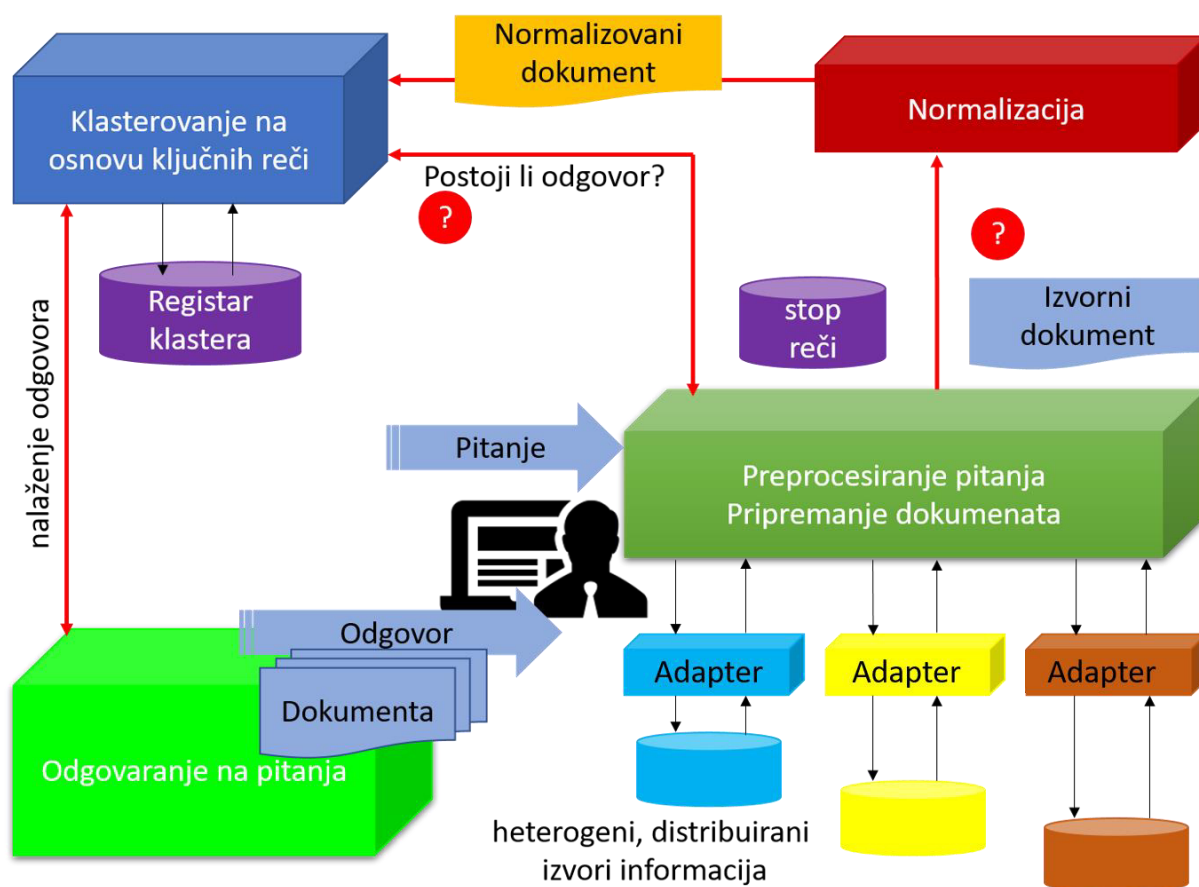
Mogućnosti ovog servisa prikazane su *UML* dijagramom slučajeja korišćena na slici Slika 34.



Slika 34. Dijagram slučajeja korišćenja za servis za automatsko odgovaranje građanima na pitanja

Arhitektura sistema za pretraživanje dokumenata za potrebe servisa pametnog zdravstva data je na slici Slika 35. Ovaj servis treba da pomogne građaninu da u svakom trenutku može brzo da dobije odgovor na svoje pitanje. U sistemu postoje tri vrste sadržaja: pitanja, formalni dokumenti i odgovori stručnjaka. Pitanja i dokumenti su na osnovu ključnih reči grupisani. Kada korisnik postavi novo pitanje računa se njegova sličnost sa pitanjima i dokumentima u izabranoj grupi. Priprema dokumenata za obradu, normalizacija, izdvajanje ključnih reči, grupisanje dokumenata i pitanja kao i pronalaženje odgovora na postavljeno pitanje se nalaze na serverskoj strani aplikacije i predstavljaju posebne komponente sistema. Detaljan opis svake od komponenti je dat u nastavku. Opis je dat hronološki i svaka naredna

komponenta se služi rezultatima prethodne, kako bi konačan rezultat, odgovor na postavljeno pitanje bio dat [77].



Slika 35. Pojednostavljeni prikaz komponenti sistema za pretraživanje dokumenata

### 5.3.1.1. Pripremanje dokumenata

Ova komponenta priprema dokument za sledeći korak – normalizaciju dokumenata. Priprema dokumenta se izvršava kad se ubacuje novi dokument (ili pitanje) u sistem. Ovaj dokument dalje prolazi i kroz ostale komponente kako bi bio spreman za pretragu. Pre nego se novi dokument ili pitanje ubace u bazu podataka oni moraju biti normalizovani. Da bi se dokument ili pitanje normalizovali, potrebna je priprema za normalizaciju i to je zadatak komponente za preprocesiranje pitanja. Priprema dokumenata obuhvata: promenu formata dokumenta, uklanjanje suvišnih i neformalnih znakova i strukturiranje dokumenta po pravilima koja odgovaraju sledećem koraku – normalizaciji. Priprema dokumenta u ovom slučaju obuhvata sledeće korake:

- Obrada dokumenata i u ćirilicnom i u latinićnom pismu,

- Obrada dokumenata u HTML formatu,
- Uklanjanje suvišnih znakova (emotikoni, sleng, ...)
- Uklanjanje stop reči,
- Prebacivanja između ASCII, UTF8 i latiničnog pisma za potrebe pripreme.

Na kraju, dokument posle normalizacije treba da bude u latiničnom pismu u UTF8 formatu.

### 5.3.1.2. Normalizacija dokumenata

Normalizacijom teksta podrazumeva se transformacija teksta u neki drugi oblik koji je pogodniji za neku vrstu kompjuterske obrade. U našem slučaju to je pretraživanje. Svrha normalizacije reči je oslobađanje od suvišnih modifikacija reči koje ne unose izmene u značenju reči, odnosno svođenje tih modifikacija na zajednički, osnovni oblik. Normalizacija se može vršiti: lematizacijom reči - otklanjanje nastavaka za oblik i tvorbenih nastavaka i svođenje reči na lemu, otkidanjem najdužeg pronađenog nastavka za tvorbu oblika za odgovarajući tip reči, izdvajanjem  $k$  prvih slova iz reči,  $n$ -gram analizom reči itd. Svaka od navedenih normalizacija ima svojih prednosti i mana. Za neke su potrebni veliki i specifični leksički resursi. Normalizacija medicinskog teksta detaljno je opisana u poglavlju 6.

### 5.3.1.3. Klasterovanje dokumenata

U ovoj komponenti grupisanje dokumenata se vrši na osnovu sličnosti sadržaja. Svaki dokument je predstavljen ključnim rečima koje su nosioci značenja dokumenta. Automatizacija izdvajanja ključnih reči je proces izdvajanja malog skupa reči i fraza koje opisuju sadržaj dokumenta. Izdvajanje ključnih reči u dokumentu zahteva postojanje korpusa dokumenata koji su sličnog sadržaja korišćenjem algoritama za indeksiranje  $tf*idf$ . U našem slučaju dokument koji se ubacuje nepoznatog je sadržaja tako da se za izdvajanje ključnih reči ne može koristiti korpus. Izdvajanje ključnih reči se vrši na osnovu njihovog pojavljivanja u rečenicama dokumenta i u celom dokumentu. Dokument koji se obrađuje je normalizovan  $n$ -gram analizom. Za osnovni skup ključnih reči izdvajaju se najfrekventniji  $n$ -grami. Dokument se sastoji iz rečenica i pojavljivanje termina se računa i na nivou rečenice. Izdvojene ključne reči su osnova daljeg traženja ključnih reči i izraza. Ukoliko se neki  $n$ -gram javlja češće u rečenicama sa određenim ključnim rečima onda on bliže određuje te ključne reči i kandidat je takođe za ključnu reč.

Za izdvajanje ovih  $n$ -grama računa se matrica uzajamnog pojavljivanja izdvojenih ključnih reči i ostalih  $n$ -grama u rečenicama u dokumentu. Matrica sadrži broj rečenica u kojima se neka ključna reči( $q$ ) i  $n$ -gram( $w$ ) javljaju zajedno. Računa se odstupanje ovog broja od očekivane vrednosti pomoću hi-kvadrat testa. Očekivana vrednost se računa kao proizvod  $nw$ , ukupnog broja reči u rečenicama u kojima se pojavljuje posmatrani  $n$ -gram( $w$ ) i  $pq$ , procenta ukupnog broja reči u rečenicama u kojima se pojavljuje ključni  $n$ -gram( $q$ ) u celom dokumentu, formula 1 [119].

$$\chi^2(w) = \sum_{g \in G} \frac{(freq(w, q) - n_w * p_q)^2}{n_w * p_q} \quad (1)$$

Ukoliko ima nekih grupa reči koje se često koriste zajedno njihovi  $n$ -grami će biti izdvojeni. Ovim algoritmom izdvajaju se  $n$ -grami ključnih reči čija je osnova duža od 4 slova, kao i  $n$ -grami fraza (grupe reči koje se koriste zajedno i tako imaju specifično značenje) koje se često koriste.

Ključne reči (frazе) čine osnovni skup potreban za grupisanje dokumenata i pitanja. Grupisanje dokumenata i pitanja se vrši algoritmom klasterovanja  $c$ -sredina [118]. Za razliku od dobro poznatog algoritma  $k$ -sredina ovaj algoritam omogućava da dokument ili pitanje pripada više klastera [117][119].

#### 5.3.1.4. Odgovaranje na pitanja (pronalaženje dokumenata)

Krajnji rezultat celog sistema je odgovor koji se dobija na postavljeno pitanje. Odgovor može biti u formi kratkog odgovora ili čitavog dokumenta koji sadrži odgovor. Zadatak ove komponente je da pronade najsljednije postojeće pitanje pitanju koje je postavljeno. Ako takvo najsljednije pitanje ne postoji ili korisnik zahteva da dobije ceo dokument kao odgovor, onda je potrebno pronaći dokument koji ima odgovor na postavljeno pitanje. Pitanja i dokumenti se traže u klasteru koji je po ključnim rečima najbliži postavljenom pitanju. Ispitivanje sličnosti između dokumenata je izvršeno pomoću algoritma kosinusne sličnosti [120]. Za izračunavanje vrednosti sličnosti, potrebni su nam podaci o broju pojavljivanja reči u pitanju (novom pitanju koje se postavlja) u poređenju sa brojem

pojavljivanja svih reči u pitanju/dokumentu (koji već postoje).  $Tf(t)$  predstavlja broj pojavljivanja datog termina  $t$  u razmatranom dokumentu, dok  $idf(t)$  se određuje na osnovu kompletnog koprusa i opisuje korpus kao celinu, a ne pojedinačne dokumente. Izračunava se primenom formule 2, gde je  $N$  broj dokumenata u korpusu, a  $df(t)$  broj dokumenata koji sadrži termin  $t$ .  $TfIdf(t)$  je statistička mera koja pokazuje koliko je termin  $t$  relevantan u dokumentu koji pripada korpusu dokumenata i računa se po formuli 3. Ove vrednosti su nam potrebne da bismo reči u njihovom pojavljivanju u pitanju/rečenici zamenili brojem, odnosno rečenicu ili pitanje u vektor, kako bi se kasnije izračunala njihova kosinusna sličnost.

$$idf(t) = 1 + \log \frac{N}{df(t)} \quad (2)$$

$$TfIdf(t) = tf(t) \times idf(t) = tf(t) * \log \frac{N}{df(t)} \quad (3)$$

$$Cosine Similarity (A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}} \quad (4)$$

Formula 4 pokazuje kako se računa kosinusna sličnost (*Cosine Similarity*), gde  $A$  i  $B$  predstavljaju dva vektora reči predstavljenih vrednostima  $TfIdf$ . Korišćenjem algoritma kosinusne sličnosti i normalizacije dokumenta koja je prethodno opisana, moguće je definisati sličnost između dva pitanja ili sličnost između pitanja i dokumenta. Postavljeno pitanje najpre treba da se normalizuje. Zatim se određuje koji klaster pitanja i dokumenata mu je najbliži i ostatak pretrage se vrši na nivou ovog klastera. Posle toga, potrebno je proći kroz svaku rečenicu svakog postojećeg dokumenta i izračunati kosinusnu sličnost pitanja sa tom rečenicom. Ako sličnost postoji, vrednost algoritma kosinusne sličnosti će biti veća od nule. U tom slučaju, rečenica se ubacuje u niz sličnih rečenica a dokument se ubacuje u niz sličnih dokumenata, na istoj poziciji indeksa u nizu na koji je rečenica upisana u svom nizu. Na kraju se primenjuje sortiranje u opadajućem redosledu po polju kosinusne sličnosti. Podaci o broju reči u dokumentu se čuvaju u bazi podataka i trenutku kada se dokument ubacuje u bazu. Ovi podaci su potrebni za računanje  $TfIdf$  vrednosti [120].



Slika 36. Prikaz realizacije servisa za pretragu dokumenata / odgovaranje na pitanja

Na slici Slika 36 prikazana je realizacija klijentskog dela servisa za pretragu dokumenta, odnosno odgovaranje na pitanja građana. Pitanje je moguće uneti direktno sa tastature ili glasovnim putem, gde se glas prepoznaje i konvertuje u tekst, korišćenjem *Speech To Text API*-ja [121], za automatsko prepoznavanje govora na srpskom jeziku. Moguće je i kliknuti na neko od najčešćih pitanja, i pretraživati njega. Kao rezultat dobija se dokument koji sadrži odgovor na pitanje.

Sa serverske strane moguće je da stručno lice unese dokument u *pdf*, *txt*, *doc* i *docx* formatu, nakon čega se dokumenti normalizuju i čuvaju u bazi podataka u obliku pogodnom za pretraživanje. Prilikom unosa dokumenta, automatski se izdvajaju ključne reči za dokument, međutim, one se mogu uneti i od strane stručnog lica koje unosi dokument, kako bi se on dodatno obogatio meta podacima.

### 5.3.2.Kontrola epidemije

U vreme epidemije je potrebno da ljudi imaju informacije o epidemiološkoj situaciji u svom mestu iz proverenih izvora, kako bi se ponašali u skladu sa istom. Nepostojanje ovakvih informacija može dovesti do širenja vesti iz neobjektivnih izvora, nesavesnog ponašanja građana, ali i širenja panike. Kontrola epidemija je jedna od tema u oblasti pametnog zdravstva u okviru pametnih gradova. Na osnovu izveštaja koji se svakodnevno skladište u medicinskim informacionim sistemima mogu se izvući razne informacije o epidemiološkoj

situaciji u gradovima tj. nekom delu grada koji gravitira konkretnom domu zdravlja. U ovoj sekciji dat je predlog servisa pametnog zdravstva namenjenih kontroli epidemija, kao i načini za njihovu realizaciju. Dati su i primeri realizacije na osnovu elektronskih medicinskih izveštaja iz DZ Niš, iz perioda epidemije morbila.

Posebna oblast u okviru pametnog zdravstva je posvećena kontroli epidemija. Postojanje elektronskih podataka o zdravstvenom stanju građana uz metodologije za njihovo korišćenje i kolektivnu inteligenciju grada radikalno poboljšavaju kompetentnost države u otkrivanju i kontroli epidemija. Vitalni znaci, aktivnosti i lokacija građana mogu se iskoristiti za otkrivanje eventualno novih slučajeva tokom epidemije, efikasno prepoznavanje mesta pojačanog rizika i efikasno upravljanje epidemijom. Takve metode se takođe primenjuju i u otkrivanju i organizaciji drugih široko rasprostranjenih zdravstvenih rizika kao što je, na primer, zagađenje vazduha [122][11].

U toku pisanja ove disertacije aktuelna je pandemija korona virusa *COVID19*, i svedoci smo koliko vesti na ovu temu interesuju javnost, koliko se govori o poklapanju javnih podataka sa onima iz informacionih sistema i sl. Na osnovu postojećeg informacionog sistema u Srbiji, *covid.rs*, dobijaju se podaci o broju umrlih, broju zaraženih i broju pacijenata na respiratoru za dan. Do 10.6.2020. bilo je moguće pratiti epidemiološku situaciju po gradovima, ali nakon toga ne. S druge strane, kriva epidemije najčešće se prati na sajtovima kao što su *Worldmeters* i *Endcoronavirus.org* itd [123][124][125].

Motivacija za kreiranje ovog servisa proizašla je iz činjenice da su postojeći servisi javnog zdravstva za kontrolu epidemija u Srbiji nedovoljni i da bi mogli pružati građanima više od onoga što danas pružaju. Doprinos realizacije ovih servisa je u identifikaciji relevantnih funkcionalnosti koje bi sistemi pametnog zdravstva za kontrolu epidemija trebalo da obezbede i predlog, dizajn i razvoj platforme pametnog zdravstva za kontrolu epidemija i njenih najznačajnijih servisa. Platforma je namenjena radu u realnom vremenu, direktno uzima podatke iz medicinskog informacionog sistema i *EMR*-ova i daje veliki broj detalja, kao što je rasprostranjenost virusa po starosnoj strukturi, zdravstvenim ustanovama i pregled broja zaraženih u odgovarajućem vremenskom intervalu.



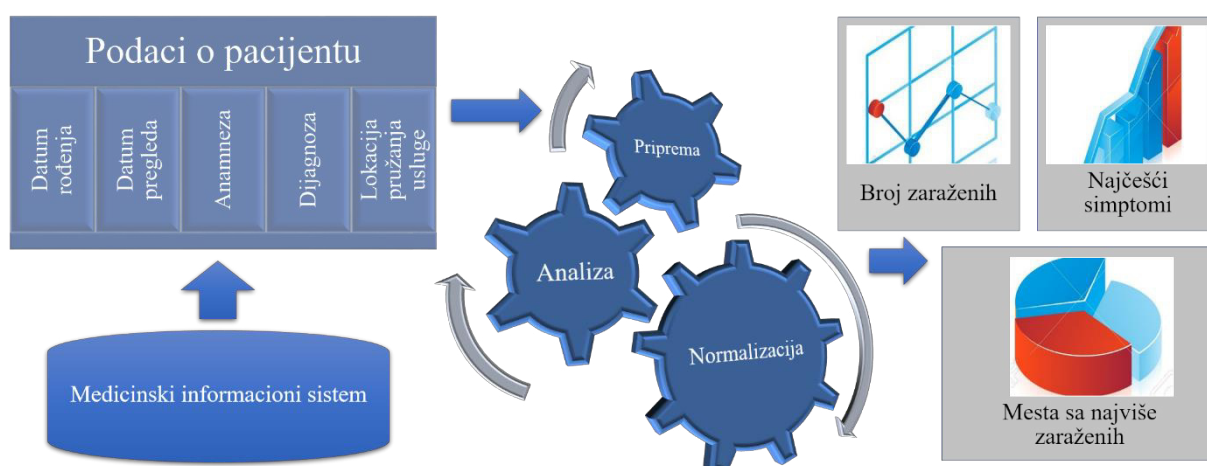
### 5.3.2.1. Stanje u oblasti

Platforme pametnog zdravstva koje rešavaju različite probleme korišćenjem informacionih tehnologija opisane su u brojnim radovima. U radu [126] opisan je pametni model zdravstvene zaštite sa pristupom modela sistema podrške odlučivanju u javnoj zdravstvenoj službi. U radu [127] opisani su servisi pametnog zdravstva bazirani na *IoT* u cilju podrške građanima u slučajevima zdravstvenih problema usled zagađenja vazduha, a u [23] je opisana platforma pametnog zdravstva koja omogućava brojne servise za interakciju s pacijentima.

Radovi na temu upotrebe informacionih tehnologija u prevenciji i kontroli epidemije variraju od onih čiji je cilj pronalaženje matematičkog modela i simulacije širenja epidemije [128], do onih koji opisuju platforme sa ciljem informisanja građana o epidemiji, i korišćenja tehnologije za praćenje lokacije pacijenata [129]. Platforma za kontrolu epidemije bazirana na *IoT* opisana je u radu [4]. U radu [130] daje se predlog nove paradigme za kontrolisanje virusa *Chickungunya*. Ova paradigma ima prednosti poput pravovremenog prepoznavanje infekcije virusom koja prati pacijente i smanjenja širenja infekcije, poboljšanja u procesu analize, praćenje ogromnog broja pacijenata bez obzira gde se nalaze, održavanje odeljenja zdravstvene zaštite i drugih agencija za rane preventivne mere za kontrolu virusa, benefite za lekare itd. U radu [131] analizirane su razlike u upotrebi informacionih tehnologija za kontrolu epidemije u Kini i zapadnim zemljama. Servisi koji su našim građanima dostupni za praćenje informacija o epidemiji virusa *Covid19* pomenuti su u prethodnom poglavlju.

### 5.3.2.2. Predlog i realizacija servisa

Na osnovu podataka dostupnih iz skupa *EMR*-ova [132], moguće je kreirati nekoliko servisa koji bi omogućili uvid u epidemiološku situaciju u realnom vremenu i bili validno sredstvo za informisanje građana. Ovi servisi bazirani su na medicinskoj analizi podataka iz medicinskih izveštaja. Oni se mogu prikazati na portalima javnog zdravlja (Slika 37) i uključivati sledeće slučajeve korišćenja (Slika 38):



Slika 37. Servis za kontrolu epidemija

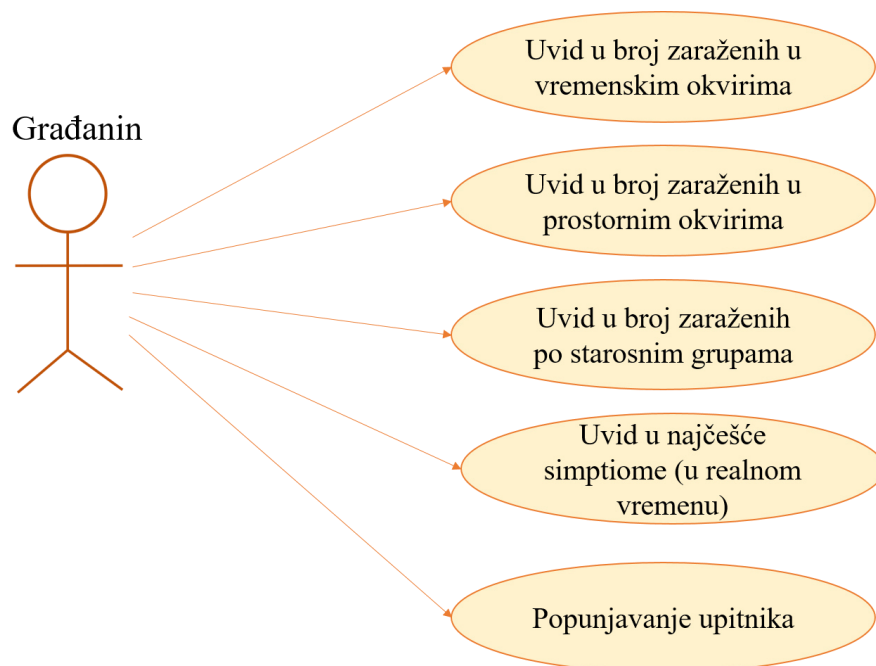
a) Izveštaj od tome koliko ljudi ima specifičnu dijagnozu u gradu, na dnevnom, nedeljnom i mesečnom nivou. Na osnovu ovog servisa, posetilac portala bio bi svestan o toku epidemije u njegovom gradu, i da li se broj obolelih povećava ili smanjuje. Ovaj izveštaj kreira se na osnovu analize podataka dobijenih unosom u strukturirano polje za datum medicinske usluge, na osnovu podataka iz *EMR*-ova.

b) Izveštaj o broju obolelih sa specifičnom dijagnozom posebno po zdravstvenim stanicama u okviru grada. Ovim bi građanin bio upoznat sa time koliko ima obolelih u njegovoj neposrednoj blizini. Ovaj izveštaj kreira se na osnovu analize podataka dobijenih unosom u strukturirano polje za lokaciju medicinske usluge, na osnovu podataka iz *EMR*-ova.

c) Izveštaj o prisustvu bolesti u različitim grupama na osnovu godina starosti. Na osnovu ovoga, građanin bi se informisao o tome da li on ili neko od njegovih bližnjih spada u rizičnu grupu pogođenu virusom. Ovaj izveštaj kreira se na osnovu analize podataka dobijenih unosom u strukturirano polje za datum rođenja pacijenta, na osnovu podataka iz *EMR*-ova.

d) Izveštaj o najzastupljenijim simptomima kod ljudi kojima je ustanovljena odgovarajuća dijagnoza. Na osnovu ovog izveštaja, građanin bi se više informisao o bolesti i mogao da prepozna simptome i u tom slučaju kontaktira lekara. Ovaj izveštaj kreira se na osnovu analize podataka dobijenih unosom u nestrukturirano polje za ostavljanje komentara lekara, na osnovu podataka iz *EMR*-ova. U ovom polju obično su beleške lekara o simptomima, laboratorijskim analizama, terapijama. Uz adekvatnu obradu ovog teksta, može se doći do simptoma i oni se mogu prikazati na dijagramu.

e) Upitnik o trenutnom zdravstvenom stanju pacijenta (prisustvo/odsustvo simptoma) kako bi proverio da li da se javlja u medicinsku stanicu radi lečenja ili ne.



Slika 38. Dijagram slučajeve korišćenja servisa za kontrolu epidemija

Svi ovi servisi omogućili bi građanima da budu stalno u toku sa epidemijom i preduzmu mere za izbegavanje ili lečenje bolesti, a na taj način će se smanjiti posledice epidemije.

Prve tri usluge zahtevaju analizu i vizuelizaciju podataka, što nije zahtevno, dok četvrta zahteva specifičnu obradu teksta za pravilno obeležavanje simptoma. Analizirajući nestrukturirani deo korišćenih *EMR*-ova, registrovane su skraćenicе, pravopisne pogreške, različiti oblici reči i sinonimi za iste simptome. Anamneza bi trebalo da se očisti od reči koje nemaju značenje, a reči od značaja treba da se svedu na isti oblik. Skraćenicе takođe treba da se obrađuju i čuvaju. Postoje i negativni simptomi u anamnezi, tako da služba ne bi pokazala istinski broj pacijenata koji imaju simptom da negacija ne bude uzeta u obzir.

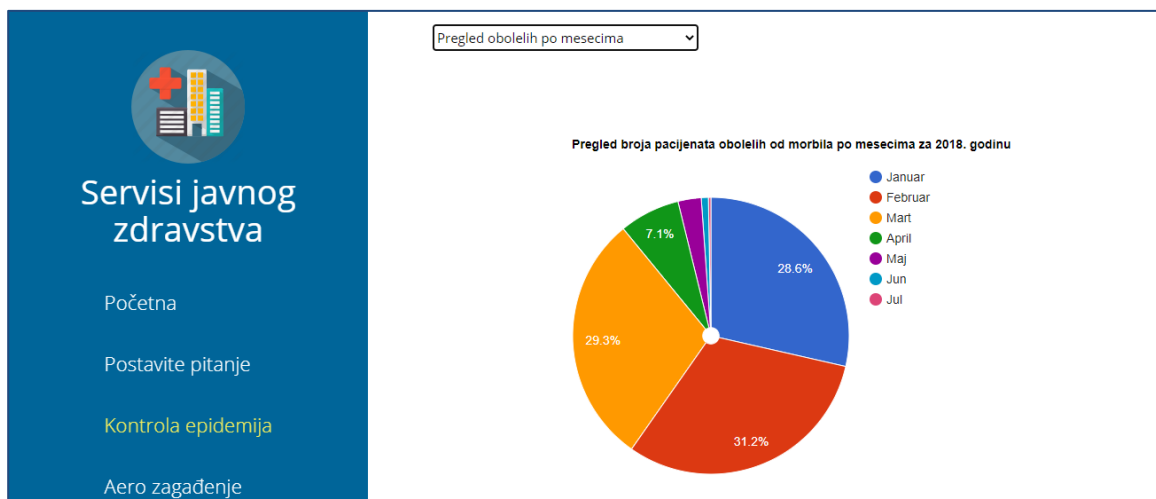
Specifičnost jezika na kojem je izveštaj napisan dodatno usložnjava proces normalizacije. Medicinska evidencija je osetljiva na istraživanje zbog poverljivosti podataka koje sadrže, pa se mora izvršiti odgovarajuća medicinska identifikacija medicinske dokumentacije i ukloniti svi lični podaci pacijenata, kao i lekara.

Za ekstrakciju informacija u slobodnom tekstu *EMR*-a potrebne su tehnike *NLP*-a. Koraci koji su neophodni za ekstrakciju informacija, u našem slučaju simptoma, za potrebe realizacije četvrtog slučaja korišćenja su:

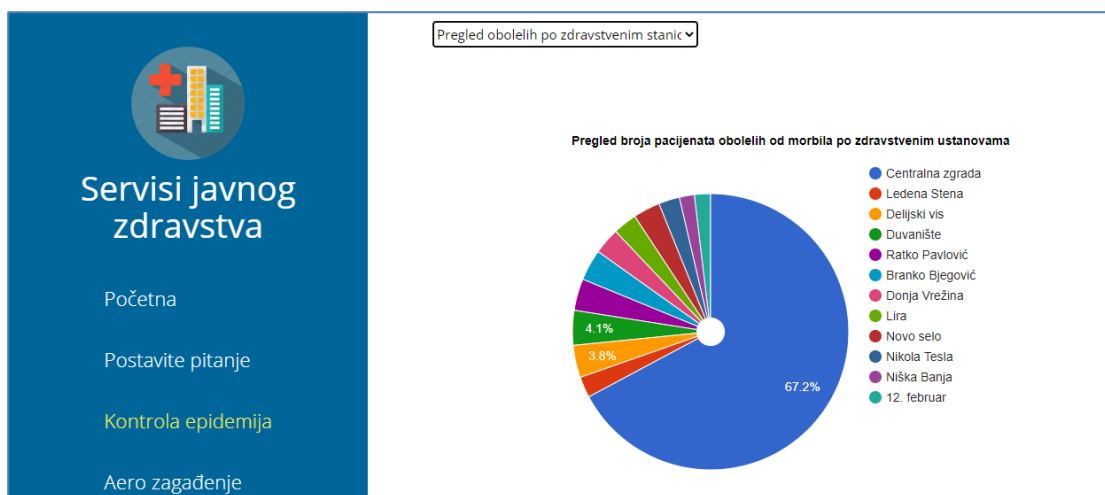
- Svođenje na jedno pismo, obrada skraćenica i tokenizacija. Kako se radi o *EMR*-ovima na srpskom jeziku, slobodni tekst se može naći napisan na latinici i ćirilici. Kako bi se izvršila transformacija podataka u standardni format, u ovom koraku se tekst prevodi na latinicu, uz posebno vođenje računa o slovima sa dijakritičkim simbolima. Nakon toga se vrši obeležavanje skraćenica, a zatim podela rečenice na tokene.
- Brisanje stop reči – u ovom koraku eliminišu se reči koje nisu nosioci značenja. One se nalaze u resursu;
- Određivanje negacije – u medicinskim izveštajima koristi se svega nekoliko simbola negacije, pa se oznaka negacije pridružuje reči u neposrednoj blizini, kako bi se označilo eventualno odsustvo simptoma;
- Svođenje na osnovu – budući da srpski jezik ima bogatu gramatiku, pa se reči mogu naći u raznim oblicima, potrebno je svođenje na osnovu. U odsustvu morfološkog rečnika, kao i za brže rezultate, osnova može biti prefiks dužine  $n$  ili stem;
- Klasifikacija. Nakon preprocesiranja teksta, može se izvršiti klasifikacija. Klasifikacija se može izvršiti korišćenjem metoda mašinskog učenja, tagera, ali i metoda zasnovanih na pravilima, ako pritom imamo označene simptome, u modelu podataka [133].

Detaljno su ovi koraci normalizacije medicinskih izveštaja na srpskom jeziku opisani u radu [134] i u poglavljima 6 i 7.

Servis za kontrolu epidemija prikazuje stanje o broju pacijenata koji imaju odgovarajuću dijagnozu, kojih su godina, gde su se javili na pregled, i koje su simptome imali. Za vizuelizaciju grafikona korišćena je *JS* biblioteka *Google Charts* [135].

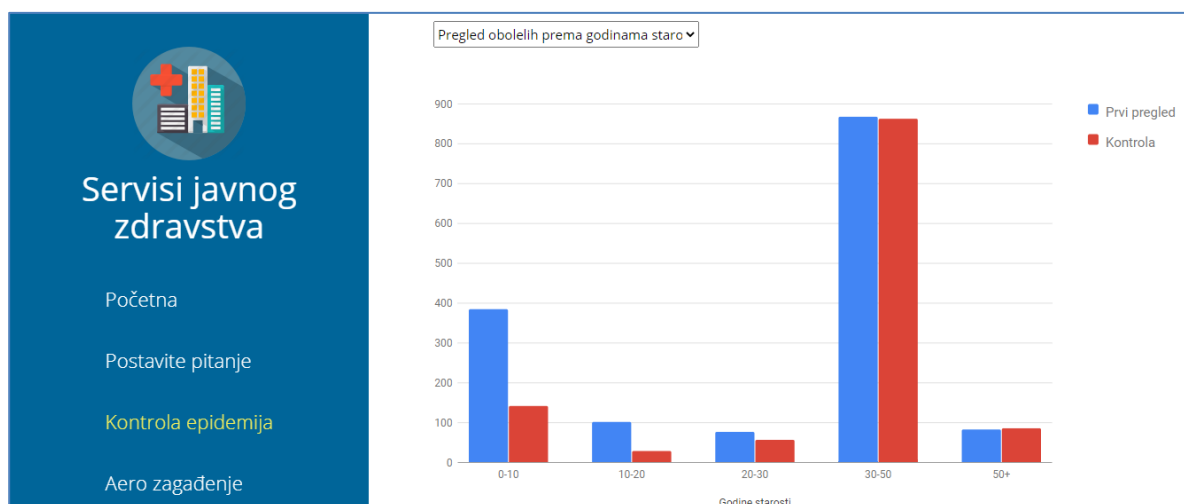


Slika 39. Pregled obolelih po mesecima

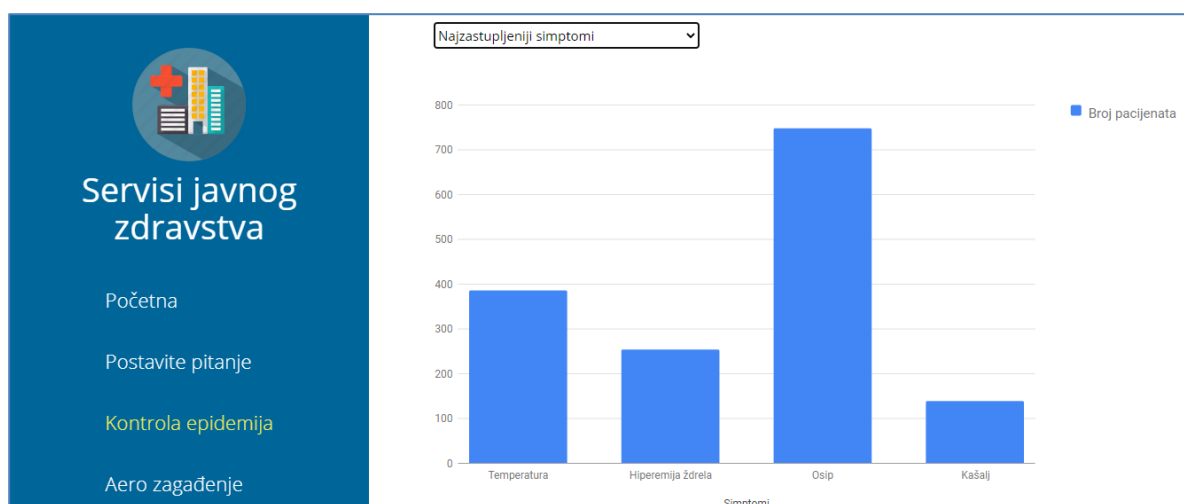


Slika 40. Pregled obolelih po zdravstvenim ustanovama

Za realizaciju ovog servisa potrebno je da deo podataka koji se skladišti u bolničkim *EMR* (elektronski medicinski rekord ili ti karton pacijenta ili istorija bolesti) sistemima bude dostupan kako bi se iz njih izvlačili podaci odnosno odgovarajuće znanje. Slika 40 daje pregled broja obolelih po zdravstvenim stanicama, a Slika 41 na osnovu godina starosti, a na slici Slika 42 prikazan je broj pacijenata sa najčešćim simptomima.



Slika 41. Pregled obolelih po godinama starosti



Slika 42. Pregled zastupljenosti simptoma po broju pacijenata

Realizacijom servisa u okviru koncepta pametnih gradova stvaraju se uslovi za poboljšanje kvaliteta života i zdravlja stanovništva. Realizacija zdravstvenih servisa je bitna komponenta u izgradnji koncepta pametnog grada, posebno u domenu kontrole epidemije, kako bi se podigla svest građana o zaštitnim merama i kako bi pratili stanje epidemije u svom mestu i na vreme se javili lekaru. U ovom poglavlju dat je predlog potrebnih zdravstvenih servisa koji bi omogućili građanima praćenje epidemije u realnom vremenu. Ovi servisi su zasnovani na postojećoj informacionoj infrastrukturi, ne iziskuju specijalna ulaganja, a obezbeđuju stanovništvu mnoštvo proverenih informacija u oblasti praćenja epidemije jer su centralizovane [136]. Na sličan način se može prikazati i stanje vakcinisanosti kod građana.

### 5.3.3. Automatska obrada upitnika i psiholoških testova

Kraudsorsing je pogodan i za primenu u psihološkim istraživanjima jer manje košta od tradicionalnih psiholoških istraživanja, skup populacije koji se obrađuje je širi, podaci se brže skupljaju i obrađuju i nude se dodatne mogućnosti. Cilj ovog pametnog servisa je razvijanje metode za automatsku obradu psiholoških testova u cilju detekcije psiholoških problema. Psihološki testovi često ograničavaju ispitanika nudeći mu mogućnost izbora neke od ponuđenih opcija, bez mogućnosti unosa slobodnog odgovora, jer ga je komplikovano kompjuterski analizirati.

Pristup velikim količinama podataka koji obuhvataju hiljade lekara, stotine hiljada pacijenata i milione pregleda, nudi priliku za povećanje kvaliteta i smanjenje troškova zdravstvene zaštite. U nekim situacijama kolektivna inteligencija može da ima veću vrednost nego mišljenje pojedinačnih eksperata. Psihološka oboljenja predstavljaju osetljivu granu medicine. Sve je veći broj ljudi koji imaju potrebu za stručnom pomoći. Studije tradicionalne psihologije zasnovane su na različitim tehnikama za prikupljanje podataka. Najčešće se prikupljanje podataka vrši putem upitnika, skala procene i sličnih validiranih psiholoških instrumenata, dok su procene psihičkih stanja potpunije ukoliko se vrše putem strukturiranog ili polustrukturiranog intervjua. Ljudi koji žive u 21. veku svedoče uticaju društvenih medija na njihov privatni i javni život. Veliki broj ljudi pokušava da svoje probleme reši putem elektronskih društvenih mreža. Internet donosi veliku količinu raspoloživih podataka, napisanih na bilo kom govornom jeziku. Mnogi psiholozi pretražuju elektronske medije kako bi metaanalizom obuhvatili raspoložive podatke i analizirali iste iz perspektive kliničke i medicinske psihologije.

U vremenu pandemije bili smo svedoci da su upitnici bili od velike važnosti za praćenje epidemije. Primer takvog upitnika je *Test samoprocene* na portalu *e-Zdravlje* [47], i ovaj test dostupan je registrovanim korisnicima (Slika 43), a sadrži pitanja postavljena u cilju identifikacije osoba kod kojih postoji osnovana sumnja za COVID-19 infekciju. Ono što se može uočiti i kod ovog testa je postojanje polja za unos slobodnog teksta za detaljniji opis simptoma, što ukazuje na to da je važno obraditi ovako uneti tekst u upitniku, kako bi se stekla konačna slika o stanju pacijenta.

Za realizaciju ovog servisa potrebno je prikupljanje informacija o psihičkom stanju, kao i o specifičnim reakcijama na odgovarajuće stresne situacije. U okviru ovog servisa pitanja su izražena u obliku standardizovanih psiholoških testova kao i dela za opis stresne situacije i izazvanih osećanja koji su predmet dalje analize. Učestalost sličnih stresnih situacija kao i izazvanih osećanja u poređenju sa rezultatima dobijenim iz psiholoških testova mogu jasnije da ukažu na postojanje odnosno nepostojanje odgovarajućeg psihološkog problema. Čovek najbolje iznosi svoje psihičko stanje na maternjem jeziku, pa je analiza ovih testova prilagođena ljudima koji govore srpski jezik. Kompleksnost obrade tekstova na srpskom jeziku je dodatni izazov.

The screenshot shows the 'eЗдравље' portal interface. At the top, there is a navigation bar with the portal logo and three tabs: 'ИНФО ТАБЛА', 'САНДУЧЕ', and 'ПУНА ВЕРЗИЈА ПОРТАЛА'. Below the navigation bar, the form is divided into several sections:

- Повишена температура:** A checkbox is followed by the text 'Повишена температура'. Below it, a text input field contains the value '36' and a unit selector '°C'.
- Болови:** A section titled 'Болови:' containing four checkboxes with corresponding labels: 'Бол у мишићима', 'Бол у грудима', 'Бол у стомаку', and 'Бол у зглобовима'.
- Symptoms List:** A vertical list of ten checkboxes with labels: 'Општа слабост', 'Дијареја', 'Кашаљ', 'Мучнина - повраћање', 'Упала грла', 'Главобоља', 'Цурење носа', 'Узнемиреност - конфузност', 'Губитак чула укуса', and 'Губитак чула мириса'.
- Друго:** A section titled 'Друго:' with a large empty text input field for additional notes.

Slika 43. Deo *Testa samoprocene* na portalu *e-Zdravlje*

Ovaj servis može se primeniti za razne vrste upitnika. U okviru ovog servisa kao primer upitnika dat je psihološki test namenjen pružanju podrške ženama u toku trudnoće i periodu posle porođaja koje mogu imati neke specifične psihološke probleme. Rođenje deteta donosi radost i blagostanje. Iscrpljenost nakon porođaja, stres i nedostatak sna u prvim nedeljama, preokret ritma života, pa i meseci života novorođenčeta možda nisu uzroci, ali kako pokazuju današnje statistike pospešuju trenutno stanje koje u 10% slučajeva vodi ka postpartalnoj depresiji. Praktični cilj ovog upitnika je da na vreme otkriju ovaj oblik depresije



putem onlajn kraudsorsing platforme za podršku trudnicama i porodiljama u stresnim situacijama. Specifičnost ovog sistema je što je rađen za srpski jezik.

Ovaj servis predstavlja sistem za automatizovano prikupljanje podataka i baziran je na kraudsorsingu. On omogućava da se prikupi velika količina podataka za kratko vreme. Osnovna ideja od koje se polazi u ovom principu jeste dobra volja grupe ljudi da preuzmu na sebe neki posao i obave ga volontirajući ili uz neku naknadu. Stoga kraudsorsing predstavlja moderan način za korišćenje kolektivne inteligencije, čije su prednosti prepoznate u brojnim oblastima, poput rešavanja problema, donošenju odluka i sl.

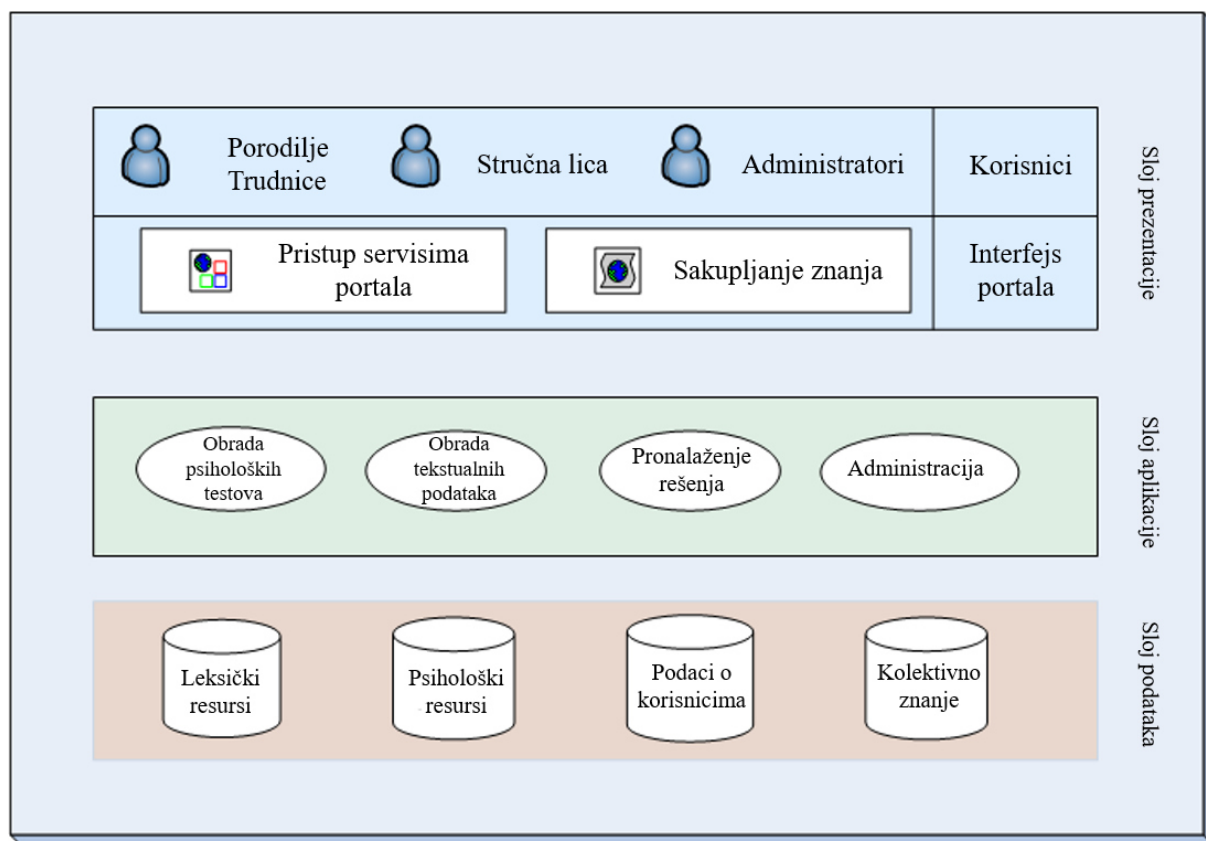
### **5.3.3.1. Stanje u oblasti**

Samostalno dijagnostikovanje (samodijagnostikovanje, engl. *self-diagnosis*) predstavlja zaključivanje o svom zdravstvenom stanju na osnovu medicinskog upitnika. Ovo zaključivanje može biti obavljeno bilo gde i bilo kada u okviru platformi za elektronsko zdravstvo. CPSQ [137] jeste program za samoispitivanje o ponašanju učenika osnovnih škola, napravljen sa ciljem da se detektuju mogući mentalni problemi u ranim stadijumima. AMPQ [138] je test za merenje nivoa brige, anksioznosti, nasilja i opsesije među mladima između 12 i 18 godina. PSS [139] je test koji se koristi za merenje nivoa stresa u periodu od jednog meseca unazad. SBQ-R [140] je test za određivanje da li je pacijent sklon samoubistvu. U radu [5] opisan je kriterijum za merenje nivoa depresije zasnovan na kraudsorsingu, a u okviru platforme za pametno zdravstvo. Nedostatak ovih programa jeste što ne podržavaju automatizaciju i nisu namenjeni za naše govorno područje.

Iskrenost je jako bitna u ovom postupku pa se zato deo u kome korisnik iznosi problem postavlja kao ulazni podatak pre nego što sam sistem na njega izvrši uticaj i smanji objektivnost odgovaranja. Dalje sledi psihološki test kojim se automatski obrađuje i daje procena postpartalne depresije na osnovu Edinburške skale postpartalne depresije (EPDS) [141]. Sistem pronalazi najbližnje situacije i osećanja prethodnih koja ih prate i daje savete koje su psiholozi dali na osnovu ulaznih podataka odgovarajućim korisnicima kao i prikaz povezanosti psihološkog stanja koje je opisano i postpartalne depresije.

Problemi koji se javljaju prilikom obavljanja kraudsorsing zadatka su ljudski faktor i platforma. Razumevanje ljudske motivacije i ponašanja umrežene grupe ljudi je značajno da bi se učesnici motivisali da dobro odrade zadatak. Zato sama platforma mora biti dizajnirana

tako da dodatnim informacijama privuče pažnju učesnica. Ne postoji nijedan sličan portal koji se bavi samodijagnostifikovanjem postpartalne depresije na srpskom pa je utoliko značaj ove platforme veći.



Slika 44. Detaljna arhitektura rešenja

### 5.3.3.2. Predlog i realizacija servisa

Na slici Slika 44. je prikazan detaljan algoritam sistema za samodijagnostifikovanje psiholoških problema kod trudnica i porodilja. Kao što je prikazano učesnici ovog sistema su trudnice i porodilje ali i psiholozi koji daju stručne savete na osnovu iznetih informacija i rezultata testa. Psiholozi ne odgovaraju na svaku pojedinačnu situaciju već počevši od onih koje imaju najmanje sličnih situacijama koje su već obrađene. Na taj način se efikasno koristi stručna pomoć. Svaka korisnica odrađuje i psihološki test na osnovu koga dobija rezultate o postpartalnoj depresiji.

### 5.3.3.3. Obrada psiholoških testova

Psihološki testovi se obrađuju automatski. Za određivanje postpartalne depresije koristi se Edinburška skala pri čemu se za skor veći od 12 uzima da ispitanica ima postpartalnu depresiju. *EPDS* je jedna od najčešće korišćenih skala za procenu depresivnih simptoma kod žena koje su rodile. Porodilje procenjuju težinu za deset različitih depresivnih simptoma u proteklih sedam dana na skali procene od 0 do 3 [142]. Mogući raspon rezultata je od 0 do 30, gde viši skorovi upućuju na težu simptomatiku. Podeoci na skali se odnose na procenu mogućnosti radovanja, doživljaja zabrinutosti, krivice, tuge, zatim procenu o samopovređivanju i slično.

Ova skala procenjuje uglavnom psihološke simptome depresivnosti, dok fiziološke tumači kroz jedan podeok koji se odnosi na procenu mogućnosti spavanja. Osetljivost ove skale je 86%, specifičnost 78%, a ukupna pozitivna prediktivna valjanost 73%. *EPDS* skala je validirana i prevedena u velikom broju zemalja i aktivno se koristi za skrining i dijagnostiku postpartalne depresije.

Trebalo bi napomenuti da je ovo skrining instrument koji se retko kod nas upotrebljava u svrhu dijagnostikovanja povišene depresivnosti. U istraživanju u radu [142], putem ovog instrumenta na slučajnom uzorku pronađeno je da je kod procene stope depresivnosti postpartalno, 29% porodilja imalo skorove iznad 12.

Pored rezultata dobijenih na osnovu testa, ispitanica dobija savet psihologa ukoliko postoji u bazi slična situacija kao i prateće psihološko stanje koje je izazvala. Ukoliko ne postoji, korisniku se daju primeri kako su se drugi osećali u datim situacijama i koliko je takvo stanje povezano sa rezultatima psihološkog testa.

### 5.3.3.4. Obrada tekstualnih podataka

Tekstualni podaci koje korisnici unose su u formatu koji nije pogodan za direktnu obradu i potrebno ga je preprocesirati kako bi bio spreman za mašinsku obradu.

Preprocesiranje je izvršeno u nekoliko koraka:

- Izbor dela teksta (opsega) koji je relevantan

- Tokenizacija – odvajanje diskretnih delova teksta (reči, brojevi, znaci interpunkcije) i uklanjanje suvišnih znakova
- Uklanjanje stop reči - reči koje ne nose značenje (prilozi, predlozi, zamenice, ...)[97].
- Svođenje različitih oblika jedne iste reči na njenu osnovu (stem ili  $n$ -gram)
- Podela teksta na rečenice

Pretraga po sličnosti. Algoritam poređenja stresnih situacija (ili osećanja) po sličnosti obuhvata sledeće korake:

1. Izabrati stresnu situaciju za poređenje;
2. Porediti tekući naziv stresne situacije sa svim ostalim nazivima stresnih situacija iz baze znanja na sledeći način:
  - a. Normalizovati uneti naziv stresne situacije
  - b. Napraviti vektor reči od izabranog naziva stresne situacije
  - c. Uporediti vektor sa svim ostalim vektorima naziva stresnih situacija po sličnosti, koristeći algoritam kosinusne sličnosti
  - d. Ako je rezultat poređenja izabranog i tekućeg vektora naziva stresnih situacija veći od nule, dodati tekući naziv stresne situacije u niz sličnih naziva stresnih situacija.
3. Ponoviti korak 2 za opise stresnih situacija.

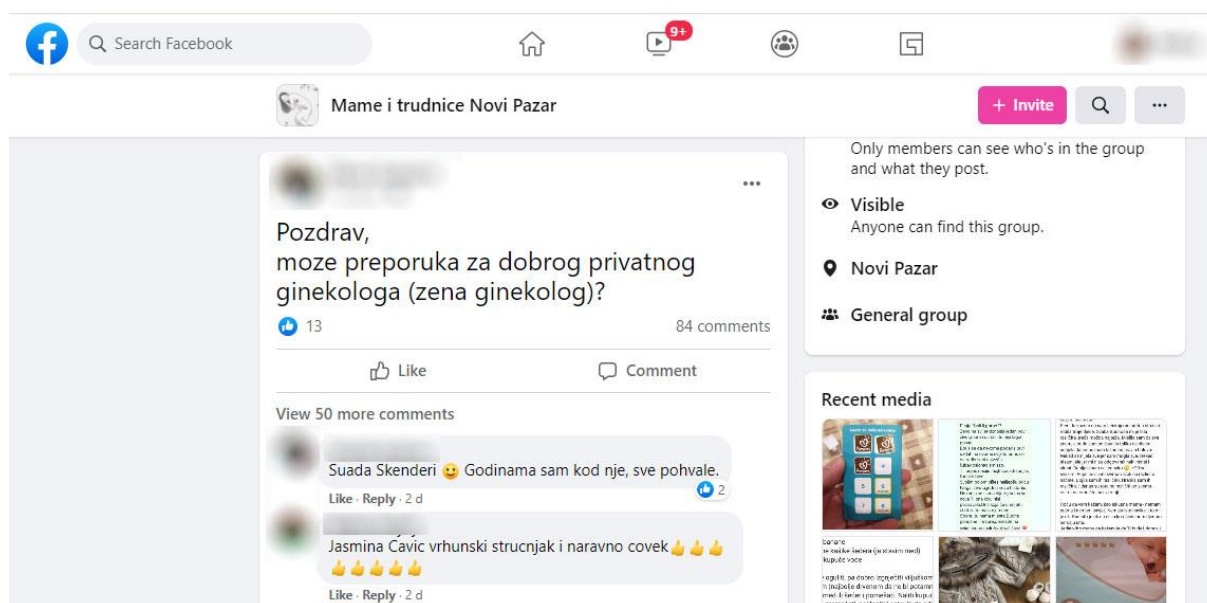
Screenshot of a web application for psychological tests. The left sidebar is blue with a white cross icon and the text "Servisi javnog zdravstva". Below it are menu items: "Početna", "Postavite pitanje", "Kontrola epidemija", "Aero zagađenje", and "Psihološki testovi". The main content area is white and titled "Psihološki testovi". It features a dropdown menu with "Test na postporodajnu depresiju". Below the dropdown is a text prompt: "Molim Vas označite odgovore koji najbolje opisiju kako ste se osećali proteklih 7 dana, a ne samo kako se osećate danas." Underneath is the instruction "U poslednjih sedam dana:". There are three numbered questions, each with four radio button options. Question 1: "Mogla sam se smežati i videti smešnu stranu stvari" with options "Kao i inače", "Ne tako često", "Sigurno manje nego inače", and "Ne uopšte". Question 2: "Radovala sam se stvarima unapred:" with options "Kao i inače", "Ne tako često", "Sigurno manje nego inače", and "Jedva". Question 3: "Nepotrebno sam sebe krivila kada bi stvari krenule loše:" with options "Da, većinu vremena", "Da, veći deo vremena", "Ne baš često", and "Ne, nikada".

Slika 45. Prikaz realizacije upitnika

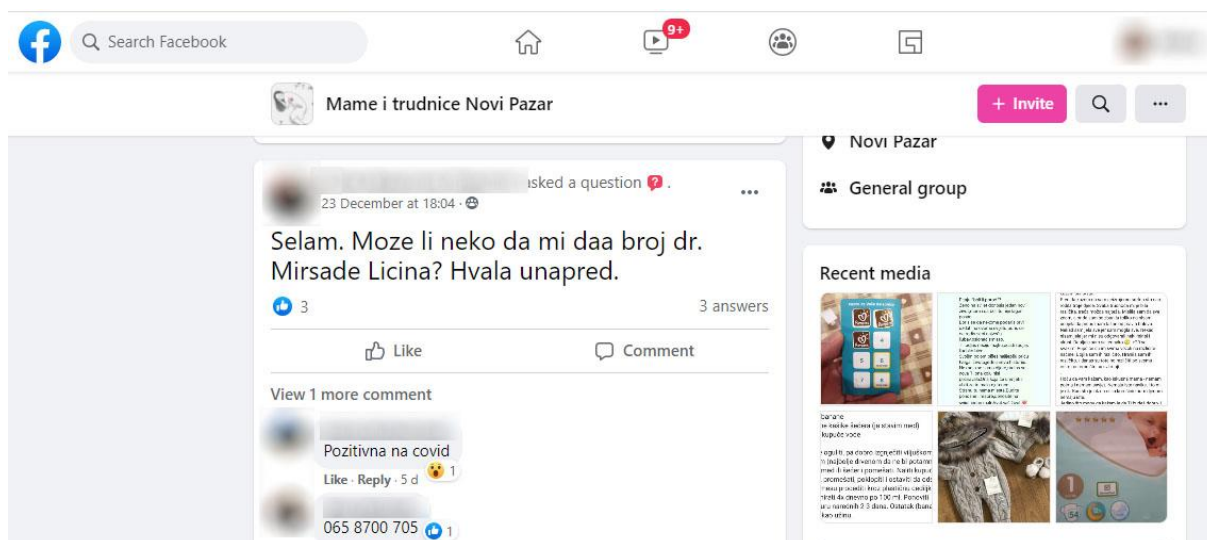
Servis realizuje onlajn kraudsorsing platformu kojom se omogućuje efikasno pružanje stručne psihološke pomoći i samodijagnostifikovanje psiholoških problema kod porodilja i trudnica [144]. Pružanje ovakve specifične podrške je jako značajno za korisnice platforme ali i za stvaranje kolektivnog znanja koje može služiti za detektovanje ovih problema u društvenim mrežama na primer.

### 5.3.4. Uvid u raspored i ocenjivanje lekara

Uvidom u specijalizovane grupe na društvenim mrežama namenjene razmeni informacija o zdravstvu, uočeno je da dosta pacijenata traži preporuku lekara ili informaciju o tome koja je smena lekar, da li je na odmoru ili radi i sl. Primeri ovakvih postova prikazani su na slikama Slika 46 i Slika 47.

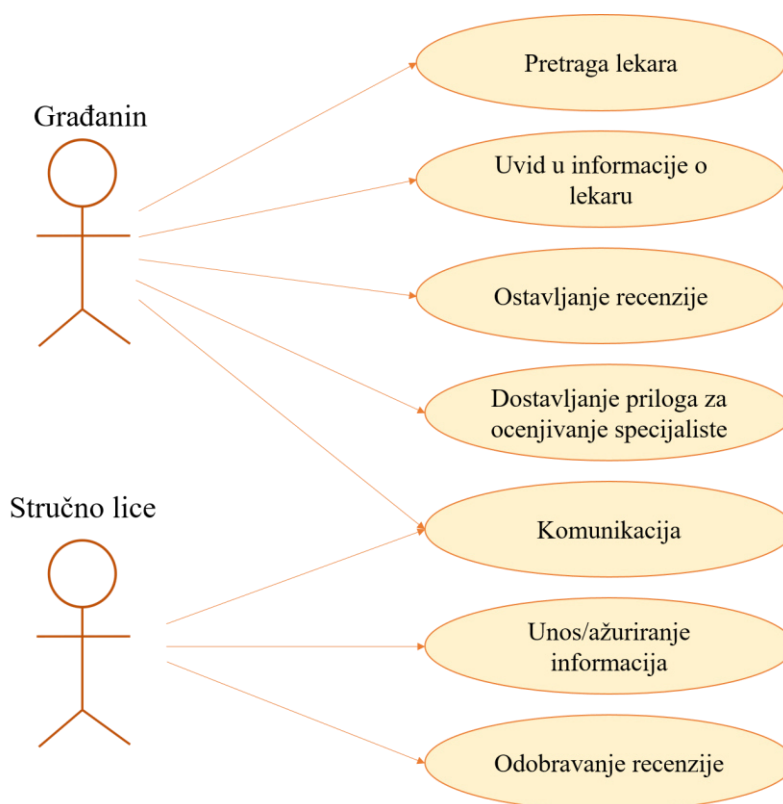


Slika 46. Traženje preporuke lekara



Slika 47. Traženje informacija o lekaru

Servis pametnog zdravstva za uvid u raspored, informacije i ocenjivanje lekara trebalo bi da korisniku omogući pretragu lekara, uvid u informacije o lekaru, broj telefona, koja je smena, uvid u ocene i komentare, mogućnost ostavljanja komentara i ocene. Dijagram slučajeva korišćenja za ovaj servis prikazan je na slici Slika 48.



Slika 48. Slučajevi korišćenja servisa za uvid u raspored i ocenjivanje lekara

Pacijent bi mogao da ostavi recenziju izabranog lekara, a ako se radi o lekaru specijalisti, da bi se ostavila recenzija, mora se ostaviti izveštaj kao dokaz posete, kako bi se sprečila zloupotreba mogućnosti ocenjivanja, a za autorizovane pacijente bila bi omogućena i komunikacija sa lekarom.

Informacije o lekaru sadržale bi njegovo ime, prezime, sliku, kratku biografiju, smenu (raspored) od datuma do datuma, prosečnu ocenu od strane korisnika i komentare, broj telefona, zdravstvenu ustanovu u kojoj je angažovan, specijalizaciju i napomenu.

Stručno lice ima mogućnost unosa informacija o doktoru (ili o sebi) i odobravanje ocene, odnosno komentara (ukoliko je korisnik podneo validna dokumenta kao uslov za ocenu lekara).

## **5.4. Servisi zasnovani na multimediji i veb tehnologijama**

Servisi za prenos multimedije preko Interneta (*IBMMS* - engl. *Internet Based Multimedia services*) [146] doživeli su ekspanziju sa pojavom i širokom primenom pametnih uređaja. Ovim servisima se obezbeđuje slanje zvuka, videa, teksta, animacija i slika preko interneta, u visokom kvalitetu. Ovi servisi mogu se iskoristiti u digitalizaciji zdravstvenih servisa, jer se pomoću njih može omogućiti komunikacija sa medicinskim osobljem u vidu audio ili video telekonsultacija, slanje medicinskih izveštaja u tekstualnom obliku ili skeniranih izveštaja u obliku slike, i ostali vidovi komunikacije u cilju informisanja u okviru realizacije skrining programa i podrške zbrinjavanju u slučaju velikih nesreća. Upravo je prikaz ovih servisa dat u narednim sekcijama ovog poglavlja.

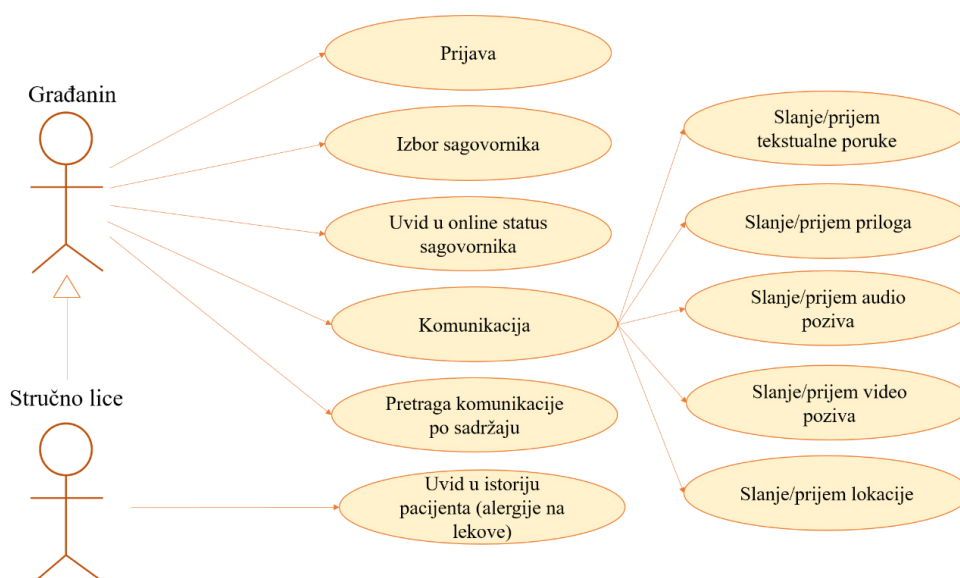
### **5.4.1. Komunikacija sa izabranim lekarima i dežurnom ekipom hitne pomoći**

Kako bi se smanjili redovi čekanja u savremenom zdravstvenom sistemu, neophodno je realizovati pametni servis koji omogućava da se deo komunikacije sa lekarom obavi

elektronskim putem. Ovo je posebno poželjno u slučajevima kada je pacijent zdrav, a potrebno mu je produženje terapije (elektronski recept), slanje rezultata laboratorijskih analiza za čitanje i evidenciju, ali i kada nije u mogućnosti da dođe zbog bolesti (npr. zahtev za otvaranje bolovanja zbog samoizolacije usled kontakta sa obolelim, ili bolesti i sl.). Takođe, na ovaj način poželjno je da se može kontaktirati i hitna služba, pre svega zbog razlike u odnosu na kontaktiranje telefonom. Na ovaj način, hitna služba ne gubi vreme na uzimanje informacija pozivaoca, već ima informacije o lokaciji i istoriji bolesti pacijenta.

Ovaj servis treba da obezbedi prijavu koja sadrži broj zdravstvene knjižice i osiguranja, i da se pri tome otvara mogućnost da se komunicira sa doktorom i sa hitnom službom. U toku izbora sagovornika, građanin ima uvid u status sagovornika, da li je aktivan na mreži ili ne. Nakon izbora sagovornika, nudi se mogućnost komunikacije na razne načine (tekstualnim putem, glasovnim ili video pozivom). Postoji i mogućnost slanja priloga, kao što su rezultati laboratorijskih analiza, izveštaji lekara specijaliste, i sl. Ako je zamena lekara u pitanju, potrebno je omogućiti građaninu komunikaciju sa lekarom koji menja izabranog lekara tog dana. Zbog toga što korisnik ima mogućnosti komuniciranja sa hitnom službom, potrebno je obezbediti i slanje lokacije sa koje se piše. Takođe, potrebno je omogućiti i pretragu tekstualnih poruka po sadržaju.

Na slici Slika 49 prikazan je *UML* dijagram slučajeva korišćenja za opisani servis.



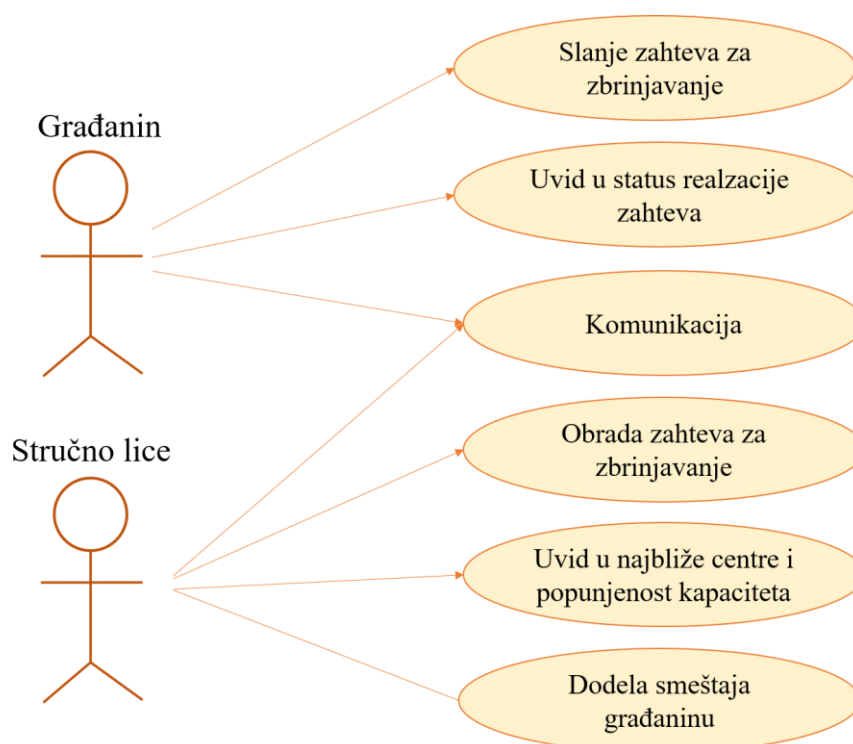
Slika 49. Dijagram slučajeva korišćenja za servis za komunikaciju sa lekarom ili hitnom službom



Na slici se vidi da stručno lice (izabrani lekar ili hitna služba) nasleđuje osobine građanina, što znači da ima sve mogućnosti koje ima i građanin, ali ima i dodatne mogućnosti kao što je uvid u istoriju pacijenta. Ova mogućnost na strani stručnog lica neophodna je usled slanja prve pomoći na lokaciju, jer se na ovaj način vide npr. alergije na lekove pacijenta ili neke druge informacije o pacijentu koje mogu ukazivati na to šta je od opreme potrebno poneti za efikasno pružanje pomoći.

### 5.4.2. Podrška zbrinjavanju u slučaju velikih nesreća

Ovaj servis pametnog zdravstva treba da omogući građaninu koji se našao u nekoj iznenadnoj situaciji velike nesreće (saobraćajne nesreće, požari, eksplozije, hemijski incidenti, poplave, zemljotresi) da se obrati za pomoć ili zbrinjavanje u centrima za ove namene. *UML* dijagram slučajeva korišćenja za ovaj servis prikazan je na slici Slika 50.



Slika 50. Dijagram slučajeva korišćenja za podršku u zbrinjavanju u slučaju velikih nesreća

Akteri u ovom dijagramu su, kao i u većini ostalih pametnih servisa, građanin i stručno lice. Građanin ima mogućnost da podnese zahtev za zbrinjavanje, koji bi bio u vidu

forme koja sadrži neophodne podatke, lokaciju i eventualne napomene. Ovaj servis bi uključivao i mogućnost komunikacije u slučaju da je građaninu neophodna medicinska pomoć ili postoji nešto što ne može izraziti kroz ponuđena polja u zahtevu. Građanin takođe može da vidi status realizacije svog zahteva za zbrinjavanje.

Stručno lice ima mogućnost obrade zahteva, uvid u centre za zbrinjavanje i popunjenost kapaciteta, dodelu smeštaja građaninu, promenu statusa zahteva i komunikaciju sa građaninom. Deo za uvid za centre za zbrinjavanje takođe se može realizovati preko GIS sistema, i odrediti najbliži smeštajni kapacitet u odnosu na lokaciju građanina koji traži zbrinjavanje.

### **5.4.3. Efikasnija realizacija skrining programa**

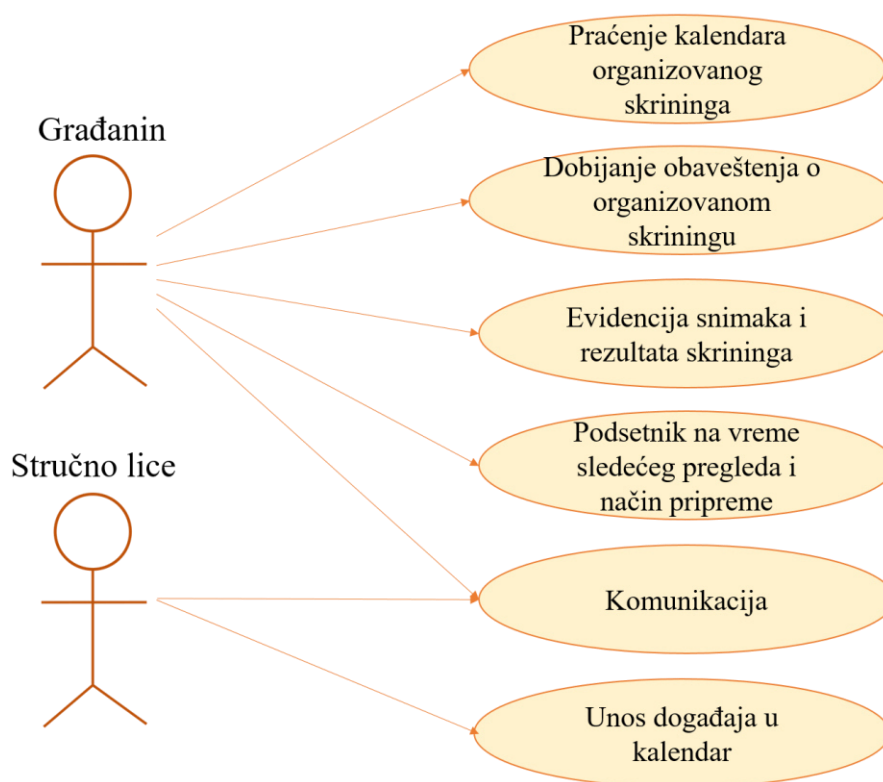
Servis pametnog zdravstva za efikasniju realizaciju skrining programa takođe je, poput servisa opisanog u poglavlju 5.1.1 namenjen prevenciji bolesti. Korišćenjem ovog servisa omogućilo bi se da se u ranoj fazi otkriju neželjene promene i na vreme spreče ozbiljnije posledice po zdravlje građanina. Veliki problem predstavlja neodazivanje ljudi na skrining programe i u nekim delovima zemlje odziv iznosi i manje od 50%. Podaci iz ovog servisa ukazali bi i na to u kojim regijama je potrebno motivisati stanovništvo da se odazove na skrining programe.

Prema [146], skrining predstavlja prepoznavanje do tada neotkrivene bolesti, korišćenjem skrining testa u prividno zdravoj, tj. asimptomatskoj ciljnoj populaciji. Organizovani skrining predstavlja masovno pozivanje ciljne populacije na testiranje i tumačenje skrining testova praćeno strogom kontrolom kvaliteta i izveštavanjem. Organizovani skrining se radi u ciklusima (u zavisnosti od vrste skrininga) u skladu sa epidemiološkim i demografskim podacima kao i kadrovskim i finansijskim kapacitetima.

Organizovani skrining je organizaciono zahtevan i složen proces pa je zato neophodno postepeno širiti teritoriju na kojoj se sprovodi u skladu sa obezbeđivanjem svih potrebnih uslova kao i edukacije zdravstvenih radnika potrebnih za njegovu realizaciju.

Skining test se primenjuje sa ciljem ranog otkrivanja bolesti. Treba da bude visoko senzitivna, specifična i lako primenljiva. Primeri skining testova su na primer citološki bris grlića materice, kolonoskopija, mamografija itd.

U skiningu je izuzetno važno prikupljanje određenih podataka, jer je samo sa kvalitetnim podacima moguće prikazati rezultate skininga u pogledu uspešnosti i efikasnosti, ali i troškova samog skininga. Podaci se prikupljaju i unose u vremenu i na mestu gde nastaju, popunjavanjem protokola za svakog učesnika u skiningu pojedinačno. Zbirni periodični podaci prosleđuju se zavodima za javno zdravlje i Kancelariji za skining u obliku izveštajnog obrasca, bilo u elektronskom ili papirnom obliku. Za uspešnost skininga neophodna je jasno definisana odgovornost pojedinaca u svakoj fazi i kvalitetna saradnja svih aktera - od domova zdravlja, instituta i zavoda za javno zdravlje, preko zdravstvenih ustanova sekundarnog i tercijarnog nivoa, do RFZO-a i Ministarstva zdravlja. Pored ovih aktera, u organizaciji, izvođenju i praćenju skininga učestvuju i mediji i lokalna zajednica. Na primer, Republika Srbija je započela sa postepenim uvođenjem organizovanog skininga raka grlića materice, raka debelog creva i raka dojke od 2012. godine.



Slika 51. Dijagram slučajeva korišćenja servisa za efikasniju realizaciju skining programa

Na slici Slika 51 prikazan je dijagram slučajeva korišćenja za servis za efikasniju realizaciju skrining programa. On treba da građaninu omogući da prati kalendar događaja skrining programa, da može dobiti obaveštenje o organizovanom skrining testu, ali i da vodi evidenciju o svojim snimcima, rezultatima, da ih pošalje izabranom ili lekaru specijalisti. Servis bi trebalo da omogući kreiranje podsetnika koji će javiti građaninu kada i na koji način da se pripremi za sledeći pregled kod lekara specijaliste. S druge strane, stručno lice treba da ima mogućnost komunikacije, pri čemu se ovo odnosi na lekara ili specijalistu kome je poslat radiološki snimak ili rezultati laboratorijskih analiza, da mogu ovaj prilog primiti i komentarisati - dati stručno mišljenje, odnosno poslati objašnjenje pacijentu. Stručnim licima zaduženim za organizaciju skrininga potrebna je mogućnost dodavanja događaja u kalendar, koji bi bio vidljiv svim zainteresovanim građanima.

## 5.5. Servisi zasnovani na blokčejnu

Termin blokčejn vezuje se za pojavu kriptovaluta. Najpopularnija među njima je bitcoin, koji je prvi put opisao naučnik ili grupa njih pod pseudonimom *Satoši Nakamoto* 2008. godine [148]. U isto vreme, kao način zaštite vlasništva nad kriptovalutama, blokčejn je predložen kao način za evidentiranje bitcoin transakcija.

Funkcionisanje blokčejna može se opisati na principu bitcoin transakcije (Slika 52).



Slika 52. Snimanje transakcije u bitcoin blokčejnu

Blokčejn, u ovom slučaju, predstavlja kolekciju digitalnih novčanika. Na primer, ako želimo da izvršimo prenos bitcoina iz novčanika A u novčanik B, svi članovi bitcoin blok lanca moraju to znati. Kreira se novi blok koji sadrži informacije o vremenu transakcije i informacije o njemu. Da bi se dodao kreirani blok u blokčejn, moraju ga verifikovati drugi u lancu. Na taj način transakcija je vidljiva svima.

To znači da su svi podaci o transakcijama u bazi podataka koja se javno distribuira svim vlasnicima, tj. ti podaci su javno dostupni. A podatke šalju korisnici koji imaju pseudonim, pa su informacije unutar bloka anonimne [148][1].



Slika 53. Struktura blokčejna i bloka u lancu

Slika 53 prikazuje strukturu blokčejna i bloka u lancu. Pored ID broja bloka i vremena kreiranja bloka (vremenska oznaka), svaki blok sadrži svoj heš ključ u zaglavlju, kao i heš ključ prethodnog bloka i naravno podatke. Podaci u slučaju kriptovalute su količina novca i transakcije, dok će u našem modelu sadržati informacije o pacijentu.

### 5.5.1. Obezbeđivanje pouzdanosti i trajnosti podataka

Za sigurnost transakcija u kriptovalutama, blokčejn se koristi kao kolekcija digitalnih novčanika. Sve promene se čuvaju na distribuiran i decentralizovan način i nemoguće je vršiti neregularne promene. Za zdravstveni sistem sa digitalnim uslugama koji uključuju različite vrste klijentskih aplikacija, važno bi bilo pročitati podatke radi njihove analize, a ako se promene, to bi bila valjana promena koju su odobrili učesnici u lancu.

Mogućnost korišćenja blokčejn tehnologije za sigurnost transakcija u kriptovalutama navela je mnoge istraživače da razmotre njenu upotrebu u bezbednosti i privatnosti podataka u drugim oblastima.

U radu [149] autori su opisali algoritme za korišćenje blokčejn tehnologije za čuvanje privatnih podataka. U radovima [150][151][152] postoje načini za primenu blokčejn tehnologije u zdravstvenom sistemu, sa naglaskom na elektronske zdravstvene usluge. U radu [153] je opisan pregled mogućnosti korišćenja blokčejn tehnologije za decentralizaciju i zaštitu podataka u uslugama pametnog grada. U radu [150] opisan je prototip sistema *MedRec* za čuvanje zdravstvenih podataka i podataka namenjenih istraživanju zasnovan na blokčejnu. U radu [154] prikazani su slučajevi upotrebe blokčejna u zdravstvu i to kroz sledeće usluge:

1. praćenje lanca snabdevanja lekovima kako bi se suzbila zloupotreba propisivanja opioida i sprečila zavisnost pacijenata od ovih lekova,
2. čuvanje podataka za usluge telemedicine,
3. kontrolisano deljenje osetljivih podataka pacijenta (praćenje kancera, terapije, uputi i sl.),
4. digitalni identitet pacijenta i lični zdravstveni karton,
5. informacije o zahtevima za pokrivanje troškova od strane zdravstvenog osiguranja.

Kolekcija digitalnih novčanika, u ovom slučaju, bila bi zbirka medicinskih kartona pacijenata jednog od pametnih gradova ili država. Ovaj način čuvanja podataka garantuje sigurnost osetljivih podataka koji se čuvaju u *EMR* (elektronski zdravstveni karton). Na slici Slika 54 dat je primer čuvanja *EMR*-a iz *MEDIS.NET* informacionog sistema. U polju za transakciju čuvaju se podaci o pacijentu iz elektronskog izveštaja za odgovarajući pregled.

Ovi izveštaji čine lanac blokova, unutar kog se podaci ne mogu menjati. Vreme za unos podataka u blokčejn nije veće od vremena za unos podataka u bilo koju bazu podatka (izuzev ako se koriste javni servisi za ove svrhe, kod kojih prosečno vreme za potvrđivanje transakcije od 6 - 10 minuta). Jednostavna implementacija je dodatni argument za integraciju u zdravstvenim servisima. U transakciji se čuvaju podaci o pacijentu, ali samo onih podataka na osnovu kojih se ne može zaključiti o identitetu pacijenta. Broj zdravstvenog osiguranja



respiratoru u toku pandemije bolesti izazvane korona virusom) mogli bi se čuvati u blokčejnu, i na taj način izbeći izmene i sumnje u tačnost prikazanih brojki.

Servis za obezbeđivanje trajnosti i pouzdanosti podataka nudi način koji omogućava skladištenje zdravstvenih podataka pomoću blokčejna iz dva razloga. Prva je distribucija podataka i nemogućnost njihove promene od strane treće strane, a druga je čuvanje podataka pod pseudonimom tako da se podaci ne mogu zloupotrebiti, a mogu se koristiti u svrhu istraživanja ili pružanja elektronskih zdravstvenih usluga u pametnom gradu.



## **6. OBRADA PODATAKA U MIS-IMA KAO OSNOVA ZA REALIZACIJU SERVISA E-ZDRAVSTVA**

Velika količina medicinskih podataka i izveštaja se generiše u medicinskim informacionim sistemima i čuva u tekstualnom obliku, kao što su simptomi, istorija bolesti i lekarska zapažanja o zdravlju pacijenata. Elektronsko snimanje podataka o pacijentima ne samo da olakšava svakodnevni rad u bolnicama, omogućava efikasnije upravljanje podacima i smanjuje materijalne troškove, već se takođe može koristiti za dalju obradu i sticanje znanja za poboljšanje javnog zdravlja. Javno dostupni zdravstveni podaci doprineli bi razvoju telemedicine, e-zdravstva, kontrole epidemija i pametne zdravstvene zaštite u pametnim gradovima. Prikaz zdravstvenih podataka u izvornom obliku ili pak njihovo direktno korišćenje i cilju realizacije sofisticiranih servisa nije moguće bez prethodne adekvatne obrade i pripreme što je izuzetno važna i zahtevna aktivnost kojoj treba posvetiti punu pažnju jer direktno utiče na kvalitet dobijenih rezultata. U ovom poglavlju opisana je važnost normalizacije tekstualnih podataka za implementaciju servisa pametnog zdravstva.

Normalizacija teksta je postupak redukcije dokumenta ili njegovih sastavnih delova u oblik pogodan za računarsku obradu. Kada su u pitanju reči, skladište reči treba, umesto celih reči, da sadrži normalizovanu formu ili skup formi za svaku reč. Ovako normalizovane reči daju i mogućnost boljeg upoređivanja sa rečima koje imaju isto ili slično semantičko značenje [155].

Predložen je algoritam za normalizaciju medicinskih podataka na srpskom jeziku kako bi se oni pripremili za dalju obradu, u ovom slučaju u okviru servisa pametnog zdravstva. Primenom ovog algoritma, pored normalizovanih medicinskih izveštaja, dobijaju se i korpusi ključnih reči i stop reči, koji su specifični za medicinski domen i koji se mogu koristiti za poboljšanje rezultata u normalizaciji medicinskih tekstualnih podataka.

U medicinskim informacionim sistemima svakodnevno se stvara i čuva velika količina podataka. Oni omogućavaju čuvanje zajedničkih podataka (broj pacijenata koje pregleda lekar, potrošnja materijala, recepti itd.), ali uključuju i medicinske izveštaje koji sadrže informacije o pacijentu kao što su anamneza, dijagnoza, simptomi itd. koji se mogu koristiti za analizu i predikciju u svrhu poboljšanja medicinskih informacionih sistema. Zbog toga je neophodno pripremiti i na odgovarajući način obraditi ove podatke. Dobro pripremljeni podaci mogu se obrađivati u različite svrhe, pa i u servisima pametnog zdravstva. Oni su važni za poboljšanje kvaliteta i efikasnosti zdravstvenih usluga. Kao takvi, smatraju se osnovom za primenu kod MIS kao neizostavni deo koncepta pametnih gradova. Činjenica je da je problem organizacije života i optimizacije, posebno u velikim gradovima, jedan od trenutnih problema koji se intenzivno rešava. Cilj je pružiti čitav niz usluga koje će život učiniti lakšim i jeftinijim. Zdravstvo zauzima značajno mesto u ovom konceptu.

Kao što je već ranije pomenuto, deo medicinske evidencije može biti obrađen i korišćen u svrhe pametnog zdravstva i njegovih usluga, kao što su kontrola epidemije, vizuelizacija podataka o vakcinaciji, prevencija bolesti, samo-dijagnoza itd.

Motivacija za kreiranje algoritma za normalizaciju medicinskog teksta je stvaranje uslova za kreiranje servisa pametnog zdravstva pomoću tehnika istraživanja teksta. Za potrebe realizacije ovih metoda stvoren je korpus koji sadrži 5.261 medicinski izveštaj. Cilj primene ove metode je brisanje nerelevantnih podataka iz ulaznog skupa medicinskih podataka i pripremu relevantnih podataka za dalju obradu. Relevantni podaci pročišćeni su od suvišnih reči, interpunkcijskih znakova i ostalih podataka koji ne nose bilo kakvu informativnu vrednost. Dobijene stop i ključne reči pamte se kako bi se koristile u normalizaciji novih podataka.

## **6.1. Stanje u oblasti normalizacije medicinskih podataka**

Istraživanja iz ove oblasti mogu se podeliti u dve grupe. U prvu grupu spadaju radovi koji opisuju normalizaciju i obradu medicinskih tekstualnih podataka koji nisu napisani na srpskom jeziku. Ovi radovi opisuju univerzalne karakteristike medicinskih podataka i njihovu

normalizaciju, koje su nezavisne od jezika. U drugu grupu povezanih istraživanja spadaju metode za normalizaciju teksta na srpskom jeziku, koji ne pripada medicinskom domenu.

Razlike između kliničkih i uobičajenih tekstova i problemi u vezi sa dobijanjem informacija iz medicinskog teksta opisani su u radu [156]. U radu [157] su predstavljene karakteristike kliničkih izveštaja iz korpusa medicinskih izveštaja u Švedskoj preuzetih od 2014-2015. Radovi [158][9] su pregledni radovi koji opisuju različite metode koje se koriste u normalizaciji elektronskih medicinskih podataka. U radu [159] se predlaže metoda za normalizaciju simptoma napisana na kineskom jeziku. Kompletan sistem koji normalizuje i izdvaja informacije iz medicinskih kartona i njegova arhitektura opisan je u [160].

S druge strane, metode opisane u prethodnim radovima nisu dovoljne da se u potpunosti primene na medicinski tekst na srpskom jeziku, i naravno, ne uključuju leksičke resurse specifične za medicinski domen. U drugu grupu spadaju radovi koji se bave normalizacijom tekstualnih dokumenata i njihovom analizom na srpskom jeziku, ali ni ovde normalizacija nije sprovedena nad kliničkim tekstovima. U radovima [155], [77] dat je opis procesa normalizacije neformalnih dokumenata na srpskom jeziku s ciljem bržeg pretraživanja. U [143] je predstavljena normalizacija teksta na srpskom jeziku korišćenjem  $n$ -gram analize.

Određeni jezički resursi potrebni su za obradu medicinskih podataka, a među njima i resursi ličnih imena i prezimena na srpskom jeziku, koji su prikazani u radovima [162][163].

Neki koraci iz ovih metoda mogu se koristiti u normalizaciji medicinskih podataka, ali većinu njih treba prilagoditi. Struktura medicinskih podataka (izveštaja) i njihov sadržaj znatno se razlikuju od ostalih vrsta tekstualnih dokumenata, pa ih je potrebno prilagoditi njihovim specifičnostima.

## 6.2. Opis korišćenog skupa podataka

Karakteristike jezika na kojem je napisan izveštaj dodatno komplikuju proces normalizacije. Potrebni su jezički korpusi za identifikaciju specifičnosti u okviru medicinskih izveštaja. Medicinski izveštaji su osetljivi za istraživanje zbog poverljivosti informacija koje

nose, pa se mora izvršiti odgovarajuća deidentifikacija medicinske dokumentacije i ukloniti svi lični podaci pacijenata, kao i lekara.

Dostupno je nekoliko korpusa na engleskom jeziku, kao što su korpusi navedeni u [164]. Ne postoji korpus na srpskom jeziku u elektronskom obliku koji je javno dostupan.

Za potrebe ovog istraživanja korišćen je korpus koji čini 5261 medicinski izveštaj. Ovi izveštaji napisani su na srpskom jeziku iz 32 zdravstvene ustanove koji pripadaju domu zdravlja Niš (DZ Niš), a prikupio ih je informacioni sistem MEDIS.NET [132]. Medicinske izveštaje napisalo je 169 različitih lekara. Ovaj korpus je napravljen prema svim etičkim standardima, uz uklanjanje identiteta pacijenata i medicinskog osoblja, kao i održavanje veza sa pripadnošću više izveštaja istom pacijentu.

### **6.3. Karakteristike medicinskih podataka**

Medicinski izveštaji se uglavnom generišu za interne potrebe zdravstvenih ustanova. Klinički izveštaji su potrebni raznim korisnicima kao što su: medicinsko osoblje kako bi pratilo svakodnevne aktivnosti, pacijenti za dokumentovanje svog zdravstvenog statusa, klinička istraživanja (medicinski istraživači, farmaceuti, epidemiolozi, itd.), rukovodstvo bolnice, da bi se pratile finansije i planiranje zaliha itd. Medicinski izveštaji mogu sadržati numeričke i tekstualne informacije. Medicinski podaci su mešovitog tipa (strukturirani, polustrukturirani i nestrukturirani) i stoga zahtevaju složeniju obradu koja uključuje postojanje odgovarajućih specijalizovanih leksičkih izvora. Strukturirani deo sadrži vrednosti za neke podatke ili parametre o pacijentu koji se unose u okviru za to predviđenih polja (npr. ime, prezime, godina), pa ih je najlakše obraditi. Polustrukturirani deo daje opisne vrednosti za neke parametre, ali struktura je i dalje poznata (temperatura, pritisak i laboratorijska analiza). Nestrukturirani deo sastoji se od slobodnog teksta koji daje lekar i sastoji se od simptoma, istorije, zapažanja, zaključaka. Nestrukturirani podaci sadrže jezički nepotpune, neformalne i nestandardne skraćenicе što otežava računarsku obradu i analizu. Za ovo je neophodno prethodno obraditi podatke pre analize da bi se oni doveli u standardizovani oblik.

Tabela 4 daje primer medicinskog izveštaja iz skupa koji se obrađuje. U njemu možemo identifikovati strukturirani deo koji sadrži datum usluge, naziv usluge, dijagnozu, kod dijagnoze, organizacionu jedinicu u kojoj je usluga pružena i lokaciju usluge. Takođe, ovaj izveštaj sadrži nestrukturirani deo koji se sastoji od slobodnog teksta u koji doktor upisuje svoja zapažanja, koji je u podacima označen kao anamneza, pa je u daljem tekstu tako i nazivan. Ovaj deo je složeniji za obradu, jer treba izdvojiti relevantne podatke i transformisati ih u standardizovani format, pogodan za dalju obradu.

Tabela 4. Primer medicinskog izveštaja iz korišćenog skupa podataka

<b>Datum pružanja usluge</b>	23-03-18
<b>Naziv usluge</b>	Ponovni pregled odraslih
<b>Anamneza</b>	Pacijent dobio sinoc osip po koži. Makulopapulозна ospa po kozi iza ušiju, čela i spušta se na trup. Vezikularni disajni šum
<b>Dijagnoza</b>	Morbili
<b>Kod dijagnoze</b>	B05
<b>Organizaciona jedinica pružanja usluge</b>	Opšta medicina
<b>Lokacija pružene usluge</b>	Centralna zgrada

Medicinske izveštaje najčešće pišu lekari. Zbog brzine kojom su napisani, često sadrže mnogo grešaka. Vrlo često su rečenice nepotpune, na primer, izostavljaju se pomoćni glagoli kao i subjekat kada je očigledno da je subjekat sam pacijent. Čak se i prilozi retko mogu naći u medicinskim podacima, izuzev kod opisa simptoma (npr. temperatura, znojenje, otežano disanje). Primer je dat u tabeli Tabela 5.

Tabela 5. Primer nepotpunih rečenica u anamnezi

<b>Originalni tekst anamneze</b>	<b>Prošireno značenje</b>
“pečenje i svrab po celom telu, difuzna ospa koja svrbi”	Pacijent ima pečenje i svrab po celom telu u obliku difuzne ospa koja svrbi

Takvi opisi su koncizni i nose najvažnije informacije koje na neki način olakšavaju obradu takvog teksta. U medicinskim izveštajima svaka reč ima veću težinu. Pored namernog izostavljanja određenih reči, postoje i vrlo česte greške i pravopisne greške kao što su:

- dve reči su spojene, a umesto razmaka je slovo,
- pogrešno napisana reč,
- netačno pisanje dijakritičkih simbola (npr. "izvestaj" umesto "izveštaj"),
- grupa slova "dj" umesto slova "đ",
- često pogrešno napisano ili izostavljeno slovo, ili slovo viška, npr. „kašljke“ umesto „kašlje“,
- postoje skraćenice medicinskih termina, interpunkcijski znakovi, zagrade, brojevi, itd.
- slovo "x" u srpskoj reči u anamnezi (npr. *extremiteti*),
- slova "z" i "y" zamenjeni zbog tastature.

U poređenju sa drugim vrstama teksta prema [165], u medicinskim izveštajima se nalazi dvostruko više pravopisnih grešaka (10%) nego u naučnim radovima, novinskim člancima, veb člancima itd. To nije iznenađujuće s obzirom na vremensko ograničenje za pregled pacijenta.

Skraćenice često nalazimo u medicinskim izveštajima. Skraćenice su načini pisanja dužih reči bez svih slova. Olakšavaju pisanje i čitanje teksta, ali samo pod uslovom da čitalac zna značenje skraćenica. Problem su nestandardizovane skraćenice. Ponekad ista skraćenica može imati dvosmisleno značenje (na primer, „feb“ za „februar“ i „febrilan“). Čest je slučaj da se za isti pojam koriste različite skraćenice: na primer za temperaturu (npr. „T\*“, „t“, „temp“, itd.). Pored skraćenica, često se koriste i akronimi. U medicinskim izveštajima koriste se standardizovani akronimi. Međutim, i ovde se mogu javiti problemi jer akronimi mogu imati više značenja, na primer: „DIK“ je akronim za „dečju infektivnu kliniku“ i „diseminovanu intravaskularnu koagulaciju“, „EEG“ za „elektroencefalografiju“ i „elektroencefalogram“, itd. Tumačenje akronima, u ovom slučaju, verovatno će biti povezano sa specijalnošću lekara koji ih piše ili sa dijagnozom. Ovo povezivanje može se obraditi ručno, ali i metodama mašinskog učenja sa odgovarajućim korpusom podataka.

Reči koje imaju latinski ili grčki koren su uobičajene u medicini. Međutim, poslednjih decenija koristi se sve više engleskih reči za koje ne postoji odgovarajući prevod na srpski jezik, pa se najčešće koriste u izvornom obliku. Ovo je posebno upečatljivo kada su u pitanju

medicinske tehnike, medicinski uređaji i određeni hirurški postupci. U zavisnosti od toga koja je svrha dalje obrade medicinskih tekstova koji se normalizuju, može biti poželjno da se pojednostave pozajmljenice (tuđice) korišćene u medicinskim izveštajima (npr. „pulmo“ - pluća, „hiperemija“ - povećanje protoka krvi u različitim tkivima u telu).

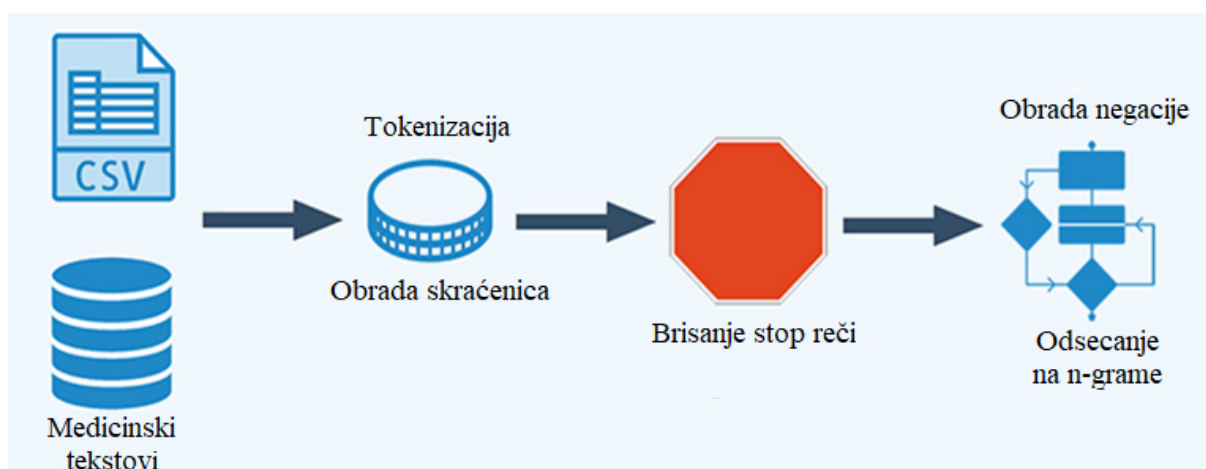
Pojava negacije u medicinskim tekstovima je vrlo česta jer isključuje postojanje nekih simptoma koji ukazuju na bolest. Prisustvo negacije u izveštaju značajno utiče na značenje samog izveštaja, s obzirom da su medicinski tekstovi kratki i sažeti. Primer anamneze sa negacijom dat je u tabeli Tabela 6. Značaj obrade negacije u medicinskim tekstovima pokazan je u engleskom korpusu *BioScope* sa medicinskim izveštajima u kojima je negacija ručno obeležena, kao i obim njenog delovanja [164]. Za srpski jezik postoje pravila negacije koja mogu značajno poboljšati obradu podataka [166].

Tabela 6. Primer anamneze sa negacijom

Anamneza s negacijom
“Izbilo ga nešto po telu, pre dva dana, temper. nema, ne kašlje, ne boli ga grlo”

## 6.4. Metoda za normalizaciju medicinskog teksta

Metoda za normalizaciju medicinskih izveštaja koja je predložena sastoji se od nekoliko koraka: tokenizacija, uklanjanje stop reči, obrada negacije, uklanjanje interpunkcije i odsecanje u  $n$ -grame. Algoritam koji je primenjen na opisane podatke prikazan je na slici Slika 55.



Slika 55. Normalizacija medicinskih izveštaja

### 6.4.1.Tokenizacija

U prvom koraku pripreme medicinskog teksta za dalju obradu, potrebno je iz anamneze izbrisati nepotrebne elemente (više praznih mesta, datume, posebne znakove itd.) da bi se identifikovale označene reči. Anamneze su napisane na srpskom jeziku latiničnim pismom koja sadrži dijakritičke simbole „ć“, „č“, „ž“, „š“, „đ“, pa se u ovom koraku menjaju simbolima „cx“, „cw“, „zx“, „sx“ i „dx“, s obzirom da se slova „x“ i „w“ ne koriste u srpskom pismu, a predložene kombinacije se ne nalaze u korpusu bolesti i srodnih zdravstvenih problema (srpska i latinska verzija) [168]. Ovo olakšava obradu reči koje sadrže dijakritičke znake.

U ovom koraku se vrši i obrada skraćenica. Skraćenice koje pripadaju standardnim medicinskim skraćenicama uglavnom su akronimi i pišu se velikim slovom. Nestandardizovane skraćenice specifične za srpski jezik izdvajaju se i svode u jedan oblik bez tačke na kraju. Interpunkcijski znakovi ostaju u ovom koraku zbog njihovog značenja u obradi negacije. Kada se izveštaji pripreme na ovaj način, mogu se podeliti na reči, odnosno tokene, na nivou na kome se odvija dalja obrada.

### 6.4.2.Brisanje stop reči

Izveštaji sadrže i reči koje ne nose značaj u rečenici i te reči se nazivaju stop rečima. Ove reči treba izbrisati iz medicinske evidencije kako bi se smanjio njen obim i zadržale samo relevantne informacije. Smanjivanjem obima podataka povećava se brzina njihove dalje obrade. Stop reči su obično prilozi, predlozi, zamenice, veznici, reči i druge reči koje nisu relevantne za određivanje značenja teksta. Proces uklanjanja stop reči započinje stvaranjem rečnika stop reči. Rečnik stop reči na srpskom jeziku, koji je rezultat prethodnog istraživanja [143], sadrži 3117 stop reči. U tabeli Tabela 7 date su neke reči iz ovog rečnika. Neke reči treba ukloniti iz skupa stop reči jer predstavljaju medicinske skraćenice ili skraćenice za hemijske elemente. Njihova obrada će se izvršiti naknadno, s obzirom na to da li odgovarajući elementi predstavljaju skraćenicu ili stop reč (bez obzira na to da li su napisana velikim slovom, da li se pojavljuju sa drugim ključnim rečima, na primer, „se“ može biti pomoćni



glagol, a „Se“ je vrednost selena iz testova krvi). Signali negacije takođe su izuzeti iz skupa stop reči kako se ne bi izgubili podaci o negaciji postojanja nekih simptoma.

Tabela 7. Neke stop reči iz rečnika

<b>Primeri stop reči</b>	
kakva	nećete
kakvih	nećemo
kakvima	nećeš
igde	neću
iako	će
htelo	ćete
htela	ćemo
hteo	ćeš

### 6.4.3.Obrada negacije

Specifičnost medicinskih izveštaja je u tome što su napisani u trećem licu i sažeti, tako da se ne pojavljuju svi signali negacije. Od signala negacije u obrađenom korpusu pronađene su sledeće reči: „ne“, „nije“, „nema“, „bez“ i „ni“. Za opseg delovanja negacije uzeće se deo anamneze od prve reči koja sledi signal negacije do interpunkcijskog znaka, ili ako ona ne postoji, uzeće se reč pre signala negacije do interpunkcijskog znaka. Reč u opsegu negacije dobija prefiks „ne“ i signal negacije se briše. Primer je dat u tabeli Tabela 8. Interpunkcijski znakovi se mogu ukloniti nakon obrade negacije.

Tabela 8. Anamneza nakon obrade skraćenica i negacije

<b>Pre obrade</b>	<b>Posle obrade</b>
Izbilo ga nešto po telu, pre dva dana, temper. nema , ne kašlje, ne boli ga grlo	izbilo telu, dana, ne_temperaturu , ne_kašlje, ne_boli grlo

#### 6.4.4. Odsecanje na $n$ -grame

Imajući u vidu bogatstvo srpskog jezika, prisustvo padeža, glasovnih promena i promena glagola prema licu, rodu i broju, dešava se da reči koje imaju isto značenje mogu da se nađu u mnogim oblicima. Primena gramatičkih pravila za svođenje reči na osnovu učinila bi problem normalizacije složenim. Takođe, bio bi potreban i morfološki rečnik srpskog jezika, koji nije javno dostupan u elektronskom obliku. Iz ovih razloga, za svođenje reči na osnovu izabrana je jezički nezavisna varijanta, koja predstavlja odsecanje na prvi  $n$ -gram (prefiks veličine  $n$ ).  $N$ -gram je podskup koji se sastoji od  $n$  elemenata datog niza. Na primer, reč „učiti“ sastoji se od sledećih  $n$ -grama: u-č-i-t-i (dužina 1), uč-či-it-ti (dužina 2), uči-čit-iti (dužina 3), učit-čiti (dužina 4) i učiti (dužina 5).  $N$ -grami se dobijaju pomeranjem okvira dužine  $n$ , čija vrednost može biti u položajima 1 do  $m - n + 1$ , gde je  $m$  dužina niza [143]. Analizom sadržaja  $n$ -grama može se uočiti korelacija između izgleda  $n$ -grama i karakteristika teksta.  $N$ -grami su pogodni za upotrebu u analizi tekstualnih dokumenata na prirodnim jezicima, zbog jezičke nezavisnosti.  $N$ -gram analiza je postupak koji se primenjuje na tekst i čiji je rezultat dobijanje skupa  $n$ -grama zadate dužine. U predloženoj metodi korišćeno je odsecanje na prvi  $n$ -gram čija je dužina 4 (4-gram). Pri odsecanju reči na 4-grame, negativni prefiks se zadržava (nema temperaturu se odseca na ne\_temp). Ova dužina je izabrana kao optimalna jer daje najbolje rezultate za srpski jezik (na datom skupu) u poređenju sa analizom 3-grama i 5-grama (poređenje je dato u sledećoj sekciji).

#### 6.4.5. Klasifikacija $n$ -grama

Klasifikacija  $n$ -grama je potrebna nakon normalizacije, gde se  $n$ -grami razvrstavaju u ključne reči i stop reči koje su specifične za medicinski domen. Ključne reči su u ovom slučaju izdvajane za potrebe servisa za prikaz podataka o najčešćim simptomima bolesti u servisu za kontrolu epidemija. Stoga one uključuju simptome koji se mogu zapamtiti za odgovarajuću vrstu bolesti. U zavisnosti od potreba za koje se izdvajaju ključne reči, to mogu biti dijagnoze, kodovi dijagnoza, vrednosti laboratorijskih analiza, itd. Medicinske stop reči su one reči koje se javljaju u većini anamneza, bez obzira na dijagnozu. To su reči koje se najčešće pojavljuju u medicinskim izveštajima („uput“, „doznaka“).

S obzirom da u ključnim rečima postoje sinonimi, oni moraju biti grupisani. Klasifikacija je proces koji zahteva posebnu pažnju i koji će biti predmet narednog poglavlja. Ovde je klasifikacija napravljena na jednostavan način da bi se pokazao uticaj normalizacije na rezultate. Klasifikacija je napravljena traženjem  $n$ -grama među simptomima i ako su pronađeni, oni su deklarirani kao ključne reči za odgovarajuću vrstu bolesti. Oni  $n$ -grami koji se najčešće pojavljuju a ne nalaze se u simptomima proglašeni su stop rečima, odnosno ne nose značaj prilikom analize podataka za realizaciju ove vrste servisa.

## 6.5. Rezultati primene metode za normalizaciju medicinskih dokumenata

Nakon normalizacije opisanih medicinskih izveštaja, anamneze su dobijene u normalizovanom obliku. Od 5261 medicinskog izveštaja, 2112 je sadržao tekst u nestrukturiranom odeljku, dok su ostali sadržali samo strukturirani deo.

Tabela 9. Rezultati različitih metoda normalizacije

Metoda normalizacije	<i>BR</i>	<i>NR</i>	<i>KNR</i>	<i>Preciznost</i>	<i>Odziv</i>	<i>F1-Rezultat</i>
<b>3-gram</b>	16606	13011	10213	0.785	0.615	0.6897
<b>4-gram</b>		10330	9488	0.9185	0.5714	0.7045
<b>5-gram</b>		8472	7891	0.9314	0.4752	0.6293
<b>Predložena metoda</b>		13375	12232	0.9145	0.7366	0.816

Nakon izbacivanja brojeva i stop reči, ostalo je 16606 reči za normalizaciju (*BR* u tabeli). Izvršena je normalizacija sa  $n$ -gramima dužina 3, 4 i 5, da bi se utvrdilo koja dužina  $n$ -grama daje najbolje rezultate. Za upoređivanje rezultata korišćene su mere *Preciznost*, *Odziv* i *F1-rezultat*, prikazane formulama 5, 6 i 7.

$$Preciznost = \frac{KNR}{NR} \quad (5)$$

$$Odziv = \frac{KNR}{RN} \quad (6)$$

$$F1 - rezultat = \frac{2 * Preciznost * Odziv}{Preciznost + Odziv} \quad (7)$$

Ovde je *KNR* broj korektno normalizovanih reči, *NR* je broj normalizovanih reči, a *RN* je broj reči koje treba normalizovati.

Rezultati normalizacije prikazani su u tabeli Tabela 9, a najbolji *F1-rezultat* dobijen je za 4-grame. Iz tabele se vidi da povećanje dužine *n*-grama povećava *Preciznost* i *Odziv* slabi. Ali za 3-grame *Preciznost* je mnogo niža, dok je za 5-grame *Odziv* manji, pa je samim tim i njihov *F1-rezultat* manji u poređenju sa 4-gramima. Očekivalo se da će *Preciznost* za 3-grame biti niža, jer ako je reč imala prefiks, koren (kao nosilac svog značenja) se uklanja normalizacijom. Ovde je korektno normalizovana reč čiji je *n*-gram deo semantičkog korena i ovo je upoređeno sa ručno označenim korpusom.

Zatim je primenjena predložena metoda iz prethodne sekcije, a dobijeni rezultati takođe su dati u tabeli Tabela 9. Može se videti da predložena metoda ima najbolji *F1-rezultat* (0,8160). Ovi rezultati su uporedivi sa metodama za normalizaciju medicinskih podataka na drugim jezicima (u [159] *F1-rezultat* = 0,6562). Reči su svedene na prvi 4-gram, a zatim prebrojane kako bi se pronašla njihova učestalost u anamnezi. Sada se anamneze pamte u prečišćenom obliku, sadrže samo relevantne podatke i njihov obim je smanjen. Tabela 10 daje primer anamneze pre i posle normalizacije. Zatim su *n*-grami sa najvećim brojem slučajeva u istorijama bolesti podeljeni u dve grupe, po značenju. Prvu grupu čine ključne reči, a to su one reči čije je značenje usko povezano sa bolešću i prikazane su u tabeli Tabela 11.

Tabela 10. Primer normalizovanih anamneza

Pre normalizacije	Posle normalizacije
Kontrola: još uvek ima malaksalost, slabost, Savet , kontrola 30. 3. 2018	kont mala slab save kont
Pacijent dobio sinoc osip po koži. Makulopapulozna ospa po koži iza ušiju, čela i spušta se na trup. Vezikularni disajni šum	paci dobi osip kozxi maku ospa kozxi usxij cwela spusx trup vezi disa sxum

Tabela 11 pokazuje da broj pojava simptoma u korpusu značajno varira u zavisnosti od primene metode normalizacije. Dakle, ako se  $n$ -grami pretražuju u korpusu, a skraćénice, sinonimi i negacije ne uzimaju u obzir, neće se dobiti tačan broj simptoma. Tabela 11 pokazuje da je procenat pojave nekih simptoma značajno povećan („osip“ zbog sinonima i „temperatura“, kako se često skraćuje). Postoje i oni termini čiji se broj pojava smanjuje (na primer, „promena“) nakon primene predložene metode, jer su pronađeni u negaciji.

Druga grupa izdvojenih  $n$ -grama sastoji se od stop reči čija značenja ne određuju bolest pacijenta i mogu se pojaviti u bilo kojoj medicinskoj istoriji ili kliničkom dokumentu. Kako je anamneza već u prečišćenom obliku bez stop reči koje su karakteristične za opšte dokumente, tako izdvojene stop reči odnose se na medicinski domen i mogu se čuvati radi poboljšanja rezultata normalizacije nove anamneze ili dokumenata. Oni su prikazani u tabeli Tabela 12. Zbog toga se ova normalizacija može koristiti u izvršavanju statistike u kontroli epidemija. Jednostavan primer vizuelizacije najčešćih simptoma epidemijske bolesti dat je na slici Slika 55.

Tabela 11. Procenat pojavljivanja za najčešćih ključnih reči

<b><math>n</math>-gram</b>	<b>Klasična 4-gram metoda</b>	<b>Predložena metoda</b>	<b>Pridružena reč</b>	<b>Pridruženi <math>n</math>-grami sinonima</b>
hipe	10,62%	12,07%	hiperemija	
kasxl	5,97%	6,63%	kašalj	kaša
kozxi	5,40%	6,01%		kože, kožn, koži
licu	7,34%	10,89%		lica, lice
morb	17,90%	17,85%		
ospa	13,92%	25,76%		osip
prom	5,16%	5,02%		
pulm	14,16%	13,92%		
telu	10,89%	12,74%		tela, telu, telo
temp	8,43%	18,32%		
zxdre	12,36%	12,41%		

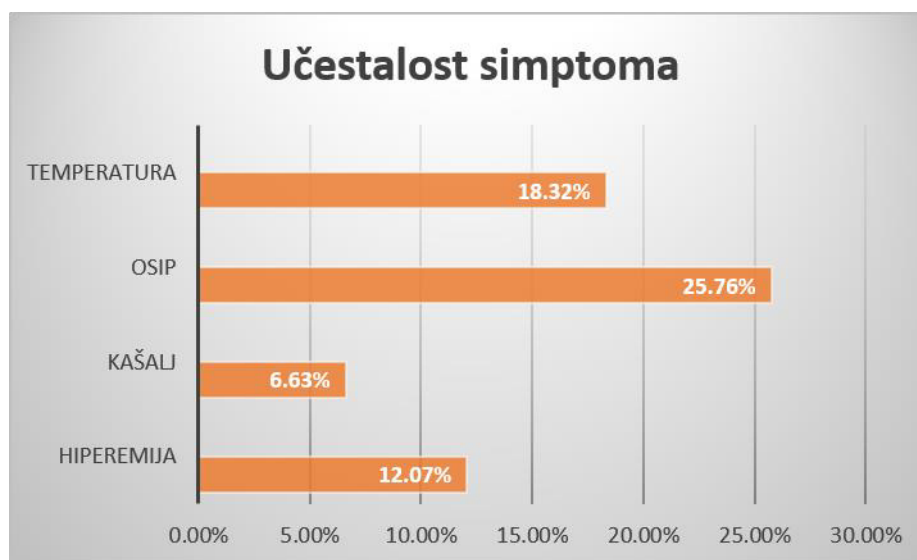
Ako uporedimo tabelu Tabela 11 sa grafikonom (Slika 56), može se zaključiti da se može izvući više informacija kako bi simptomi bili tačniji. Na primer, u tabeli su pomenute „koža“, „lice“, „telo“, koje sadrže odredbe o osipu. To znači da bi se zajednička pojava  $n$ -

grama u anamnezi mogla utvrditi kako bi se dobile fraze, koje će biti predmet daljeg istraživanja.

Ključne reči koje se najčešće javljaju povezane su sa simptomima bolesti („osip“, „temperatura“, „kašalj“ itd.), a najčešće stop reči su medicinski izrazi koji ne ukazuju na stanje pacijenta („kontrola“, „izveštaj“ itd.).

Tabela 12. Medicinske stop reči koje se najčešće pojavljuju

<i>n</i> -gram	Procenat pojavljivanja u anamnezama	Pridružena reč
kont	21,77%	kontrola
izve	13,01%	izveštaj
bolo	11,36%	bolovanje
dana	10,46%	dan
dozn	9,28%	doznaka
uput	7,00%	uput
nala	5,73%	nalaz
preg	5,30%	pregled



Slika 56. Statistika u kontroli epidemija

Ovo poglavlje opisuje kako se medicinski izveštaji u elektronskom obliku koji se svakodnevno čuvaju mogu iskoristiti u svrhu realizacije servisa pametnog zdravstva, kao i problemi sa kojima se mogu susresti u njihovoj analizi i obradi.

Pokazan je značaj normalizacije teksta za servise pametnog zdravstva i predstavljen način normalizacije kliničkih tekstualnih informacija napisanih na srpskom jeziku.

U eksperimentalnom delu su predstavljeni rezultati primene predloženog algoritma za normalizaciju u anamnezama prikupljenim korišćenjem medicinskog informacionog sistema MEDIS.NET. Opisana metoda izdvaja relevantne podatke iz medicinskih izveštaja kako bi se mogli koristiti u razne svrhe, uključujući servise javnog zdravstva. Podaci sačuvani u prečišćenom obliku mogu se koristiti za dalju obradu. Primenom predložene metode nad prikupljenim podacima dobijen je *F1-rezultat* (0,816). S obzirom da na srpskom jeziku ne postoji odgovarajuća metoda za poređenje, u poređenju sa metodama za normalizaciju medicinskih podataka na drugim jezicima [159], predložena metoda daje dobre rezultate. Na ovaj način može se formirati korpus koji sadrži stop reči za medicinski domen na srpskom jeziku koji se mogu koristiti za obradu medicinskog teksta u razne svrhe. Ova normalizacija predstavlja dobru osnovu za klasifikaciju medicinskih pojmova i obeležavanje entiteta u medicinskim kartonima i njihovu primenu u servisima pametnog zdravstva.

Servisi iz predložene platforme pametnog zdravstva koji koriste normalizaciju medicinskih podataka, u ovom ili modifikovanom obliku su servis za kontrolu epidemija, servis za automatsko odgovaranje na pitanja i pretragu medicinskih dokumenata i servis za obradu psiholoških testova.

## **7. IMPLEMENTACIJA SERVISA E-ZDRAVSTVA KORIŠĆENJEM ALGORITAMA VEŠTAČKE INTELIGENCIJE**

Podaci uključeni u *EMR* su strukturirani (ime, prezime, starost, dijagnoza, itd.) i nestrukturirani (opis stanja pacijenta koji se ne može izraziti kroz ponuđena strukturirana polja, poput simptoma, anamneze, istorije bolesti itd.). Pored podataka, sve informacije se koriste u razne svrhe u cilju naučnog istraživanja, pregleda stanja ili poboljšanja procesa u zdravstvenim ustanovama. Ove analize često uključuju različite metode i alate kao što su statističke metode, poslovna inteligencija, metode veštačke inteligencije i druge. Poznato je da je lakše tražiti i izvlačiti znanje iz strukturiranih podataka u poređenju sa nestrukturiranim. Na primer, ako postoji kod za svaki simptom koji korisnik može imati, biće lakše tražiti simptome po kodu. Međutim, ako su napisani na prirodnom jeziku, tada su mogući sinonimi i skraćenice, uključujući različite prezentacije istog simptoma. Nestrukturirani deo takođe sadrži značajne informacije, stoga je izuzetno važno omogućiti njegovu analizu. Po pravilu su potrebni znatni naponi da bi se do njih došlo, a obično se mora raditi ručno označavanje reči.

Najjednostavniji i najzahtevniji način za rešavanje ovog problema u smislu utrošenog vremena (iako je praktično beskoristan u slučaju masovnih podataka) je imati stručnjake koji čitaju i dodeljuju značenje bilo kojem medicinskom dokumentu [169]. Nemoguće je koristiti ovaj pristup u realnom vremenu, s obzirom na količinu dokumenata koji se prenose u bazu podataka *EMR*-ova.

Da bi se rešio ovaj problem, neophodno je koristiti tehnike istraživanja medicinskog teksta usmerene na izdvajanje znanja iz teksta. Za izdvajanje znanja iz nestrukturiranog dela podataka potrebne su metode *NLP* (obrada prirodnog jezika). *NLP* softver dizajniran je za pretvaranje slobodnog teksta u mašinski čitljive strukturirane podatke i sadrži metode koje utvrđuju sentiment teksta i latentne karakteristike prirodnog jezika [10]. U ovom poglavlju



prikazane su razvijene metode za označavanje pojmova u medicinskim izveštajima koje daju bolje rezultate u odnosu na rezultate dobijene primenom postojećih alata i metoda. Korišćenjem predloženih metoda, sve reči sadržane u slobodnom delu medicinskih izveštaja kategorisane su kao termini koji pripadaju medicinskom domenu (simptomi, opisi simptoma, dijagnoze, biohemijske analize, latinske reči, anatomska imena organa, terapije i drugi medicinski izrazi) i one izvan medicinskog domena (brojevi, negativni simboli, stop reči i druge reči bez značaja za zdravstveno stanje pacijenta).

Označavanje simptoma, dijagnoza i drugih medicinskih termina u nestrukturiranom delu *EMR*-a može doprineti izdvajanju nespecifičnih simptoma, što može ubrzati dijagnostiku u vreme epidemija.

Ne postoje odgovarajući javno dostupni leksički izvori ili medicinski korpusi za srpski jezik koji se mogu koristiti za označavanje medicinskih termina. Stoga je za ovo istraživanje pripremljen skup rečnika sa medicinskim (dijagnoze, lekovi, latinski izrazi itd.) i nemedicinskim terminima (stop reči i neke vlastite imenice).

Primenjeno je ručno označavanje medicinskih izraza u nestrukturiranom delu iz 4112 medicinskih izveštaja. Predložene su tri metode: metoda zasnovana na dobijenim rečnicima medicinskih termina i dve metode koje koriste ručno obeleženi skup medicinskih izveštaj, od kojih je jedna hibridna i primenjuje pravila za obradu skraćenica i grešaka. Predstavljeni su rezultati za sve metode, zajedno sa njihovim varijacijama kada se primenjuju stemer i odsecanje na prefiks dužine  $n$  da bi se reč svela na osnovni oblik.

Zbog specifičnosti jezika koji se koristi u pisanju medicinskih izveštaja (upotreba skraćenica, prisustvo grešaka, latinski izrazi itd.), pokazano je da metoda zasnovana na rečniku ne daje najbolje rezultate. Rezultati su bolji ako se automatsko obeležavanje vrši na osnovu ručno označenih korpusa i pravila za obradu grešaka i skraćenica. Pokazalo se da hibridna metoda zasnovana na mašinskom učenju i pravilima daje najbolje rezultate od svih predloženih metoda bez svođenja reči na osnovni oblik.

Predloženi hibridni metod koristi pravila za ispravljanje nedostataka u anamnezi *EMR*-ova (grešaka i skraćenica) i omogućava označavanje medicinskih i nemedicinskih pojmova u *EMR*-u, tako da se mogu lakše analizirati. Ovo je jedan od glavnih doprinosa disertacije. Zbog

važnosti tačnosti u donošenju medicinskih odluka, metode čistog mašinskog učenja nisu primenljive u praksi, ali se veći značaj pridaje metodama zasnovanim na pravilima. Ponekad se netačni izrazi mogu označiti kao tačni upotrebom mašinskog učenja, što kasnije može dovesti do netačnih zaključaka. Sporedni doprinos istraživanja je priprema podataka za sistem brze dijagnostike na osnovu simptoma, koji u vreme epidemija može ubrzati dijagnostiku i / ili se može koristiti kao kontrolna tabla u službi kontrole epidemije namenjena građanima pametnih gradova. Aktuelna epidemija COVID-19 je dobar primer važnosti kreiranja alata za brzu dijagnostiku. Ovaj sistem bi mogao da se primeni kao javni zdravstveni servis kako bi se građanima pomoglo da sami opisuju simptome i vide verovatnoću da su zaraženi i da li treba da odu u zdravstvenu ustanovu ili ostanu kod kuće i potraže samo onlajn savet lekara.

## 7.1. Pregled postojećih medicinskih resursa

Povezana istraživanja na ovu temu opisuju metode za normalizaciju i izdvajanje znanja iz medicinske evidencije koje nisu direktno povezane sa primenom na srpskom jeziku. Većina istraživanja odnosi se na korpuse engleskog govornog područja i leksičke resurse koji su javno dostupni. Jedan od najpoznatijih kliničkih korpusa za engleski jezik je „Informatics for Integrating Biology & the Bedside (i2b2)“ [156], „MIMIC II“ [157], kao i korpus biomedicinskih tekstova sa označenim opsegom negacije „The BioScope“ [158]. Korpus za medicinski domen pod nazivom „HEALTH BANK“ [160] dostupan je na švedskom jeziku. Za bugarski jezik postoje rezultati izdvajanja informacija iz velikog korpusa nestrukturiranih podataka i njihovog pribavljanja u strukturiranom obliku [170].

Normalizacija je prvi korak koji se koristi u klasifikaciji i obeležavanju medicinskih termina, a postojeće metode za normalizaciju prikazane su u prethodnom poglavlju. Mnogi sistemi koji su izgrađeni za obradu slobodnog teksta u medicinskom domenu koriste metode obrade prirodnog jezika (*NLP*) za njihovu dalju primenu u sistemu zdravstvene zaštite. Obrada medicinskog izveštaja sastoji se od nekoliko koraka kao što su prečišćavanje podataka, integracija, transformacija, redukcija i na kraju zaštita podataka. Glavna svrha je prevođenje polustrukturiranih i nestrukturiranih medicinskih izveštaja u računarski čitljive informacije korišćenjem *NLP* metoda. Ključne metode u ovom procesu su prepoznavanje

imenovanih entiteta (*NER*) i ekstrakcija odnosa (*RE*) [156]. Izvlačenjem kliničkih odnosa iz medicinskih izveštaja mogu se identifikovati odnosi između referenci na lekove i njihovih atributa [171]. Pregled medicinskih informacionih sistema [157] pokazuje da 60% komercijalnih sistema koristi metode zasnovane na pravilima, dok se u naučnim istraživanjima više koriste metode mašinskog učenja. Prema [172], bolji rezultati u izdvajanju teksta iz medicinskih izveštaja dobijaju se metodama zasnovanim na pravilima. Najpopularniji sistemi za obeležavanje medicinskog teksta su *CTAKES* i *CLAMP* sistemi. Kompletna arhitektura sistema *CTAKES* opisana je u [160][173], a sličan sistem *CLAMP* zasnovan na *NLP*-u u radu [174]. Autori su se u radu bavili identifikacijom medicinskih pojmova u tekstovima koje su napisali pacijenti, koristeći *crowdsourcing* [175]. U [176] je prikazan jedan pristup ispravljanju grešaka u slobodnom tekstu medicinskih izveštaja.

Srpski jezik ima složenu gramatiku i blizak je ostalim slovenskim jezicima, pa se sličan pristup obradi teksta može primeniti i na druge jezike koji pripadaju ovoj grupi. Automatsko označavanje dijagnoza u nestrukturiranom medicinskom tekstu sa odgovarajućim leksičkim izvorima za medicinske izraze na bugarskom jeziku izvršeno je u [170]. Na jezicima bivše Jugoslavije nisu pronađeni radovi koji bi se bavili procesom slobodnog teksta u medicinskim izveštajima i javno dostupnim leksičkim izvorima u medicinskom domenu. „Wordnet za biomedicinske nauke“ [176] [177] za srpski jezik sadrži skupove sinonimnih reči ili tačnije različite delove govora (engl. *PoS, parts of speech*) sa novim konceptom za šest ontoloških kategorija (genetika, virus, bakterije, ćelije, naučna polja i mikroorganizam).

Ne postoje javno dostupni leksički izvori klasifikovanih medicinskih termina (dijagnoze, lekovi, simptomi itd.). Nezavisno od medicinskog domena, u radu [179] je opisana metoda za označavanje semantičkih uloga na slovenačkom i hrvatskom jeziku. Složena gramatika i dva pisma (latinica i ćirilica) otežavaju normalizaciju dokumenata na srpskom jeziku. Algoritam za normalizaciju medicinskih dokumenata na srpskom jeziku predstavljen je u [180]. Od radova koji se bave ovim problemom, nema radova koji se direktno odnose na medicinsku oblast. Prepoznavanjem imenovanih entiteta u srpskom jeziku autori su se bavili u radu [178].

Pristup obrade prirodnog jezika zasnovan na tagerima i metodama mašinskog učenja primenjenim na tekstove na srpskom jeziku prikazan je u radovima [181][182]. Tageri se mogu koristiti za klasifikaciju pojmova u grupe sa različitim oznakama. Ovi radovi opisuju

javno dostupne alate, *Stanford POS Tagger* [183] i *Tree Tagger* [184]. Zbog toga su korišćeni za poređenje sa predloženim metodama, kao što je prikazano u odeljku sa eksperimentima i rezultatima.

## 7.2. Skup podataka

Ne postoji ni javno dostupan korpus medicinskih izveštaja na srpskom jeziku, niti rad koji se bavi automatskim izdvajanjem informacija iz medicinskih tekstova na srpskom jeziku. Iz medicinskih centara grada Niša u Srbiji korišćena su dva skupa podataka. Ovi medicinski izveštaji su prikupljeni iz informacionog sistema *MEDIS.NET* u periodu 2012-2018. Primarni skup podataka sadrži 2212 medicinskih izveštaja, koji je korišćen i u prethodnom poglavlju. U jednom delu ovog perioda bila je epidemija malih boginja u gradu Nišu, pa ovaj skup podataka sadrži *EMR*-ove za ovu dijagnozu. Obrada ovog skupa podataka je važna za analizu epidemije. Kontrolni skup podataka sastoji se od 2000 medicinskih izveštaja sa deset različitih vrsta dijagnoza koje se koriste za validaciju rezultata nadgledane metode. Korpusi su stvoreni u skladu sa etičkim standardima, uz deidentifikaciju pacijenta i medicinskog osoblja. Korpusi su ručno označeni od strane četiri nezavisna anotatora, koji su medicinski stručnjaci. Slaganje anotatora je 93,3% i izračunato je mernom nominalnom skalom za slaganje između više anotatora [185]. Svakom *EMR*-u dodeljen je dijagnostički kod, kao što je prikazano u tabeli Tabela 4 iz prethodnog poglavlja. U primarnom skupu, svim dijagnozama je dodeljen kod B05, dok je drugi skup sastavljen od *EMR*-a koji su spojeni sa 10 različitih dijagnoza (B00, B01, B02, H10, H650, H66, J11, J18, N390, S60) ravnomerno raspoređenih u skupu (10% za svaku klasu). Pored dijagnoza povezanih sa *EMR*-om, u nestrukturiranom delu *EMR*-a, postoje dijagnoze bolesti koje se javljaju kao komplikacije odgovarajuće bolesti, kao i dijagnoze pratećih bolesti pacijenta. Dakle, u primarnom skupu, iako su zapisi povezani sa istim kodom dijagnoze B05, pojmovi (koji se odnose na dijagnoze 29, simptome 116, terapije 158 i specijalnosti 11) pojavljuju se u različitim oblicima (kod, ime, sinonim, skraćena).

### 7.2.1.1. Karakteristike primarnog skupa podataka

Nestrukturirani deo ovih izveštaja sastoji se od anamneze napisane u vidu slobodnog teksta. Iz ovog dela teže je izvući relevantne podatke jer se on mora transformisati u standardizovani format, pogodan za dalju obradu.

Pre označavanja, bilo je neophodno da se izvrši normalizacija anamneza naznačena u sekciji 7.3.2.1. Na ovaj način izdvojeno je 25425 reči, pa je prosečan broj reči po anamnezi bio 11,49. Sve reči sadržane u nestrukturiranom delu medicinskih izveštaja kategorisane su kao reči iz medicinskog domena (simptomi, opisi simptoma, dijagnoze, biohemijske analize, latinske reči, anatomski imena organa, terapije i drugi medicinski izrazi) i one izvan medicinskog domena (brojevi, simboli negacije, stop reči i druge reči bez značaja za zdravstveno stanje pacijenta). Ručnim obeležavanjem, mali broj (2,74%) reči nije klasifikovan ni u jednu grupu, ni u greške ili skraćenice.

Prilikom obeležavanja korpusa primećeno je da su skraćenice i greške prilično prisutne u tekstu (12,9% skraćenica i 4,46% grešaka, što ukupno čini 17,36%). Najčešće skraćenice bile su standardne (često za laboratorijske analize, npr. „wbc“ - bela krvna zrnca), ali bilo je i puno ličnih skraćenica (npr. „izv“ (izveštaj), „temp“ (temperatura) itd.). Za isti pojam korišćene su različite vrste skraćenica (npr. „t“, „T\*“, „temp“ za termin „temperatura“). Greške koje su se dogodile klasifikovane su u 11 vrsta u cilju njihovog lakšeg korigovanja. Vrste grešaka koje su se javile su: izostavljeno dvostruko slovo (T1), zamena mesta slova (T2), dodatna slova (T3), nedostajuća slova (T4), zamena sličnom rečju (T5), spojene reči bez razmaka (T6), spojene reči sa slučajnim slovom umesto razmaka (T7), izostavljen (zamenjen) dijakritički simbol (npr. „c“ umesto „č“ ili „ć“) (T8), netačno slovo (T9), upotreba slova koja ne pripadaju srpskoj abecedi („x“ umesto „ks“) (T10) i više grešaka u jednoj reči (T11).

Slobodno napisan deo medicinskog izveštaja često uključuje brojeve koji ukazuju na vrednosti laboratorijskih analiza, terapija i opise nekih specifičnosti pacijenata (kao što su detalji lične istorije prethodnog lečenja, porodične istorije, zaposlenja itd.), izveštaji o oporavku gde se zaključuje bolovanje i daju doznake pacijentu.

## 7.3. Korišćeni resursi

Pristup medicinskim informacijama je ograničen i teško ih je pronaći u formatu pogodnom za upotrebu u automatskoj obradi medicinskog teksta. Nedostaju i specijalizovani medicinski resursi koji sadrže medicinske termine i izraze.

Za potrebe ovog istraživanja kreiran je skup leksičkih izvora na srpskom jeziku koji su specifični za medicinsku obradu teksta. Mogu se podeliti u četiri grupe: dijagnoze, terapije, simptomi i dijagnostika, specijalnosti.

Neki opšti leksički izvori (lična imena i imena naselja, stop reči i simboli negacije) takođe su korišćeni za uklanjanje reči koje nisu medicinski relevantne. Svi predstavljeni medicinski izvori dobijaju se automatskom obradom medicinskih dokumenata koji se mogu naći u bilo kom medicinskom informacionom sistemu ili su javno dostupni. Značaj ovog pristupa je u jezičkoj nezavisnosti procesa za dobijanje takvih resursa i sposobnosti da se lako prilagode drugom jeziku.

Cilj stvaranja ovih medicinskih resursa je označavanje pojmova u elektronskim medicinskim izveštajima. Nestrukturirani deo podataka sadrži opisne informacije do kojih je došao lekar. Simptomi, dijagnostika, dijagnoza i propisana terapija ističu se kao važne medicinske informacije. Predloženi leksički izvori mogu se koristiti za identifikaciju navedenih medicinskih pojmova na nivou reči. Korišćenjem istih resursa moguće je identifikovati imenovane entitete, što nije obuhvaćeno ovim istraživanjem i koje će biti predmet interesovanja u budućnosti.

### 7.3.1. Nemedicinski resursi

Obrada teksta na prirodnom jeziku zahteva resurse za izdvajanje bitnih reči od nebitnih činjenica. Svi lični podaci koji se odnose na pacijente i lekare moraju se izdvojiti iz podataka, za koje postoje dostupni resursi [163][186]. Često medicinski izveštaji uključuju imena i lokaciju zdravstvenih centara i gradova; stoga je stvoren resurs sa imenima naselja u Srbiji. Pretpostavka je da imenice, pridevi i glagoli nose sadržaj teksta, a da su druge vrste

reči manje važne za informativnu vrednost teksta (stop reči). Ovaj skup reči sadrži reči sa velikom frekvencijom u većini dokumenata, veznike, uzvike, rečce itd. Ovaj skup je prikupljen u svrhu normalizacije dokumenata na srpskom jeziku [143].

Označavanje opsega negacije je važno u medicinskim kartonima. U svrhu označavanja simbola negacije, korišćen je skup simbola negacije zajedno sa skupom pravila predloženih u radu [166].

Svrha obeležavanja nije bila da se označi opseg negacije; ali upotrebom tokenizacije za označavanje simbola negacije, tokenizacija se prilagođava pravilima negacije. Dakle, naznačen je i opseg delovanja odgovarajućeg simbola negacije. Važno je napomenuti da se u medicinskim izveštajima najčešće pojavljuju samo jedan ili dva oblika negacije koji se lako obrađuju.

### **7.3.2.Obrada medicinskih leksičkih resursa**

Ne postoje javno dostupni specijalizovani medicinski leksički resursi za srpski jezik. Stoga su dostupni dokumenti u različitim formatima korišćeni za obradu pojmova specifičnih za medicinski domen. Tekstualni podaci u izvornom obliku moraju biti podvrgnuti predobradi da bi se mogli koristiti.

#### **7.3.2.1.Normalizacija**

Proces pripreme medicinskih resursa uključuje uklanjanje nebitnih informacija. Još jedna specifičnost srpskog jezika je upotreba dva pisma (latinice i ćirilice). Zbog toga je bilo potrebno sve podatke prebaciti u jedno pismo kako bi se omogućila dalja uniformna obrada. Srpska abeceda razlikuje se od engleske jer postoje slova sa dijakritičkim simbolima (ć, č, đ, ž, š i dž). U srpskom pismu se ne koriste slova „x“, „y“, „z“ i „q“.

Proces normalizacije sastoji se od sledećih koraka:

(1) uklanjanje svih suvišnih delova iz teksta i priprema za tokenizaciju (podela teksta na reči):

(1.1) konverzija iz ćirilice u latinicu,

(1.2) promena dijakritičkih znaka u kombinacije slova (ć, č, ž, š, đ) ⇒ (cx, cw, zx, sx, dx); ovo mapiranje je korektno jer slova „x“ i „w“ ne pripadaju srpskoj abecedi,

(1.3) označavanje separatora rečenica ('.', '!', ';', '-', '!', '?', '(', ')', ',', ' '),

(1.4) uklanjanje nepotrebnih razmaka, svih ostalih simbola, osim slova, brojeva i separatora rečenice;

(2) podela originalnog teksta na rečenične jedinice,

(3) podela rečeničnih jedinica na reči uz održavanje veze sa rečeničnom jedinicom,

(4) obeležavanje stop reči,

(5) obeležavanje simbola negacije i

(6) označavanje numeričkih vrednosti.

### 7.3.3. Medicinski resursi

Leksički izvor specifičan za određeni medicinski domen je skup različitih izraza koji se izdvajaju obradom odgovarajućih medicinskih dokumenata, kao što je opisano u prethodnom odeljku.

#### 7.3.3.1. Dijagnoze

*ICD-10* (Međunarodna statistička klasifikacija bolesti i srodni zdravstveni problemi) je standard za klasifikaciju i kodiranje bolesti i medicinskih problema i koristi se u medicinske i administrativne svrhe [167][168]. Javno je dostupan i preveden na nekoliko jezika. Sadrži šifre bolesti, opis bolesti, simptome i znake, socijalne prilike i spoljne uzroke bolesti i još mnogo toga. Početna klasifikacija sadrži oko 14000 dijagnoza. Proširene verzije i nacionalna izdanja ove klasifikacije sadrže više dijagnoza.

Inicijalni set je skup dijagnoza datih u [168], koji sadrži dijagnoze prevedene na srpski jezik i uključuje 14194 dijagnoze. Prethodnom obradom imena dijagnoza izdvojen je skup od 72652 reči, nakon čega je smanjen na 7942 različite reči. Treba naglasiti da naziv dijagnoze takođe uključuje simptome, anatomske delove, uzroke bolesti i mnoge medicinske i nemedicinske izraze. Međutim, ovaj skup takođe sadrži mnoge medicinske izraze koji se



često javljaju, pa je teško razlikovati nemedicinske pojmove koji nisu značajni za obeležavanje. Potrebni su dodatni resursi za označavanje ovih medicinskih termina.

Pored naziva dijagnoze, postoji i resurs sa *ICD-10* kodovima dijagnoze koji se takođe koriste prilikom pisanja medicinskih izveštaja. Izdvojen je set od 14194 šifre. U medicinskim izveštajima upotreba latinskog jezika je česta, pa je tako stvoren resurs sa latinskim imenima dijagnoza za one sa latinskim imenom (3794 dijagnoze). Ovaj resurs takođe uključuje neka latinska imena za anatomske organe, simptome i još mnogo toga. Ovaj set za naknadnu obradu (tokenizacija bez izuzimanja srpskih nemedicinskih termina) sadrži 2844 latiničnih izraza.

### **7.3.3.2. Terapije**

Terapija ili medicinsko lečenje je pokušaj popravljavanja zdravstvenog stanja nakon dijagnoze. Jedna vrsta terapije uključuje propisivanje lekova i ampula. Lekovi su opisani fabričkim nazivom, generičkim nazivom i ATC (Anatomsko-terapijski sistem hemijske klasifikacije). Izvori za terapiju sastoje se od 2232 leka i 1317 ampula. Nakon obrade ovih osnovnih podataka, formirano je 2255 različitih termina koji čine resurs terapija.

### **7.3.3.3. Simptomi i dijagnostika**

Mnogi simptomi bolesti već su obuhvaćeni nizom dijagnoza. Za potrebe ovog istraživanja prikupljeni su dokumenti koji opisuju simptome morbila (B05) [187][188] zbog specifičnosti skupa podataka. Početni skup sastoji se od 28 odvojenih opisa simptoma koji su obradom svedeni na 95 različitih reči. Pored simptoma, za određivanje dijagnoze korišćeni su i biohemijski rezultati. Obraden je resurs sa različitim vrstama biohemijskih analiza dostupnih u posmatranom informacionom sistemu *MEDIS.NET* i dobijeno je 63 izraza iz domena biohemije krvi.

### **7.3.3.4. Specijalnosti**

Medicinski izveštaji često uključuju izveštaje sa prethodnih tretmana ili drugih medicinskih ustanova, kao i preporuke drugim specijalistima za dalje lečenje u drugim specijalizovanim odeljenjima. Da bi se identifikovala ova vrsta reči, obrađeni su nazivi dostupnih specijalizacija, kao i zdravstvenih usluga različitih specijalnosti koje su preuzete iz

medicinskog informacionog sistema. Dobijeni resursi sadržali su 41 termin koji se odnosi na specijalizaciju i 41 termin povezan sa zdravstvenom službom dotične specijalnosti.

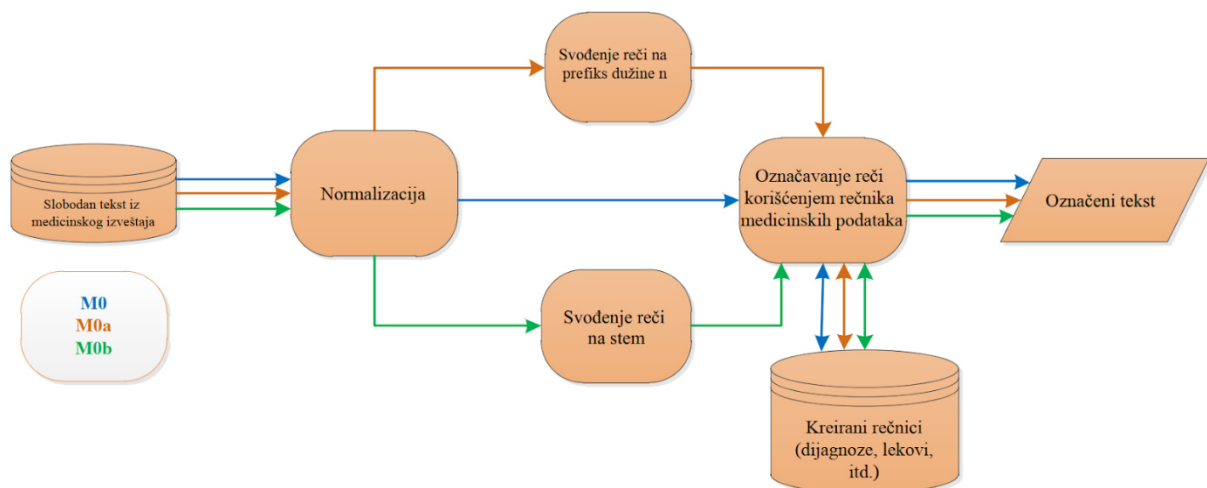
## **7.4. Predložene metode za obeležavanje medicinskih termina u slobodnom tekstu**

Nakon označavanja, skup podataka podeljen je u dve grupe. Dve trećine anamneza se koriste kao trening skup. Poslednja trećina podataka koristi se kao test, a zadatak je obeležiti ovu grupu pomoću tri predložene metode. Prva metoda (metoda *M0*) zasniva se na rečnicima medicinskih i nemedicinskih termina. Druga metoda (metoda *M1*) označava pojmove pomoću trening skupa, a treća metoda (metoda *M2*) je hibridna, koristi i trening skup i pravila za obeležavanje pojmova. Metode *M1* i *M2* ne zahtevaju rečnike medicinskih izraza. Za svaku metodu kreirane su njene varijacije. Razlike u varijacijama su način na koji se reč svodi na njen osnovni oblik (*a* - odsecanje na prefiks dužine *n*, *b* - upotrebom stema). Razlog zašto se to radi je postojanje deklinacija (promena imenskih reči po padežu) i konjugacija (promena glagolskih oblika po vremenima, rodu i broju) u srpskom jeziku, pri čemu se promenljive vrste reči mogu naći u mnogim oblicima. U predloženim metodama, prefiks dužine *n* je niz od *n* prvih slova reči. Odsecanjem na prefiks dužine *n*, uklanja se sufiks duge reči, a zatim se ostatak reči traži u resursima kao početak reči. Reči kraće od *n* slova takođe su potrebne kao početak reči. Tako se 4 koristi za vrednost *n*, jer se veliki broj predloga na srpskom jeziku sastoji od tri slova, kao i iz razloga objašnjenih u prethodnom poglavlju o normalizaciji reči. Stemovanje je metoda zavisna od jezika, koja se koristi za identifikovanje i zamenu sufiksa (nastavak za oblik reči) odgovarajućim nastavkom za osnovni oblik reči (stem). Postupak stemovanja vrši se upotrebom stemera za srpski jezik [189].

### **7.4.1. Metoda zasnovana na jezičkim resursima**

U ovoj metodi (*M0*) tekst je obeležen korišćenjem kreiranih medicinskih i nemedicinskih izvora opisanih u odeljku 7.3. Slobodni tekst iz podataka iz testa treba automatski označiti. Uzet je slobodan tekst iz anamneze, a na anamnezi se izvode tokenizacija

i normalizacija opisane u odeljku 7.3.2. Ovaj deo je zajednički i metodama  $M1$  i  $M2$  i njihovim varijacijama koje će biti opisane u narednim poglavljima. Nakon tokenizacije, svaka reč se traži u rečnicima. Čim se pronade u jednom od resursa, pretraga se završava i obrađuju se sledeće reči. Metod  $M0a$  je varijacija ove metode gde se originalna reč odseca posle prva četiri slova i traži se kao početak reči u rečnicima. Metod  $M0b$  je varijacija metode  $M0$ , gde je reč zamenjena svojim stemom i pretraživana u rečnicima. Postupak za izvršenje metoda  $M0$ ,  $M0a$  i  $M0b$  detaljno je prikazan na slici Slika 57.

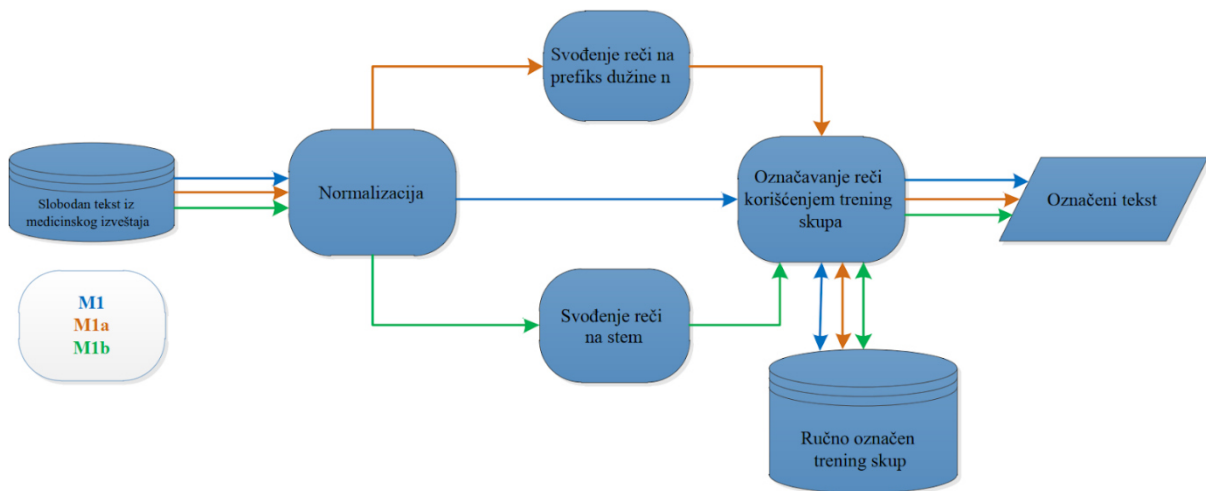


Slika 57. Koraci obrade u metodi  $M0$

#### 7.4.2. Metoda zasnovana na nadgledanom učenju

U metodi  $M1$ , nakon normalizacije, reč se označava korišćenjem podataka iz trening skupa. Označavanje reči  $v$  oznakom  $l$  vrši se uzimajući u obzir pravilo izdvojeno iz modela sa najvećim poverenjem ( $Conf$ ), predstavljeno u formuli 8. Ako je reč označena kao greška u modelu, ona se ne uzima u obzir, tj. reči u testu su označene samo kada se zasnivaju na tačnim rečima u modelu. Metoda  $M1$  ne uzima u obzir lične skraćenice, jer one ne pripadaju domenu tačno napisanih reči. U metodi  $M1$  pretražuje se reč u izvornom obliku; u metodi  $M1a$  pretražuje se prefiks reči dužine 4, dok se u metodi  $M1b$  vrši pretraga stema u označenom modelu. Postupak izvođenja metoda  $M1$ ,  $M1a$  i  $M1b$  detaljno je prikazan na slici Slika 58.

$$Conf(v \rightarrow l) = \frac{\text{broj reči } v \text{ označenih labelom } l \text{ u modelu podataka}}{\text{broj reči } v \text{ u modelu podataka}} \quad (8)$$

Slika 58. Koraci obrade u metodi *M1*

### 7.4.3. Metoda zasnovana na nadgledanom učenju i pravilima

U metodi *M2* reč se označava na osnovu obeleženog skupa podataka, ali su dodatno napravljena pravila za prepoznavanje grešaka i skraćenica i njihovu obradu (Slika 59). Ova metoda predstavlja proširenje metode *M1*. U ovoj metodi reči koje su pravilno napisane označene su na osnovu modela kao u metodi *M1*. Međutim, sve reči bez oznake nakon tog koraka prolaze kroz još nekoliko koraka da bi se pronašla njihova oznaka. Stoga neobeležena reč prolazi kroz sledeće korake:

a. Ispravka greške T8. Najčešći tip grešaka je izostavljeni dijakritički simbol. U ovom koraku se proverava da li neobeležena reč sadrži jedno od slova „c“, „s“, „z“, a ako je tako, zamenjuje se svojim dijakritičkim parom „ć“, „č“, „š“, „ž“, i posebna grupa slova „dj“ se menja znakom „đ“. Ako posle promene slova reč ostane neobeležena, postupak se ponavlja sve dok se ne zameni svaki potencijalni dijakritički simbol. Ovde su primenjena slična pravila kao u radu [190].

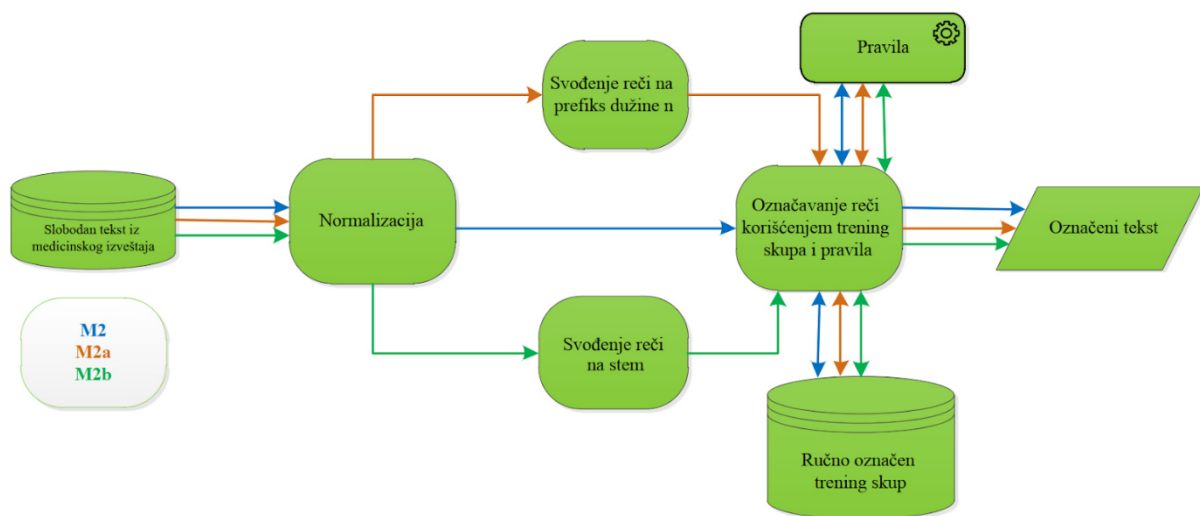
b. Ispravka greške T10. Slova poput „x“, „w“, „q“ i „y“ ne pripadaju srpskoj abecedi. Primećuje se da postoje primeri njihove netačne upotrebe u srpskim rečima (slovo „x“ je skraćeno od grupa slova „ks“, slovo „y“ zbog specifičnosti tastature piše se umesto „z“). U ovom koraku, ako tražena reč nije obeležena kao latinski izraz, slova se zamenjuju grupom

odgovarajućih slova ili jednim slovom iz srpske abecede i pokušava se naći oznaka za promenjenu reč.

c. Označavanje ličnih skraćenica. Lične skraćenice se obrađuju pretraživanjem na početku reči iz modela podataka, pa ako se takva reč pronađe, neobeležena reč označava se kao pronađena reč.

d. Označavanje reči sa greškama T6 i T7. U ovom koraku obeležavaju se dve spojene reči, traženjem reči iz modela podataka koji je podniz ovog skupa od dve reči; ako je pronađena, skup je označen kao pronađena reč.

e. Označavanje reči sa drugim vrstama grešaka. Ako je nakon svih prethodnih koraka reč i dalje neobeležena, u poslednjem koraku pokušava se da se označi neobeležena reč na osnovu reči koje zvuče kao neobeležena reč. To se postiže izmenom funkcije *SOUNDEX* prilagođenom pravilima za sličnost slova u srpskom jeziku [162]. Na taj način se ispravljaju neke greške nastale zamenom dva slova, dupliranjem slova, odsustvom ili dodavanjem jednog slova ili zamenom sličnom rečju. Kako reč može dobiti pogrešnu oznaku u ovom koraku, važno je da se ovaj korak završi na kraju obrade, nakon što se iscrpe sve druge mogućnosti za označavanje.



Slika 59. Koraci obrade u metodi M2

Pravila su data po redu primene. Ako se na jednu reč može primeniti više pravila, pravila se primenjuju po prioritetu kako bi se reč manje modifikovala, tj. da bi se smanjila mogućnost dodeljivanja pogrešne oznake. Metoda *M2* takođe ima varijacije *M2a* i *M2b* (Slika 59).

Primeri obeleženih delova anamneze primenom metode *M2* prikazani su na slici Slika 60.

Originalna anamneza	Označena anamneza
Ždrelo hiperemicno,nadplucioma grubljre disanje	zxdrelo hiperemixno nadpluxima grubljre disanje
Kaslje od juce,promukla,slina oci joj krmeljaju,temp. nije imala	kaslje od juce promukla slina oci joj krmeljaju temperatura nije imala
Izvestaj dermatologa	izvesxtaj dermatologa
<b>Oznake:</b> <i>anatomski organ</i> , <i>simptom</i> , <i>specijalnost</i> , <i>drugi medicinski temin</i> , <i>nemedicinski izraz</i> , <i>neobeleženo</i> , <i>ispravljeno korišćenjem metode M2</i>	

Slika 60. Primeri obeležene anamneze primenom metode *M2*

## 7.5. Metrike korišćene za evaluaciju metoda

Metrike koje su korišćene za upoređivanje rezultata metoda *M0*, *M1* i *M2* su *Preciznost*, *Odziv*, *F1-rezultat* i *Tačnost*. One su definisane formulama 9, 10, 11, 12, respektivno.

$$Preciznost = \frac{BKOT}{BOT} \quad (9)$$

$$Odziv = \frac{BKOT}{BTO} \quad (10)$$

$$F1 - Rezultat = \frac{2 * Preciznost * Odziv}{Preciznost + Odziv} \quad (11)$$

$$Tačnost = \frac{BKOT}{BT} \quad (12)$$

Ovde je *BKOT* broj tačno označenih termina, *BOT* je broj označenih termina, *BTO* je broj termina koji treba da budu označeni, a *BT* je ukupan broj termina. Za klasifikaciju u više klasa, sve kategorije pojmova uzimaju se zajedno koristeći mikro metrike (metrike se

izračunavaju globalno brojeći ukupne istinite pozitivne, lažne negativne i lažne pozitivne rezultate).

Rezultati dobijeni samo za medicinske termine su takođe predstavljeni, korišćenjem metrika definisanim formulama 13 i 14.

$$Preciznost = \frac{BKOMT}{BOMT} \quad (13)$$

$$Odziv = \frac{BKOMT}{BMTO} \quad (14)$$

$$RelativnoPoboljšanje = \frac{PoredbenaMetoda - OsnovnaMetoda}{OsnovnaMetoda} * 100 \quad (15)$$

Ovde je *BKOMT* broj tačno označenih medicinskih termina, *BOMT* je broj termina označenih kao medicinski, a *BMTO* je broj medicinskih termina koji treba da budu označeni. Za upoređivanje metoda korišćena je i relativna mera poboljšanja, data formulom 15 i ovde se upoređuju *F1-rezultati* metoda.

## 7.6. Evaluacija metoda za obeležavanje medicinskih termina na srpskom jeziku

Rezultati primene metoda i njihovih varijacija, koji pokazuju uspešnost metoda u označavanju medicinskih i nemedicinskih termina, predstavljeni su u tabeli Tabela 13.

Metoda *MO* koristi resurse medicinskih izraza koji pripadaju skupovima dijagnoza, ampula, lekova, specijalizacija itd. Reči u tim izvorima napisane su tačno, sa dijakritičkim simbolima. Rezultati eksperimenta pokazuju da je na osnovu takvih resursa manji broj reči iz skupa podataka označen u poređenju sa drugim metodama. *Preciznost* metode *MO* je najbolja kada se reč pretražuje bez svođenja na osnovni oblik. Metodama *MOa* i *MOb*, reč se svodi na

osnovu, povećavajući tako broj označenih pojmova, ali je *Preciznost* niža, jer se osnova reči često svrstava u slične reči bez oznake. Ipak, *F1-rezultat* i *Tačnost* su najbolji za *M0a* (relativno poboljšanje vrednosti *F1-rezultat* u odnosu na metodu *M0* iznosi 19,29%) gde se reč traži kao prefiks dužine 4 kao početak reči iz resursa.

Tabela 13. Rezultati označavanja medicinskih i nemedicinskih termina

Metoda	<i>BT</i>	<i>BTO</i>	<i>BKO</i> <i>T</i>	<i>BOT</i>	<i>Preciznost</i>	<i>Odziv</i>	<i>F1-rezultat</i>	<i>Tačnost</i>
<i>M0</i>	10969	10883	5119	5895	0.86836	0.47037	0.6102	0.46668
<i>M0a</i>			7717	10320	0.74777	0.70909	0.72792	0.70353
<i>M0b</i>			6965	9013	0.77277	0.63999	0.70014	0.63497
<i>M1</i>			8415	8429	0.99834	0.77322	0.87148	0.76716
<i>M1a</i>			9424	10341	0.91132	0.86594	0.88805	0.85915
<i>M1b</i>			9234	9927	0.93019	0.84848	0.88746	0.84183
<i>M2</i>			9612	10285	0.93457	0.88321	0.90816	0.87629
<i>M2a</i>			9569	10656	0.89799	0.87926	0.88853	0.87237
<i>M2b</i>			9609	10522	0.91323	0.88294	0.89783	0.87601

Razlog tome je što se traženjem osnove reči označavaju i lične skraćenice i neke vrste reči sa greškama 6 i 7 (pošto je greška koja se dogodi na kraju reči odsečena), mada ponekad i netačno. Razlog zbog kojeg metoda *M0b* daje lošije rezultate je sledeći: nastavak za oblik reči se ne eliminiše već se zamenjuje odgovarajućim nastavkom za oblik da bi se dobila osnova, a reči su ređe označene u tom osnovnom obliku. Medicinski resursi korišćeni u metodi *M0* takođe sadrže nemedicinske izraze kako je opisano u odeljku 7.3. Uobičajena vrsta greške u ovoj metodi je označavanje nemedicinskih izraza kao medicinskih. Većina nemedicinskih izraza koegzistira sa drugim medicinskim terminima u opisu dijagnoze. Na primer, u anamnezi „Dobio boginje od majke“, nemedicinski termin „majke“ je označen kao medicinski zbog dijagnoze O75.0 („iscrpljenost majke tokom porođaja“). Resurs dijagnoza sadrži najveći broj reči, što je razlog pogrešnog označavanja. Od ukupnog pogrešnog označavanja, ova vrsta pogrešnog označavanja iznosila je 83,12%. Metodom *M1* postižu se bolji rezultati od metode *M0* (relativno poboljšanje vrednosti *F1-rezultat* u odnosu na metodu *M0* iznosi 42,82%). Takođe, od svih metoda, *M1* ima najbolju *Preciznost* jer je reč pronađena u modelu u originalnom obliku pravilno označena. *Odziv* je takođe veći nego u *M0*, što ukazuje da je dovoljno obeležiti jedan deo izveštaja za bolest, jer se reči u njima obično ponavljaju.



Svođenjem reči na osnovni oblik, metode *M1a* i *M1b* daju bolje rezultate od metode *M1*. *M1a* pokazuje najveće poboljšanje jer odsecanje na prefiks veličine  $n$  rešava deo problema sa skraćenicama i greškama, kako je objašnjeno u ovoj varijaciji u metodi *M0*.

Metoda *M2* označava tačnu reč na isti način kao i *M1*. Međutim, metoda *M2* se posebno bavi označavanjem reči koje su pogrešno napisane ili skraćene. *Preciznost* je nešto niža nego u *M1*, međutim, *Odziv* se značajno povećava kao i *F1-rezultat* (relativno poboljšanje vrednosti *F1-rezultat* u poređenju sa metodom *M1* iznosi 4,21%). To znači da su mnoge skraćene i netačne reči ispravljene i, u većini slučajeva, pravilno označene. *M2a* i *M2b* ne pružaju bolje rezultate nego *M2*, što ukazuje da predložena pravila o greškama i skraćenicama delimično pokrivaju svođenje reči na osnovu, pa ovaj korak postaje beskoristan i postupak označavanja je ubrzan. Za metodu *M2* koja ima najbolji *F1-rezultat*, vrši se i trostruka unakrsna validacija. Razlika dobijena primenom metode na različitim delovima skupa podataka nije statistički značajna ( $p = 0.013185 > 0.01$ , upareni T.Test [191]). Ovo pokazuje da je metoda *M2* jezički nezavisna, jer se obeležavanje vrši na osnovu označenog modela podataka i pravila za obradu grešaka i skraćenica, koja se mogu prilagoditi prema jezičkim specifičnostima. Za slične jezike metoda se može primeniti direktno ili sa manjim adaptacijama.

Tabela 14 prikazuje rezultate primene metoda samo nad medicinskim terminima. U ovom slučaju *M2* ponovo pruža najbolje rezultate, vrednosti *F1-rezultata* su bliske rezultatima iz prethodne tabele. Nešto niži rezultati za *Preciznost* i *Odziv* za metodu *M0* prouzrokovani su pogrešnim označavanjem nemedicinskih izraza kao medicinskih, kao što je ranije opisano.

Tabela 14. Rezultati primene metoda nad medicinskim terminima

Metoda	BMTO	BKOMT	BOMT	Preciznost	Odziv	F1-rezultat
<i>M0</i>	5692	1920	2565	0.74854	0.33732	0.46506
<i>M0a</i>		3890	5125	0.75902	0.68342	0.71924
<i>M0b</i>		3174	4108	0.77264	0.55763	0.64776
<i>M1</i>		3994	4007	0.99676	0.70169	0.82359
<i>M1a</i>		4746	5075	0.93517	0.8338	0.88158
<i>M1b</i>		4793	5262	0.91087	0.84206	0.87511
<i>M2</i>		4731	4865	0.97246	0.83117	0.89628
<i>M2a</i>		4787	5139	0.9315	0.84101	0.88394
<i>M2b</i>		5013	5528	0.90684	0.88071	0.89358

Korišćenjem MekNemarovog testa [191], izračunava se statistički značaj razlike u označavanju termina različitim metodama. Tabela 15 prikazuje razliku u primeni metode sa i bez svođenja reči na osnovni oblik. Kao što se može videti iz tabele u slučaju metoda  $M0$  i  $M1$ , ova razlika je značajna u korist metode koja koristi odsecanje na prefiks veličine  $n$ . U slučaju metode  $M2$ , ne postoji statistički značajna razlika u rezultatima dobijenim metodama  $M2$ ,  $M2a$  i  $M2b$ .

Tabela 15. Statistički značaj poboljšanja metode  $M2$  u odnosu na druge metode

Osnovna metoda	Poredbena metoda	Bolja metoda	<i>Hi kvadrat</i>	<i>P</i>
$M0$	$M0a$	$M0a$	1832.094	0
$M0$	$M0b$	$M0b$	1174.503	0
$M0a$	$M0b$	$M0a$	248.9091	0
$M1$	$M1a$	$M1a$	437.9769	0
$M1$	$M1b$	$M1b$	344.4994	0
$M1a$	$M1b$	$M1a$	31.42694	0
$M2$	$M2a$	$M2$	2.441485	0.11817
$M2$	$M2b$	$M2$	0.082234	0.77429
$M2a$	$M2b$	$M2b$	2.410023	0.12465

Tabela 16. Statistički značaj razlika u označavanju pojmova

Poredbena metoda	Osnovna metoda	Bolja metoda	<i>Hi kvadrat</i>	<i>P</i>
$M2$	$M0$	$M2$	4010.383	0
$M2$	$M0a$	$M2$	1117.108	0
$M2$	$M1$	$M2$	1020.6	0
$M2$	$M1a$	$M2$	23.44428	0.000001

Kako su najbolji *F1-rezultat* i *Tačnost* dobijeni metodom  $M2$ , Tabela 16 prikazuje rezultate statističkog značaja poboljšanja obeležavanja dobijenih ovom metodom u odnosu na metode  $M0$ ,  $M0a$ ,  $M1$ ,  $M1a$ . U sva četiri slučaja postoji statistički značajno poboljšanje  $M2$  (bez svođenja na osnovni oblik) u poređenju sa drugim metodama.

Važnost metoda  $M1$  i  $M2$  dodatno je proverena na skupu od 2000 medicinskih izveštaja, koji opisuju različite dijagnoze. Za *F1-rezultat* su dobijeni slični rezultati kao u primarnom skupu podataka (Tabela 17). Ovde je dobijen najbolji *F1-rezultat* za  $M2$ , sa relativnim poboljšanjem od 8,34% u poređenju sa  $M1$ . Dakle, može se zaključiti da  $M2$  daje

najbolje rezultate nezavisno od broja dijagnoza u korpusu. Dobijeni rezultati su uporedivi sa rezultatima za druge jezike (npr. bugarski i engleski) [170][175].

Tabela 17. Rezultati nadgledanih metoda nad kontrolnim skupom podataka

<b>Metoda</b>	<b><i>M1</i></b>	<b><i>M1a</i></b>	<b><i>M1b</i></b>	<b><i>M2</i></b>	<b><i>M2a</i></b>	<b><i>M2b</i></b>
<b><i>BOT</i></b>	4478	5422	5258	5528	5671	5570
<b><i>BKOT</i></b>	4467	4891	4492	5332	5045	4664
<b><i>Preciznost</i></b>	0.9975	0.9021	0.8543	0.9645	0.8896	0.8373
<b><i>Odziv</i></b>	0.7640	0.8365	0.7683	0.9119	0.8628	0.7977
<b><i>F1-rezultat</i></b>	0.8653	0.8680	0.8090	0.9375	0.8760	0.8170

U nastavku su predstavljeni rezultati primene metoda i alata koji se koriste za obeležavanje pojmova na srpskom jeziku [181][182], kao i metode nadgledane klasifikacije u više klasa koju vrši alat *Weka 3.8.4* [192]. Ove metode su primenjene na primarni obeleženi skup podataka proširen dodatnim atributima. Tabela 18 prikazuje rezultate metoda zavisnih od jezika, nadgledane klasifikacije na osnovu *POS* oznaka i metoda nezavisnih od jezika, za prilagođeno označavanje i nadgledane klasifikacije na osnovu *n*-gram karaktera.

Metoda (*NBMT + POS oznake*) koristi *Stanford POS* alat za označavanje koja ima model za srpski jezik. *Stanford POS tager* je alat koji svakoj reči (i drugim tokenima) dodeljuje delove govora, kao što su imenica, glagol, pridev itd. Obeleženi skup podataka proširen je *POS* atributom i primenjena je *NBMT* (engl. *Naive Bayes Multinomial Text*) metoda za klasifikaciju pojmova u odgovarajuće oznake. Metoda označavanja tagerom za prilagođeno označavanje koristi alat za označavanje (engl. *Tree Tagger*) koji ima mogućnost treniranja modela sa prilagođenim oznakama, sa označenim leksikonom i listom oznaka kao ulazom. Tada se dobija istrenirani model koji se može koristiti za obeležavanje podataka u test skupovima. Nadgledana klasifikacija se zasniva na *n*-gram karakteru koristeći metode *naivni Bajes (NB)*, *logističku regresiju (LR)*, *metodu podržanih vektora (SMO)* i *metodu slučajnih šuma (RF)*, istrenirane za svaku oznaku u označenom skupu podataka sa dodatnim atributima 3-gram (3-gram klasifikator) i 4-gram (4-gram klasifikator). *F1-rezultati* dobijeni korišćenjem ovih metoda predstavljeni su u tabeli Tabela 19, a najbolji rezultat dobijen je označavanjem pomoću alata *Tree Tagger*. U poređenju sa predloženim *M1* i *M2* metodama, sve metode daju niži *F1-rezultat*. Presentovane performanse za sve nadgledane metode klasifikacije dobijene su pomoću trostruke unakrsne validacije.

Tabela 18. Rezultati primene nekih postojećih metoda za označavanje pojmova

<b>Metoda klasifikacije</b>	<b><i>F1-rezultat</i></b>
<i>NBMT + POS Tags</i>	0.716255
<i>TreeTagger</i>	0.80409
<i>NB 3-gram klasifikator</i>	0.749155
<i>NB 4-gram klasifikator</i>	0.749936
<i>LR 3-gram klasifikator</i>	0.773655
<i>SMO 3-gram klasifikator</i>	0.776511
<i>SMO 4-gram klasifikator</i>	0.77482
<i>RF 3-gram klasifikator</i>	0.787349
<i>RF 4-gram klasifikator</i>	0.785962

Tabela 19. *F1-rezultati M2, TreeTagger i RF metode za različite kategorije pojmova*

<b>Kategorija termina</b>	<b><i>Metoda M2</i></b>	<b><i>TreeTagger</i></b>	<b><i>RF klasifikator</i></b>
Simptom	0.963895	0,940355	0.933064
Biohemijska analiza	0.950119	0,333333	0.625698
Dijagnoza/bolest	0.934211	0,918103	0.901099
Lek/Terapija	0.791155	0,566327	0.771218
Anatomski organ	0.912621	0,856525	0.937468
Simbol negacije	1	0,99115	0.901763
Broj	0.897315	0,892541	0.558538
Prilog za vreme	0.965347	0,903226	0.856802
Opis simptoma	0.858908	0,857438	0.935256
Nemedicinski termin od manjeg značaja	0.89172	0,56213	0.886491
Specijalnost	0.930788	0,904357	0.886364
Institucija/Mesto	0.722348	0,443946	0.944591
Snizilica	0.990476	0,898305	0.978355
Povisilica	1	0,891566	0.991781
Drugi administrativni medicinski termin	0.753986	0,455006	0.832224
Latinska reč	0.737589	0,268293	0.453258
Stop reč	0.975724	0,9977	0.677451

Tabela 19 prikazuje performanse obeležavanja različitih metoda odvojeno za svaku od proučavanih kategorija pojmova. Tabela 18 prikazuje dve vrste postojećih metoda označavanja, i to metode zasnovane na tagerima i nadgledane metode mašinskog učenja. Od metoda zasnovanih na tagerima, najbolja vrednost za *F1-rezultat* dobija se za metodu *Tree*

*Tagger*. Najbolja vrednost *F1-rezultata* među nadgledanim klasifikatorima dobija se za *RF* klasifikator.

Stoga se rezultati ove dve metode upoređuju sa rezultatima dobijenim korišćenjem predložene metode *M2* (koja pruža najbolje rezultate od svih predloženih metoda). Tabela 19 pokazuje da metoda *M2* daje najveći *F1-rezultat* u većini kategorija.

## 8. ZAKLJUČAK

Pametno zdravstvo je veoma važna komponenta pametnih gradova, i mnoge tehnologije i metode koje se upotrebljavaju u savremenim informacionim sistemima mogu se upotrebiti i za kreiranje servisa pametnog zdravstva. Suština ovih servisa je da se velike količine podataka koje se svakodnevno sakupljaju od građana na adekvatan način obrade i da se iz njih izvuče znanje koje je opet korisno građanima, i direktno utiče na poboljšanje zdravstvenih usluga i generalno kvaliteta života. Stalno povećanje procenta dela ljudske populacije koji živi u gradovima stavlja u fokus upravo način obezbeđenja kvaliteta života. Praktična implementacija koncepta pametnih gradova je usmerena u tom pravcu. Stoga, u konceptu pametnog grada sve je dominantniji deo koji se odnosi na servise pametnog zdravstva. Najveći izazovi u domenu pametnog zdravstva su uvođenje novih servisa, integracija sa postojećim servisima, obezbeđivanje skalabilnosti, efikasno prikupljanje i analiza obilja heterogenih podataka, vizuelizacija podataka i informacija, edukacija stanovništva, obezbeđenje sigurnosti i pouzdanosti podataka, itd. Ova disertacija je upravo pokušaj da se za neke od ovih izazova ponude nova i kvalitetna rešenja i pripreme uslovi za još intenzivniji razvoj servisa pametnog zdravstva i njihovo potpuno uključivanje u koncept pametnog grada, a sve sa ciljem obezbeđenja boljeg kvaliteta života. A to iz dana u dan postaje sve teže ostvarivo zbog stalnog pojavljivanja novih izazova i teškoća kao što je, na primer, aktuelna pandemija virusom COVID-19, koja je otvorila mnoga nova pitanja i pokazala koliko je važan neprestani razvoj koncepta i servisa pametnih gradova.

U ovoj doktorskoj disertaciji predstavljene su osnovne karakteristike pametnog grada i pametnog zdravstva i tehnologije na kojima oni počivaju. Prikazan je pregled servisa u svetu i u Republici Srbiji koji se koriste u ove svrhe. Kako se postavljaju pitanja privatnosti i sigurnosti ovih podataka, i njihova obrada treba da bude usklađena sa zakonskom regulativom, prikazane su preporuke u vidu metoda koje omogućavaju ispunjavanje i zadovoljavanje ovih uslova.

Prema načinu porekla podataka, servisi pametnog zdravstva prikazani u ovoj disertaciji svrstani su četiri grupe:

- servisi zasnovani na podacima prikupljenim sa senzora, odnosno kraudsensingom,
- servisi zasnovani na podacima dobijenim iz direktnog učešća građanina tako što daje svoj opis neke situacije, odgovor, obaveštenje, odnosno na podacima koji su prikupljeni kraudsorsingom,
- servisi zasnovani na podacima iz medicinskih informacionih sistema,
- servisi zasnovani na podacima iz dokumenata iz IKT infrastrukture pametnog grada.

Prvi izazov na koji se naišlo u postupku istraživanja je adaptiranje ovih podataka koji dolaze iz različitih izvora u jednu uniformnu strukturu koja će služiti da se postojeći servisi integrišu u koncept pametnog grada, i na taj način obezbediti skalabilnost ovih servisa. Načini integracije heterogenih servisa pametnog zdravstva i predlog uniformnog veb servisa za realizaciju servisa pametnog zdravstva opisani su u doktorskoj disertaciji, a mogu se primeniti i na ostale servise pametnog grada. Takođe su dati i predlozi i primeri kolaboracije postojećih sistema sa novim servisima pametnih gradova, uz minimalne adaptacije, korišćenjem postojećih rešenja poput sistema poslovne inteligencije i sistema sa mehanizmima notifikacije.

Prema tehnologijama na kojima su zasnovani, servisi pametnog zdravstva prikazani u ovoj doktorskoj disertaciji, svrstani su u pet kategorija:

- servisi zasnovani na *LMS* sistemima i društvenim mrežama,
- servisi zasnovani na *IoT* i geografskim informacionim sistemima,
- servisi koji koriste metode veštačke inteligencije i procesiranja prirodnih jezika,
- servisi zasnovani na multimediji i veb tehnologijama,
- servisi zasnovani na blokčejnu.

Pokazano je da se sistemi za upravljanje učenjem i društvene mreže mogu iskoristiti u cilju prevencije bolesti, informisanja i komunikacije sa stručnim licima. Od servisa baziranih

na senzorskim tehnologijama izdvajaju se i servisi zasnovani na geografskim informacionim sistemima i *IoT*, što nam govori da znajući lokaciju građanina, ustanove ili nekog događaja možemo unaprediti servise pametnog zdravstva. U servisima se može desiti da je potrebno kombinovanje obrade podataka iz različitih izvora, i da su bazirani na više pomenutih tehnologija. Primer za to je servis za praćenje zagađenja vazduha, koji koristi podatke sakupljene kraudsensinogom, kraudsorsingom, a zasnovan je na GIS sistemu. Servisi za prenos multimedije takođe su našli primenu u servisima pametnog zdravstva, kroz obezbeđivanje servisa za telekonsultacije i podršku u realizaciji skrining programa i zbrinjavanju u slučaju velikih nesreća. Primena blokčejn tehnologije sa svojim prednostima za očuvanje pouzdanosti i trajnosti podataka takođe mogu biti od koristi u servisima pametnog zdravstva, za čuvanje osetljivih podataka. Svi ovi pomenuti servisi mogu se na sličan način primenjivati na globalnom nivou. Za razliku od ovih servisa, poseban izazov je kreiranje servisa koji zahtevaju obradu prirodnih jezika, odnosno zasnovani su na metodama veštačke inteligencije.

Posebna pažnja posvećena je servisima baziranim na obradi teksta, a to su servis za automatsko odgovaranje na pitanja, servis za automatsku obradu psiholoških testova i servis za vizuelizaciju podataka o epidemiji. Ovi servisi za ulaz imaju medicinske tekstove ili medicinske izveštaje koji su na srpskom jeziku. Iako se način njihovog kreiranja može ponoviti za druge jezike, ipak su specifičnosti jezika ono što traži ulaganje velikih napora za efikasno rešavanje ovih problema. Mora se posvetiti pažnja sakupljanju i obradi jezičkih resursa. Kako se radi i o obradi izveštaja iz medicinskih informacionih sistema, mora se voditi računa o deidentifikaciji i privatnosti podataka.

Pored strukturiranog dela, medicinski izveštaji uključuju i tekst u slobodnoj formi koji sadrži važne informacije o zdravlju pacijenta. Zbog toga je izuzetno važno obezbediti analizu ove vrste podataka. Srpski jezik je gramatički vrlo složen i zbog toga je veoma izazovan za analizu, pa verovatno zbog toga ne postoje radovi konkretno vezani za analizu teksta iz medicinskog domena na srpskom jeziku, niti javno dostupni elektronski medicinski rečnici sa klasifikovanim medicinskim terminima na srpskom jeziku. Jedan od glavnih doprinosa disertacije je predlog tri metode za obeležavanje teksta u medicinskim izveštajima (zasnovana na rečniku, metoda zasnovana na učenju na trening skupu i hibridna metoda, zasnovana na učenju na trening skupu i pravilima) i njihovo upoređivanje.



Pored predloženih metoda, rezultati istraživanja u okviru ove disertacije uključuju i veliki broj obeleženih i normalizovanih resursa kreiranih za medicinske i nemedicinske termine na srpskom jeziku koji će se koristiti u daljem istraživanju, ali mogu biti od koristi i drugim istraživačima. Takođe, kreirani se ručno označeni korpusi elektronskih medicinskih izveštaja na srpskom jeziku.

Rezultati primene predloženih metoda pokazuju da se bolji rezultati dobijaju kada se označavanje vrši na osnovu učenja nad ručno obeleženim skupom podataka. Dobri rezultati već su postignuti upotrebom korpusa sa oko 2000 označenih izveštaja. Rezultati pokazuju da su greške i skraćenice vrlo česte u medicinskim izveštajima, pa je potrebno stvoriti pravila za njihovu obradu. Za najuspešnije prepoznavanje medicinskih izraza u ovom tekstu, najbolja metoda je hibridna metoda (*M2*) koja koristi označeni korpus kao model podataka i pravila za ispravljanje grešaka i skraćenica. Metoda koja daje najbolje rezultate ne zahteva medicinski rečnik ili morfološku obradu teksta, što je čini manje zahtevnom. Ova predložena metoda može se prilagoditi uz male adaptacije za upotrebu u drugim sličnim jezicima i onima sa složenom gramatikom.

Rezultati metoda su takođe uporedivi sa metodama za označavanje pojmova u drugim jezicima. Prikazani su i rezultati primene standardnih uobičajenih tehnika na skupu podataka i pokazano je da one daju lošije rezultate u odnosu na predložene metode. Dobijeni dobri rezultati mogu podstaći druge istraživače da rade slična istraživanja za svoje jezike, koji su takođe složeni.

Doprinosi ove doktorske disertacije su sledeći:

- dat je predlog arhitekture IS pametnog grada namenjenog realizaciji servisa pametnog zdravstva,
- dat je način integracije heterogenih servisa pametnog grada,
- predložena je arhitektura uniformnog veb servisa za realizaciju servisa pametnog zdravstva,
- data je klasifikacija servisa pametnog zdravstva prema poreklu podataka koje obrađuju i prema tehnologijama na kojima su zasnovani,

- pokazane su primene *LMS* sistema, društvenih mreža, GIS sistema, blokčejn tehnologija, metoda veštačke inteligencije i obrade prirodnih jezika i multimedije i web tehnologija u servisima pametnog zdravstva,
- isprojektovano je četrnaest servisa pametnog zdravstva, od kojih su neki i implementirani [193]. Dati su *UML* dijagrami i načini implementacije servisa pametnog zdravstva:
  - servis za informisanje, komunikaciju i prevenciju bolesti,
  - servis za pretragu i prikaz najbližih zdravstvenih ustanova,
  - servis za praćenje kvaliteta vazduha,
  - servis za unapređenje pružanja medicinskih usluga,
  - servisi za inteligentni transport pacijenata,
  - servis za mobilni *SOS* sistem,
  - servis za automatsko odgovaranje građanima na pitanja i pretragu medicinskih dokumenata,
  - servis za kontrolu epidemije,
  - servis za automatsku obradu upitnika i psiholoških testova,
  - servis za uvid u raspored i ocenjivanje lekara,
  - servis za komunikaciju sa izabranim lekarima i dežurnom ekipom hitne pomoći,
  - servis za podršku u zbrinjavanju u slučajevima velikih nesreća,
  - servis za efikasniju realizaciju skrining programa,
  - servis za obezbeđivanje pouzdanosti i trajnosti podatka.
- prikazani su rezultati primene postojećih metoda za obeležavanje medicinskog teksta na srpskom jeziku,
- kreirane su potpuno nove metode za obradu podataka na srpskom jeziku, a koje pripadaju medicinskom domenu, koje predstavljaju osnovu za kreiranje pojedinih servisa pametnog zdravstva koji obrađuje tekstualne podatke, odnosno podatke prikupljene kraudsorsingom, tekstualne dokumente iz medicinskog domena i elektronske medicinske izveštaje. Ovom prilikom kreirane su sledeće metode:
  - metoda za normalizaciju medicinskih podataka

- metode za obeležavanje termina u medicinskim izveštajima:
  - metoda bazirana na rečniku,
  - metoda nadgledanog mašinskog učenja,
  - hibridna metoda bazirana na nadgledanom mašinskom učenju i pravilima.
- kreirani su resursi na srpskom jeziku koji su potrebni za primenu metoda obeležavanja medicinskog teksta na srpskom jeziku,
- kreirani su obeleženi korpusi elektronskih medicinskih izveštaja na srpskom jeziku.

## 9. PRAVCI DALJEG RAZVOJA

Predloženi servisi pametnog zdravstva mogu se i dalje unapređivati. Iz novih kreiranih podataka može se izvlačiti znanje koje će biti korisno za građane, u pogledu unapređenja njihovog zdravlja i kvaliteta pruženih zdravstvenih usluga. Pored podataka sa senzora za lokaciju, zagađenost vazduha, tekstualnih podataka i glasa, mogu se analizirati i drugi signali poput slike i zvuka, i iskoristiti potencijali njihove obrade u svrhu kreiranja novih servisa pametnog zdravstva. Recimo, obradom većeg broja slika sa karakterističnim osipom može se kreirati alat za prepoznavanje slike koja sadrži taj osip. Obradom zvuka mogu se kreirati servisi za otkrivanje teškog disanja, zatim gužve u saobraćaju, pa se vozila hitne pomoći mogu preusmeriti na drugu rutu itd. Takođe, samo pitanje zaštite podataka je oblast koja nosi puno izazova.

Dobijeni rezultati metoda za obeležavanje termina u medicinskim tekstovima su podsticaj da se nastavi sa radom i, pre svega, da se uloži dodatni napor da se poveća skup analiziranih medicinskih izveštaja kako bi se dobili još bolji rezultati.

Što se tiče dela disertacije, u kome su predloženi algoritmi obrade prirodnog jezika za doprinos efikasnijem istraživanju medicinskih tekstualnih podataka, tu su moguće tri varijante u daljim istraživanjima. Prva od njih je istraživanje mogućnosti povezivanja skupa ličnih skraćenica sa lekarom, budući da se nestrukturirani deo izveštaja koji su napisali različiti lekari razlikuje u broju reči i stilu pisanja. Takođe, ovo je uvod u stvaranje sistema označavanja medicinskih entiteta. Takav sistem će takođe moći da označava medicinske izveštaje i druge tekstove koji sadrže medicinske izraze. Kako se greške i skraćenice mogu otkriti i ispraviti metodom zasnovanom na nadgledanom učenju i pravilima, jedan od pravaca za dalji rad je stvaranje servisa za automatsku korekciju grešaka u medicinskim dokumentima i izveštajima.

Činjenica je da različiti IS arhiviraju ogromne količine različitih podataka koje se realno samo usko ciljano koriste, a nose sa sobom značajano veći potencijal. Primena metoda

veštačke inteligencije nad ovim podacima pravac je u kome će se u budućnosti očekivati najveći rezultati u razvoju koncepta pametnih gradova i od kojih će građani imati velike koristi. Servisi koji će koristiti mogućnosti veštačke inteligencije obrađivaće heterogene podatke iz različitih IS u okviru jednog grada. Kombinacijom podataka iz više sistema mogu se obezbediti značajni servisi, što smo imali prilike da vidimo u okviru servisa vezanih za aktuelnu pandemiju korona virusa.

## Literatura

- [1] S. P. Mohanty, U. Choppali, Elias Kougianos, Everything you wanted to know about smart cities: The internet of things is the backbone, *IEEE Consumer Electronics Magazine*, 2016, 5(3), 60-70.
- [2] A. Avdić, D. Janković, Značaj e-Zdravstva u konceptu pametnih gradova, 17th International Symposium INFOTEH-JAHORINA, 2018, 383-387.
- [3] A. Solanas, C. Patsakis, M. Conti, I. S. Vlachos, V. Ramos, F. Falcone, F., ..., A. Martinez-Balleste, Smart health: a context-aware health paradigm within smart cities, *IEEE Communications Magazine*, 2014, 52(8), 74-81.
- [4] P. D. S. Tavares, E. B. Rodrigues, IoT-Based Architecture for Data Analytics of Arboviruses in Smart Cities, In 2018 IEEE Symposium on Computers and Communications (ISCC), June 2018, 952-957.
- [5] J. C. Kim, K. Chung, Depression index service using knowledge based crowdsourcing in smart health, *Wireless Personal Communications*, 2017, 93(1), 255-268.
- [6] N. A. Adzharuddin, L. H. Ling, Learning management system (LMS) among university students: Does it work, *International Journal of e-Education, e-Business, e-Management and e-Learning*, 2013, 3(3), 248-252.
- [7] F. B. Davies, Object-based geographic information system (GIS), U.S. Patent No. 6,247,019. 12 Jun. 2001.
- [8] M. Mettler, Blockchain technology in healthcare: The revolution starts here. e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on IEEE, 2016, 1-3.
- [9] I. Yoo, P. Alafaireet, M. Marinov, K. Pena-Hernandez, R. Gopidi, J. F., Chang, L. Hua, Data mining in healthcare and biomedicine: a survey of the literature, *Journal of medical systems*, 2012, 36(4), 2431-2448.

- [10] W. Chapman, M. Nadkarni, L. Hirschman, W. D'Avolio, K. Savova, O. Uzuner, Overcoming barriers to NLP for clinical text: the role of shared tasks and the need for additional creative solutions, *J Am Med Inform Assoc*, 2011, 18(5), 540-543.
- [11] J. Dutta, F. Gazi, S. Roy, C. Chowdhury, AirSense: Opportunistic Crowd-Sensing Based Air Quality Monitoring System for Smart City, In 2016 IEEE SENSORS, IEEE, 2016. 1-3.
- [12] S. Fox, M. Duggan, *Mobile health 2010*, Washington, DC, Pew Internet & American Life Project, 2010.
- [13] M. Moraitou, A. Pateli, S. Fotiou, Smart Health Caring Home: A Systematic Review of Smart Home Care for Elders and Chronic Disease Patients, In *GeNeDis*, Springer, 2017, 255-264.
- [14] G. Chen, D. Kotz, *A survey of context-aware mobile computing research*, 2000.
- [15] Our World in Data, <https://ourworldindata.org/urbanization>.
- [16] A. Zanella, N. Bui, A. Castellani, L. Vangelista, M. Zorzi, Internet of things for smart cities, *IEEE Internet of Things journal*, 2014, 1(1), 22-32.
- [17] M. Batty, Big data, smart cities and city planning, *Dialogues in human geography*, 2013, 3(3), 274-279.
- [18] A. Avdić, D. Rančić, Ž. Spalević, J. Achkoski, S. Bojanić, Značaj mobilnog poslovanja u pametnim gradovima, *ZBORNIK RADOVA UNIVERZITETA SINERGIJA*, 2017, 18(3).
- [19] H. Al-Aqrabi, L. Liu, R. Hill, N. Antonopoulos, Cloud BI: Future of business intelligence in the Cloud. *Journal of Computer and System Sciences*, 2015, 81(1), 85-96.
- [20] A. Sebaa, A. Nouicer, A. Tari, R. Tarik, O. Abdellah, Decision support system for health care resources allocation. *Electronic physician*, 2017, 9(6), 4661.
- [21] M. L., Bruce, G. J. McAvay, P. J. Raue, E. L. Brown, B. S. Meyers, D. J. Keohane, ..., C. Weber, Major depression in elderly home health care patients, *American Journal of psychiatry*, 2002, 159(8), 1367-1374.
- [22] Dž. Avdić, A. Avdić, MEDIS upitnik - primena mobilnog računarstva u medicini, *Društvo za informacione sisteme i računarske mreže YUINFO*, 2016, 321-324.

- [23] A. Avdić, U. Marovac, D. Janković, Interaktivna softverska platforma za realizaciju javnih zdravstvenih servisa, 6. Međunarodna konferencija Upravljanje znanjem i informatika, Kopaonik, 13.-14. januar, 2020, 46-52.
- [24] A. Solanas, A. Martinez-Balleste, J. M. Mateo-Sanz, Distributed architecture with double-phase microaggregation for the private sharing of biomedical data in mobile health, *IEEE Transactions on Information Forensics and Security*, 2013, 8(6), 901-910.
- [25] A. Martinez-Balleste, P. A. Pérez-Martínez, A. Solanas, The pursuit of citizens' privacy: a privacy-aware smart city is possible, *IEEE Communications Magazine*, 2013, 51(6), 136-141.
- [26] A. Avdić, D. Janković, Healthcare in Smart Cities – Privacy and Security Issues, Serbia, 5th International Conference Contemporary Problems of Mathematics, Mechanics and Informatics, 17-19 Jun, 2018, Novi Pazar.
- [27] EU General Data Protection Regulation (EU-GDPR), <http://www.privacy-regulation.eu/>.
- [28] GDPR portal, <https://www.eugdpr.org/>.
- [29] Službeni glasnik, <http://www.pravno-informacionisistem.rs/SlGlasnikPortal/.../zakon/2018/87/13/reg>.
- [30] Vebinar: Jeste li usklađeni sa GDPR-om, [https://www.youtube.com/watch?v=RD7TW5\\_6wOY&](https://www.youtube.com/watch?v=RD7TW5_6wOY&)
- [31] Direktiva 95/46/EC Evropskog parlamenta, <http://eur-lex.europa.eu/legal->.
- [32] A. Pajić, D. Bečejski-Vujaklija, Metamodel of the Artifact-Centric Approach to Event Log Extraction from ERP Systems, *International Journal to Decision Support System Technology*, 2016, 8(2), 18-28.
- [33] W. M. Van der Aalst, Extracting event data from databases to unleash process mining, In *BPM-Driving innovation in a digital world*, Springer, Cham, 2015, 105-128.
- [34] N. Badovinac, Osvrt na GDPR uredbu i sugestije za razvoj aplikativne podrške, *InfoM-Časopis za informacione tehnologije i multimedijalne sisteme*, 2018, 2018(65).
- [35] D. Ding, M. Conti, A. Solanas, A smart health application and its related privacy issues, In *Smart City Security and Privacy Workshop*, 2016, 1-5.



- [36] Y. Mehta, M. M. Pai, S. Mallisery, S. Singh, Cloud enabled air quality detection, analysis and prediction-a smart city application for smart health, In Big Data and Smart City (ICBDSC), 3rd MEC International Conference on IEEE, 2016, 1-7.
- [37] T. Adame, A. Bel, A. Carreras, J. Melià-Seguí, M. Oliver, R. Pous, CUIDATS: An RFID–WSN hybrid monitoring system for smart healthcare environments, *Future Generation Computer Systems*, 2018, 78, 602-615.
- [38] K. Natarajan, B. Prasath, P. Kokila, Smart health care system using internet of things, *Journal of Network Communications and Emerging Technologies (JNCET)*, 2016, 6(3), 37-42.
- [39] K. Ullah, M. A. Shah, S. Zhang, Effective ways to use Internet of Things in the field of medical and smart health care, In *Intelligent Systems Engineering (ICISE)*, International Conference on IEEE, 2016, 372-379.
- [40] A. Gomaa, C. Zhang, M. Hasan, M. B. Roche, S. Hynes, Supportive Glucose Sensing Mobile Application to Improve the Accuracy of Continuous Glucose Monitors, In *International Conference on Smart Health*, Springer, 2014, 206-212.
- [41] G. Muhammad, S. M. M. Rahman, A. Alelaiwi, A. Alamri, Smart health solution integrating IoT and cloud: a case study of voice pathology monitoring, *IEEE Communications Magazine*, 2017, 55(1), 69-73.
- [42] G. Sprint, D. J. Cook, R. Fritz, M. Schmitter-Edgecombe, Using smart homes to detect and analyze health events, *Computer*, 2016, 49(11), 29-37.
- [43] S. Takatori, S. Matsumoto, S. Saiki, M. Nakamura, A proposal of cloud-based home network system for multi-vendor services, In *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, 15th IEEE/ACIS International Conference on IEEE, 2014, 1-6.
- [44] H. Jung, K. Chung, Sequential pattern profiling based bio-detection for smart health service, *Cluster Computing*, 2015, 18(1), 209-219.
- [45] R. C. Park, H. Jung, D. K. Shin, G. J. Kim, K. H. Yoon, M2Mbased smart health service for human UI/UX using motion recognition, *Cluster Computing*, 2015, 18(1), 221-232.

- [46] S. Hu, M. Huang, W. Feng, Y. Zhang, A smart health service model for elders based on ECA-S rules, In Software Engineering Research, Management and Applications (SERA), 15th International Conference on IEEE, 2017, 93-97.
- [47] E-Zdravlje, <https://www.e-zdravlje.gov.rs/>.
- [48] Moj Doktor, <https://www.mojdoktor.gov.rs/>.
- [49] Mediatelly Baza Lekova, <https://mediately.co/rs/drugs>
- [50] Otvoreni Podaci, <https://data.gov.rs>.
- [51] A. Milenković, D. Janković, M. Stojković, A. Veljanovski, P. Rajković, Kolaboracija mobilnih senzorskih aplikacija i medicinskog informacionog sistema, INFOTEH-Jahorina, 2014, 13, 879-884.
- [52] HELIANT Health, <https://heliant.rs/>.
- [53] ZipSoft e-Karton, <https://www.zipsoft.rs/>.
- [54] Comtrade, <https://www.comtrade.com/>.
- [55] R. T. Fielding, Architectural Styles and the Design of Network-based Software Architectures, PhD thesis, University of California, Irvine, 2000.
- [56] D. Romero, G. Hermosillo, A. Taherkordi, R. Nzekwa, R. Rouvoy, F. Eliassen, RESTful integration of heterogeneous devices in pervasive environments, In IFIP International Conference on Distributed Applications and Interoperable Systems ,Springer, Berlin, Heidelberg, 2010, 1-14.
- [57] E. Handoyo, M. Arfan, Y. A. A. Soetrisno, M. Somantri, A. Sofwan, E. W. Sinuraya, Ticketing chatbot service using serverless NLP technology. In 2018 5th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE) 2018, 325-330.
- [58] Zapier, <https://zapier.com/app/dashboard>
- [59] W. Rice, H. William, Moodle, Birmingham, Packt publishing, 2006.
- [60] Doktor.rs, [www.doktor.rs](http://www.doktor.rs).
- [61] Stetoskop.info, [www.stetoskop.info](http://www.stetoskop.info).
- [62] M. Nichols, A theory for eLearning, Journal of Educational Technology & Society, 2003, 1-10.
- [63] U. Brandes, L. C. Freeman, D. Wagner, Social networks, 2013, 805-839.

- [64] M. Zuckerberg, Facebook. People Own and Control Their Information, <http://blog.facebook.com/blog.php>.
- [65] A. Avdić, D. Janković, Dž. Avdić, Socijalne mreže u službi e-učenja – primer integracije Moodle LMS-a i Facebook-a, YU INFO Kopaonik, 2015, 36-40.
- [66] Službeni glasnik,  
<http://www.pravno-informacioni-sistem.rs/SlGlasnikPortal/eli/rep/sgrs/skupstina/zakon/2016/15/7/reg>
- [67] Službeni glasnik,  
<http://www.pravno-informacioni-sistem.rs/SlGlasnikPortal/eli/rep/sgrs/skupstina/zakon/2019/25/2>
- [68] M.I. Hosny, S. Fatima, Facebook in Education: Students, Teachers, and Library Perspectives, Journal Of Computing, 2012, 4(6), 78-86.
- [69] S. Kotzer, Y. Elran, Learning and teaching with Moodle-based E-learning environments, combining learning skills and content in the fields of Math and Science & Technology, 1st Moodle ResearchConference, Heraklion, Crete-Greece, 2012, 124-131.
- [70] V. Radosavljević, N. Vugdelija, Društvene mreže kao web 2.0 alat u nastavnom Procesu, INFOTEH-JAHORINA, 2013, 12, 856-859.
- [71] N. Petrovic, V. Jeremic, M. Cirovic, Z. Radojicic, N. Milenkovic, Facebook Vs. Moodle: What Do Students Really Think?, ICICTE 2013 Proceedings, 2013, 413-421.
- [72] A. Hanna, Using Moodle as a dynamic multi-purpose “e-Health Learning Management System”: Tailored patients education, efficient doctor-patient communication, and risk-free professional development for health students.
- [73] Google Maps, <https://developers.google.com/maps/documentation>.
- [74] H. Butler, M. Daly, A. Doyle, S. Gillies, S. Hagen, T. Schaub, The geojson format. Internet Engineering Task Force (IETF), 2016.
- [75] GEOJSON, <https://leafletjs.com/examples/geojson/>.
- [76] E. K. Cromley, S. L. McLafferty, GIS and public health, Guilford Press, 2011.

- [77] E. Kajan, A. Avdić, U. Marovac, A. Ljajić, G.Šimić, J. Stanković, Enhancing local economic development using collective intelligence, In Telecommunications Forum Telfor (TELFOR), 2015, 882-885.
- [78] R. K. Ganti, F. Ye, H. Lei, Mobile crowdsensing: current state and future challenges, IEEE communications Magazine, 2011, 49(11), 32-39.
- [79] AIR NOW, <https://airnow.gov>.
- [80] EPA, <https://www.epa.gov/outdoor-air-quality-data>.
- [81] EPA, Clean Air Act, <https://www.epa.gov/clean-air-act-overview>.
- [82] SEPA, <http://www.amskv.sepa.gov.rs/stanicepodaci.php>.
- [83] N. Shadbolt, Knowledge acquisition and the rise of social machines, International Journal of Human-Computer Studies, 2013, 71(2), 200-205.
- [84] V.S. Revathy, K. Ganesan, K. Rohini, S.T. Chindhu, T. Boobalan, Air pollution monitoring system, IOSR Journal of Electronics and Communication Engineering, 2016, 11(2), 27-40.
- [85] F. Corno, T. Montanaro, C. Migliore, P. Castrogiovanni, SmartBike: an IoT Crowd Sensing Platform for Monitoring City Air Pollution, International Journal of Electrical and Computer Engineering (IJECE), 2017, 7(6), 3602-3612.
- [86] S.R. Garzon, S. Walther, S. Pang, B. Deva, A. Küpper, Urban air pollution alert service for smart cities, In: Proceedings of the 8th International Conference on the Internet of Things, ACM, 2018, p. 9.
- [87] S.R. Khodve, A.N. Kulkarni, Web Based Air Pollution Monitoring System (Air Pollution Monitoring Using Smart Phone), International Journal of Science and Research (IJSR), 2016, 5(3), 266-269.
- [88] A. Alshamsi, Y. Anwar, M. Almulla, M. Aldohoori, N. Hamad, M.R. Awad, Monitoring pollution: Applying IoT to create a smart environment, International Conference on Electrical and Computing Technologies and Applications (ICECTA), 2017, 1-4.
- [89] J. Calbimonte, J. Eberle, K. Aberer, Semantic Data Layers in Air Quality Monitoring for Smarter Cities, In: Proc. of the 6th Workshop on Semantics for Smarter Cities S4SC 2015, at ISWC 2015.

- [90] B. Braem, S. Latré, P. Leroux, P. Demeester, T. Coenen, P. Ballon, Designing a smart city playground: Real-time air quality measurements and visualization in the City of Things testbed, IEEE International Smart Cities Conference (ISC2), 2016, 1-2.
- [91] J. Dutta, C. Chowdhury, S. Roy, A. I. Middy, F. Gazi, Towards Smart City: Sensing Air Quality in City based on Opportunistic Crowd-sensing, In Proceedings of the 18th International Conference on Distributed Computing and Networking ACM, 2017, p. 42.
- [92] Y. Zheng, F. Liu, H. Hsieh, U-Air: when urban air quality inference meets big data, In: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2013, 1436-1444.
- [93] V.A. Buregio, Z. Maamar, S.R. Meira, An architecture and guiding framework for the social enterprise, IEEE Internet Computing, 2015, 19(1), 64-68.
- [94] A. Nigam, N.S. Caswell, Business Artifacts: An Approach to Operational Specification, IBM Systems Journal, 2013, 42(3).
- [95] E. Kajan, V.A. Buregio, Z. Maamar, M. Sellami, E. Ugljanin, Knowledge as a service to support smart cities, ICIST, Dubai, 2017, 17-26.
- [96] A. Avdić, E. Kajan, D. Janković, D., Kontekstno-svesna platforma pametnog zdravstva zasnovana na IoT i Crowdsourcing-u, 18th International Symposium INFOTEH-JAHORINA, 2019, 414-418.
- [97] A. Beaudry, G. Laporte, T. Melo, S. Nickel, Dynamic transportation of patients in hospitals, OR spectrum, 2010, 32(1), 77-107.
- [98] J. Kim, H. H. Jeong, J. P. Cho, W. C. Jeon, K. J. Oh, S. C. Choi, Study on Heuristic Transportation Routes of Patients with Acute Dysbarism for the Best Prognosis, Journal of the Korean Society of Emergency Medicine, 2016, 27(1), 118-125.
- [99] Y. Kergosien, C. Lente, D. Piton, J. C. Billaut, A tabu search heuristic for the dynamic transportation of patients between care units, European Journal of Operational Research, 2011, 214(2), 442-452.
- [100] Č. Vasić, B. Predić, D. Rančić, P. Spalević, Dž. Avdić, Dynamic Relocation of Emergency Ambulance Vehicles Using the AVL Component of the GPS/GPRS Tracking System, Acta Polytechnica Hungarica, 2014, 11(9), 39-59.

- [101] Z. Zhang, M. Liu, A. Lim, A memetic algorithm for the patient transportation problem, *Omega*, 2015, 54, 60-71.
- [102] I. M. Hains, A. Marks, A. Georgiou, J. I. Westbrook, Non-emergency patient transport: what are the quality and safety issues? A systematic review, *International Journal for Quality in Health Care*, 2010, 23(1), 68-75.
- [103] N. Theodore, B. Aarabi, S. S. Dhall, D. E. Gelb, R. J. Hurlbert, C. J. Rozzelle, M. N. Hadley, Transportation of patients with acute traumatic cervical spine injuries, *Neurosurgery*, 2013, 72(3), 35-39.
- [104] Ž. Jovanović, A. Avdić, D. Janković, Transportation in Smart Cities - Tracking and Improving Driving Comfort, In: Konjović, Z., Zdravković, M., Trajanović, M. (Eds.) *ICIST 2018 Proceedings*, 2018, 2, 346-349.
- [105] Ž. Jovanović, D. Janković, A. Peulić, Mobilni sistem za određivanje nivoa udobnosti u vozilu, *INFOTEH-JAHORINA*, 2016, 15, 442-446.
- [106] Z. Jovanovic, R. Bacevic, R. Markovic, S. Randjic, D. Jankovic, Information system for generating road comfort maps, In *Proceedings of the 60th International Conference ETRAN*, Zlatibor, Serbia, 2016, RT 5.6.
- [107] Z. Jovanovic, R. Bacevic, R. Markovic, S. Randjic, D. Jankovic, Information System for the Centralized Display of the Transport Comfort Information, *Facta Universitatis, Series: Electronics and Energetics*, 2017, 30(4), 557-570.
- [108] Z. Jovanovic, R. Bacevic, R. Markovic, S. Randjic, Android application for observing data streams from built-in sensors using RxJava, In *Proceedings of the 23rd Telecommunication Forum Telfor*, Belgrade, Serbia, 2015, 918-921.
- [109] Ž. Jovanović, A. Avdić, D. Janković, D. Vujičić, Smart Transportation in the Service of Improving Healthcare in Smart Cities, 2018, *IcETLAN*.
- [110] Ž. Jovanović, *Primena IT u unapređenju kvaliteta transporta pacijenata*, doktorska disertacija, Elektronski fakultet, Univerzitet u Nišu, 2020.
- [111] A. Avdić, Dž. Avdić, U. Marovac, E. Kajan, A. Ljajić, A concept of efficient parking in smart cities, In *Proceedings of the 25th Telecommunication Forum Telfor*, Belgrade, Serbia, 2017, 1-4.
- [112] S. Russell, P. Norvig, *Artificial intelligence: a modern approach*, 2002.

- [113] D. J. Hand, N. M. Adams, Data mining. Wiley StatsRef: Statistics Reference Online, 2014, 1-7.
- [114] M. W. Berry, J. Kogan, Text mining. Applications and Theory. West Sussex, PO19 8SQ, UK: John Wiley & Sons, 2010.
- [115] G. G. Chowdhury, Natural language processing. Annual review of information science and technology, 2003, 37(1), 51-89.
- [116] R. Rosales, Method for Automatic Labeling of Unstructured Data Fragments from Electronic Medical Records. U.S. Patent Application, 2009, 12/469,745.
- [117] A. Pljasković, D. Avdić, U. Marovac, A. Crnišanin, D. Rančić, Pretraživanje dokumenata na srpskom jeziku za potrebe m-Uprave, ETRAN, 2013, RT4.6.
- [118] S. Ghosh, S. K. Dubey, Comparative analysis of k-means and fuzzy c-means algorithms. International Journal of Advanced Computer Science and Applications, 2013, 4.4.
- [119] U. Marovac, E. Kajan, G. Šimić, A solution of semantic clustering of text documents, CPPMI 2012, Novi Pazar, jun 2012.
- [120] B. Li, L. Han, Distance weighted cosine similarity measure for text classification, In International Conference on Intelligent Data Engineering and Automated Learning, Springer, Berlin, Heidelberg, 2013, 611-618.
- [121] Speech to Text, <https://cloud.google.com/speech-to-text>.
- [122] M. Al-Azzam, M. B. Alazzam, Smart city and smart-health framework, challenges and opportunities, Int. J. Adv. Comput. Sci. Appl, 2019, 10(2), 171-176.
- [123] COVID19, <https://covid19.rs/>.
- [124] Worldometers, <https://www.worldometers.info/coronavirus/>.
- [125] End Coronavirus, <https://www.endcoronavirus.org/countries>.
- [126] A. M. H. Pardede, Framework for Patient Service Queue System for Decision Support System on Smart Health Care, 2018.
- [127] A. Avdić, E. Kajan, D. Janković, Dž. Avdić, Towards Context-Aware Smart Healthcare Platform, International Journal of Electrical Engineering and Computing, 2019, 3(1), 26-31.

- [128] J. Wang, X. Wang, J. Wu, Inferring Metapopulation Propagation Network for Intra-City Epidemic Control and Prevention, In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, 830-838.
- [129] G. Pan, G. Qi, W. Zhang, S. Li, Z. Wu, L. T. Yang, Trace Analysis and Mining for Smart Cities: Issues, Methods, and Applications, IEEE Communications Magazine, 2013, 51(6), 120-126.
- [130] S. Rani, S. H. Ahmed, S. C. Shah, Smart Health: A Novel Paradigm to Control the Chickungunya Virus, IEEE Internet of Things Journal, 2018, 6 (2), 1306-1311.
- [131] R. K. R. Kummitha, Smart Technologies for Fighting Pandemics: The Techno - and Human-Driven Approaches in Controlling the Virus Transmission, Government Information Quarterly, 101481, 2020.
- [132] A. M. Milenkovic, P. Rajkovic, T. Stankovic, D. S. Janković, Application of Medical Information System MEDIS.NET in Professional Learning, 19th Telecommunications Forum (TELFOR) Proceedings of Papers, 2011, 1474-1477.
- [133] A. Avdić, U. Marovac, D. Janković, Automated Labeling of Terms in Medical Reports in Serbian, Turkish Journal of Electrical Engineering & Computer Sciences, 2020, 28(6), 3285-3303.
- [134] A. Avdić, U. Marovac, D. Janković, Normalization of Health Records in the Serbian Language with the Aim of Smart Health Services Realization, Facta Universitatis, Series Mathematics and Informatics, 2020, 35(3), 825-841.
- [135] Google Charts, <https://developers.google.com/chart>.
- [136] A. Avdić, U. Marovac, D. Janković, Smart health Services for Epidemic Control. In: 2020 55th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), IEEE, 2020. 46-49.
- [137] Y. Huh, D. H. Ahn, J. H. Choi, J. Y. Kang, Y. Y. Kim, K. J. Oh, Development of a childproblem-behavior screening test, Journal of Korean Neuropsychiatric Association, 2003, 6(42), 724-735.
- [138] S. Y. Bhang, H. K. Yoo, J. H. Kim, B. Kim, G. H. Bahn, D. Ahn, Revision of adolescent mental health and problem behavior screening questionnaire: Development of adolescentmental health and problem behavior screening questionnaire-II, Journal of Korean Academy of ChildAdolescent Psychiatry, 2011, 22(4), 271–286.



- [139] S. Cohen, T. Kamarck, R. Mermelstein, A global measure of perceived stress. *Journal of Health and Social Behavior*, 1983, 24(4), 385–396.
- [140] A. Osman, C. L. Bagge, P. M. Gutierrez, L. C. Konick, B. A. Kopper, F. X. Barrios, The suicidal behaviors questionnaire-revised (SBQ-R): Validation with clinical and nonclinical samples. *Assessment*, 2001, 8(4), 443–454.
- [141] J. L. Cox, J. M. Holden, R. Sagovsky, Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry*, 1987, 150, 782-786.
- [142] A. Fazlagić, Psihosocijalni korelati depresivnosti u postpartalnom periodu. Doktorska disertacija, Univerzitet u Istočnom Sarajevu, Filozofski fakultet.
- [143] U. Marovac, A. Pljaskovic, A. Crnisanin, E. Kajan, N-gram analysis of text documents in Serbian language. In *Telecommunications Forum (TELFOR)*, 2012, 1385-1388.
- [144] U. Marovac, A. Ljajić, A. Avdić, A. Fazlagić, Automation of psychological testing of stressful situations in the Serbian, In: Trajanović, M., Zdravković, M., Konjović, Z. (Eds.) *ICIST 2019 Proceedings*, 2019, 1, 102-106.
- [145] Smart Band For Seniors With SOS function ST-01, <https://canyon.eu/product/cne-st01bb/>.
- [146] S. Memon, K. Khownbati, S.R. Hussain, Internet based multimedia services and technologies in the context of e-government: A conceptual framework. *Information Technology Journal*, 2007, 16(1), 903-908.
- [147] Skrining Srbija, <http://www.skriningsrbija.rs/srl/opste-informacije-o-skriningu/>.
- [148] N. Satoshi, Bitcoin: A Peer-to-Peer Electronic Cash System, 2008.
- [149] G. Zyskind, O. Nathan, Decentralizing privacy: Using blockchain to protect personal data. *Security and Privacy Workshops (SPW)*, 2015, 180-184.
- [150] A. Ekblaw, A. Azaria, J.D. Halamka, A. Lippman, A Case Study for Blockchain in Healthcare: “MedRec” prototype for electronic health records and medical research data. In *Proceedings of IEEE open & big data conference*, 2016, 13, 13.
- [151] L.A. Linn, M. B. Koo, Blockchain for health data and its potential use in health it and health care related research, *ONC/NIST Use of Blockchain for Healthcare and Research Workshop*. Gaithersburg, Maryland, United States: ONC/NIST, 2016.

- [152] S. Angraal, H. M. Krumholz, W. L. Schulz, Blockchain technology: applications in health care, *Circulation: Cardiovascular Quality and Outcomes*, 2017, 10.
- [153] K. Biswas, V. Muthukkumarasamy, Securing smart cities using blockchain technology. High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS), 2016 IEEE 18th International Conference on. IEEE, 2016, 1392-1393.
- [154] P. Zhang, D. C. Schmidt, J. White, G. Lenz, Blockchain technology use cases in healthcare, In *Advances in computers*, 2018, 111, 1-41.
- [155] A. Ljajić, U. Marovac, A. Avdić, Optimizacija dokumenata za njihovu bržu pretragu, Nove informacione tehnologije za analitičko odlučivanje u biološkim, ekonomskim i sociološkim sistemima, 2014, 68-88.
- [156] S. M. Meystre, G. K. Savova, K. C. Kipper-Schuler, J. F. Hurdle, Extracting information from textual documents in the electronic health record: a review of recent research, *Yearbook of medical informatics*, 2008, 17(1), 128-144.
- [157] H. Dalianis, Characteristics of Patient Records and Clinical Corpora. In: *Clinical Text Mining*, Springer, Cham, 2018.
- [158] W. Sun, Z. Cai, Y. Li, F. Liu, S. Fang, G. Wang, Data processing and text mining technologies on electronic medical records: a review, *Journal of healthcare engineering*, 2018, 2018, 1-10.
- [159] Y. Wang, Z. Yu, Y. Jiang, Automatic symptom name normalization in clinical records of traditional Chinese medicine, *BMC Bioinformatics*, 2010, 11, 40.
- [160] G. K. Savova, J. J. Masanz, P. V. Ogren, J. Zheng, S. Sohn, K. C. Kipper-Schuler, C. G. Chute, Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications, *Journal of the American Medical Informatics Association*, 2010, 17, 5, 507-513.
- [161] A. Ljajić, U. Marovac, M. Stanković, Comparison of the influence of different normalization methods on tweet sentiment analysis in Serbian language, *Facta Universitatis (NIS) Ser. Math. Inform.* 2018, 33, (5), 683-696.

- [162] P. Rajković, D. Janković, D. Vučković, Adaptation and Application of Daitch – Mokotoff SoundEx Algorithm on Serbian Names, Conf. PRIM (book of abstracts), 2006, 21.
- [163] P. Rajković, D. Janković, D. Vucković, Using String Comparison Algorithms for Serbian Names, Proceedings XLI International scientific conference on Information, communication and energy systems and technologies – ICEST, 2006, 221-224.
- [164] V. Vincze, G. Szarvas, R. Farkas, G. Móra, J. Csirik, The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes, BMC Bioinformatics, 2008, 9(11), S9.
- [165] C. Ehrentraut, H. Tanushi, J. Tiedemann, H. Dalianis, Detection of hospital acquired infections in sparse and noisy Swedish patient records, In Proceedings of the Sixth Workshop on Analytics for Noisy Unstructured Text Data (AND 2012) Held in Conjunction with Coling 2012, Bombay, ACM Digital Library.
- [166] A. Ljajić, U. Marovac, Improving Sentiment Analysis for Twitter Data by Handling Negation Rules in the Serbian Language, Computer Science and Information Systems, 2019, 16(1), 289-311.
- [167] ICD-10, <https://www.who.int/standards/classifications/classification-of-diseases>.
- [168] Međunarodna Statistička Klasifikacija Bolesti i Srodnih Zdravstvenih Problema, Deseta revizija, editor Dr Miljan Ljubičić, Institut za javno zdravlje Srbije “Dr Milan Jovanović Batut”, edition 2010, 2013, 1.
- [169] M. Buckley, B. Coopey, J. Sharko, F. Polubriaginof, B. Drohan, K. Belli, ..., C. Specht, The feasibility of using natural language processing to extract clinical information from breast pathology reports. Journal of pathology informatics, 2012, 3, 23.
- [170] S. Boytcheva, G. Angelova, Z. Angelov, D. Tcharaktchiev, Text mining and big data analytics for retrospective analysis of clinical texts from outpatient care, Cybernetics and Information Technologies, 2015, 15(4), 58-77.
- [171] I. Alimova, E. Tutubalina, Multiple features for clinical relation extraction: A machine learning approach. Journal of Biomedical Informatics, 2020, 103, 103382.

- [172] J. Gorinski, H. Wu, C. Grover, R. Tobin, C. Talbot, Named entity recognition for electronic health records: A Comparison of Rule-based and Machine Learning Approaches, 2019.
- [173] V. Garla, L. Re, Z. Dorey-Stein, F. Kidwai, M. Scotch, J. Womack, ..., C. Brandt, The Yale cTAKES extensions for document classification: architecture and application; *Journal of the American Medical Informatics Association*, 2011, 18(5), 614-620.
- [174] E. Soysal, J. Wang, M. Jiang, Y. Wu, S. Pakhomov, H. Liu, H. Xu, CLAMP - a toolkit for efficiently building customized clinical natural language processing pipelines, *Journal of the American Medical Informatics Association*, 2017, 25(3), 331-336.
- [175] D. L. MacLean, H. Jeffrey, Identifying medical terms in patient-authored text: a crowdsourcing-based approach, *Journal of the American Medical Informatics Association*, 2013, 20(6), 1120-1127.
- [176] H. Lai, M. Topaz, R. Goss, L. Zhou, Automated misspelling detection and correction in clinical free-text records, *Journal of biomedical informatics*, 2015, 55, 188-195.
- [177] S. Antonic, C. Krstev, Serbian Wordnet for biomedical sciences, In *INFORUM*; Prague, Czech Republic, 2008, 28-30.
- [178] C. Krstev, I. Obradović, M. Utvić, D. Vitas, A system for named entity recognition based on local grammars, *Journal of Logic and Computation*, 2014, 24(2), 473-489.
- [179] P. Gantar, D. Štrkalj, S. Krek, N. Ljubešić, Towards semantic role labeling in Slovene and Croatian. In *Proceedings of the Conference on Language Technologies Digital Humanities*, Ljubljana, Slovenia, 2018, 92-98.
- [180] A. Avdić, U. Marovac, D. Janković, Dž. Avdić, Normalization of medical records written in Serbian. In: Konjović, Z., Zdravković, M., Trajanović, M. (Eds.) *ICIST 2019 Proceedings*, Kopaonik, Serbia, 2019, 1, 72-75.
- [181] Z. Popović, Taggers applied on texts in Serbian. *INFOtheca - Journal of Informatics & Librarianship*, 2010, 11(2).
- [182] B. Šandrih, C. Krstev, R. Stankovic, Development and evaluation of three Named Entity Recognition systems for Serbian - The case of personal names. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing RANLP*, Varna, Bulgaria, 2019, 1060-1068.

- [183] K. Toutanova, D. Klein, C. Manning, Y. Singer, Feature-rich Part-of-Speech tagging with a cyclic dependency network, In Proceedings of HLT-NAACL, Edmonton, Canada, 2003, 252-259.
- [184] H. Schmid, Improvements in Part-of-Speech tagging with an application to German. Proceedings of the ACL SIGDAT-Workshop, Dublin, Ireland, 1995.
- [185] J. L. Fleiss, Measuring nominal scale agreement among many raters, Psychological bulletin, 1971, 76(5), 378.
- [186] C. Krstev, D. Vitas, S. Gucul, Recognition of personal names in Serbian texts. International Conference Recent Advances in Natural Language Processing, Borovets, Bulgaria, 2005, 288-292.
- [187] D. Đurić-Petković, E. Ristanović, N. Kuljić-Kapulica, Virus malih boginja, MD-Medical Data, 2017, 9(3), 181-184.
- [188] S. Krstić, Kliničke karakteristike malih boginja kod dece lečene u službi za pedijatriju Opšte bolnice Leskovac, Apollinem medicum et aesculapium, 9.
- [189] N. Milošević, Stemmer for the Serbian language, arXiv 1209.4471, 2012.
- [190] C. Krstev, R. Stankovic, D. Vitas, Knowledge and rule-based diacritic restoration in Serbian, In Proceedings of the Third International Conference Computational Linguistics, Sofia, Bulgaria, CLIB-2018, 41-51.
- [191] J. Rice, Mathematical Statistics and Data Analysis(3rd ed.). USA, Duxbury Advanced, 2006.
- [192] Weka, <https://www.cs.waikato.ac.nz/ml/weka/>.
- [193] Servisi pametnog zdravstva (demo verzija), <https://www.shealth.tagstudio.rs/servisi/>.

## SPISAK KORIŠĆENIH SKRAĆENICA

Skraćenica	Engleski termin	Srpski termin / značenje
API	<i>Application Programming Interface</i>	programski interfejs aplikacije
ATC	-	anatomsko-terapijsko-hemijska klasifikacija
BI	<i>Business Intelligence</i>	poslovna inteligencija
CSV	<i>Comma Separated Values</i>	vrednosti razdvojene zarezom
DZ	-	Dom zdravlja
EHR	<i>Electronic health record</i>	elektronski zdravstveni izveštaj
EMR	<i>Electronic medical record</i>	elektronski medicinski izveštaj
EMS	<i>Emergency medical services</i>	hitni medicinski servisi
EPA	<i>Environmental Protection Agency</i>	Agencija za zaštitu sredine
EPDS	<i>Edinburgh Postnatal Depression Scale</i>	Edinburška skala postpartalne depresije
e-zdravstvo	<i>e-Health</i>	elektronsko zdravstvo
FB	<i>Facebook</i>	Fejsbuk
GDPR	<i>General Data Protection Regulation</i>	Opšta uredba o zaštiti podataka
GIS	<i>Geographic Information System</i>	geografski informacioni sistem

GPS	<i>Global Positioning System</i>	globalni pozicioni sistem
HTTP	<i>Hypertext Transfer Protocol</i>	protokol za prenos hiperteksta
h-zdravstvo	<i>h-Health (home health)</i>	kućna nega
ICD	<i>International Statistical Classification of Diseases and Related Health Problems</i>	Međunarodna statistička klasifikacija bolesti i srodnih zdravstvenih problema
IDF	<i>Inverse Document Frequency</i>	inverzna učestalost u dokumentu
IKT	<i>Information and Communication Technologies</i>	informacione i komunikacione tehnologije
IKV	<i>AQI (Air Quality Index)</i>	indeks kvaliteta vazduha
IoT	<i>Internet of Things</i>	Internet stvari
IS	<i>Information system</i>	informacioni sistem
ISO	<i>International Organization for Standardization</i>	međunarodna organizacija za standardizaciju
JMBG	-	jedinstveni matični broj građana
JSON	<i>JavaScript Object Notation</i>	-
LB	<i>Location Based</i>	zasnovan na lokaciji
LBO	-	lični broj osiguranika
LMS	<i>Learning Management System</i>	sistem za upravljanje učenjem
LR	<i>Logistic Regression</i>	logistička regresija

Spisak korišćenih skraćenica

M2M	<i>Machine to Machine</i>	-
MAC	<i>Media Access Control</i>	kontrola pristupa mediju
MIS	<i>Medical Information System</i>	medicinski informacioni sistem
m-zdravstvo	<i>m-Health (mobile health)</i>	mobilno zdravstvo
NB	<i>Naïve Bayes</i>	naivni Bajes
NBMT	<i>Naïve Bayes Multinomial Text</i>	Naivni Bajes multiterminski tekst
NER	<i>Named Entity Recognition</i>	prepoznavanje imenovanih entiteta
NLP	<i>Natural Language Processing</i>	procesiranje prirodnih jezika
PM10	<i>Particulate matter</i>	Čestice
POS	<i>Part of Speech</i>	delovi govora
p-zdravstvo	<i>s-Health (smart health)</i>	pametno zdravstvo
RE	<i>Relation Extraction</i>	izdvajanje odnosa
REST	<i>REpresentational State Transfer</i>	prenos stanja reprezentacije
RF	<i>Random Forest</i>	metoda slučajnih šuma
RFID	<i>Radio Frequency Identification</i>	identifikacija radio frekvencijom
RFZO	-	Republički fond za zdravstveno osiguranje
RMS	<i>Root Mean Square</i>	kvadratna sredina



SEPA	<i>Serbian Environmental Protection Agency</i>	Agencija za zaštitu sredine Srbije
SMO	<i>Sequential Minimal Optimization</i>	sekvencijalna minimalna optimizacija
SMS	<i>Short Message Service</i>	sistem za slanje kratkih poruka
TF	<i>Term Frequency</i>	učestalost termina
UML	<i>Unified Modeling Language</i>	uniformni jezik za modelovanje
URL	<i>Uniform Resource Locator</i>	uniformni lokator resursa
WSN	<i>Wireless Sensor Network</i>	bežična senzorska mreža
XML	<i>eXtensible Markup Language</i>	proširivi jezik za označavanje

## SPISAK SLIKA

Slika 1. Procenat stanovništva u gradovima (preuzeto iz [15]).....	17
Slika 2. Komponente, atributi, teme i infrastruktura pametnih gradova .....	18
Slika 3. Komponente pametnih gradova .....	19
Slika 4. Primer čuvanja zdravstvenih podataka u kubovima.....	21
Slika 5. Odnos mobilnog i pametnog zdravstva.....	24
Slika 6. Postojeći servisi na portalu e-Zdravlje .....	32
Slika 7. Prikaz ekrana postojećih servisa <i>MojDoktor.rs</i> i <i>Mediatelly</i> .....	33
Slika 8. <i>Medis.NET</i> informacioni sistem.....	34
Slika 9. Slojevita arhitektura IS koji objedinjuje heterogene servise pametnog zdravstva.....	36
Slika 10. Struktura uniformnog veb servisa za realizaciju servisa pametnog zdravstva.....	40
Slika 11. Primer dela <i>JSON</i> -a koji sadrži deo elektronskih izveštaja iz MIS .....	40
Slika 12. Primer povezivanja dva različita sistema korišćenjem alata <i>Zapier</i> .....	42
Slika 13. Primer razgovora pacijenata o bolesti na <i>FB</i> grupi.....	44
Slika 14. Primer informisanja o radu doktora na <i>FB</i> grupi .....	44
Slika 15. Post na forumu <i>Doktor.rs</i> .....	45
Slika 16. Dijagram slučajeva korišćenja servisa za informisanje i prevenciju bolesti.....	48
Slika 17. Koraci u kolaboraciji <i>Moodle</i> sistema i <i>FB</i> aplikacije .....	49
Slika 18. Primer geografskih podataka u <i>GeoJSON</i> formatu .....	51
Slika 19. Dijagram slučajeva korišćenja za servis za pretragu i prikaz najbližih zdravstvenih ustanova i apoteka .....	52
Slika 20. Izbor apoteka i bolnica na <i>Google</i> mapama.....	52
Slika 21. Dodatne informacije o izabranoj bolnici.....	53
Slika 22. Slojevita arhitektura servisa za praćenje zagađenja vazduha.....	57
Slika 23. Dijagram slučajeva korišćenja za servis za praćenje kvaliteta vazduha .....	59

Slika 24. Predlog izgleda mape zagađenja i vizuelizacija vrednosti zagađivača .....	60
Slika 25. Vizuelizacija rešenja korišćenjem dostupnog <i>AQP API</i> -ja .....	60
Slika 26. Dijagram slučajeva korišćenja servisa za unapređenje medicinskih usluga .....	62
Slika 27. Merenje vibracija korišćenjem <i>VibroDroid</i> aplikacije .....	66
Slika 28. Prikaz jedne kreirane mape udobnosti .....	67
Slika 29. Arhitektura rešenja aplikacije za pametno parkiranje .....	68
Slika 30. Prikaz dela <i>JSON</i> podataka o parkinzima .....	69
Slika 31. Prikaz ekrana aplikacije .....	69
Slika 32. Prikaz ekrana aplikacije – moguće rute .....	70
Slika 33. Primer nosivog uređaja za svrhe pametnog zdravstva [145] .....	71
Slika 34. Dijagram slučajeva korišćenja za servis za automatsko odgovaranje građanima na pitanja .....	74
Slika 35. Pojednostavljeni prikaz komponenti sistema za pretraživanje dokumenata .....	75
Slika 36. Prikaz realizacije servisa za pretragu dokumenata / odgovaranje na pitanja .....	79
Slika 37. Servis za kontrolu epidemija .....	82
Slika 38. Dijagram slučajeva korišćenja servisa za kontrolu epidemija .....	83
Slika 39. Pregled obolelih po mesecima .....	85
Slika 40. Pregled obolelih po zdravstvenim ustanovama .....	85
Slika 41. Pregled obolelih po godinama starosti .....	86
Slika 42. Pregled zastupljenosti simptoma po broju pacijenata .....	86
Slika 43. Deo <i>Testa samoprocene</i> na portalu <i>e-Zdravlje</i> .....	88
Slika 44. Detaljna arhitektura rešenja .....	90
Slika 45. Prikaz realizacije upitnika .....	92
Slika 46. Traženje preporuke lekara .....	93
Slika 47. Traženje informacija o lekaru .....	94
Slika 48. Slučajevi korišćenja servisa za uvid u raspored i ocenjivanje lekara .....	94
Slika 49. Dijagram slučajeva korišćenja za servis za komunikaciju sa lekarom ili hitnom službom .....	96
Slika 50. Dijagram slučajeva korišćenja za podršku u zbrinjavanju u slučaju velikih nesreća .....	97

Slika 51. Dijagram slučajeva korišćenja servisa za efikasniju realizaciju skrining programa.	99
Slika 52. Snimanje transakcije u bitcoin blokčejnu .....	100
Slika 53. Struktura blokčejna i bloka u lancu.....	101
Slika 54. Struktura blokčejna koji sadrži podatke iz <i>EMR</i> -a.....	103
Slika 55. Normalizacija medicinskih izveštaja.....	111
Slika 56. Statistika u kontroli epidemija .....	118
Slika 57. Koraci obrade u metodi <i>M0</i> .....	131
Slika 58. Koraci obrade u metodi <i>M1</i> .....	132
Slika 59. Koraci obrade u metodi <i>M2</i> .....	133
Slika 60. Primeri obeležene anamneze primenom metode <i>M2</i> .....	134

## SPISAK TABELA

Tabela 1. Usklađivanje aktivnosti nad podacima sa <i>GDPR</i> .....	29
Tabela 2. Zaštita podataka šifriranjem i maskiranjem [34].....	30
Tabela 3. Uticaj indeksa kvaliteta vazduha na zdravlje .....	54
Tabela 4. Primer medicinskog izveštaja iz korišćenog skupa podataka.....	109
Tabela 5. Primer nepotpunih rečenica u anamnezi.....	109
Tabela 6. Primer anamneze sa negacijom .....	111
Tabela 7. Neke stop reči iz rečnika .....	113
Tabela 8. Anamneza nakon obrade skraćenica i negacije .....	113
Tabela 9. Rezultati različitih metoda normalizacije .....	115
Tabela 10. Primer normalizovanih anamneza .....	116
Tabela 11. Procenat pojavljivanja za najčešćih ključnih reči.....	117
Tabela 12. Medicinske stop reči koje se najčešće pojavljuju.....	118
Tabela 13. Rezultati označavanja medicinskih i nemedicinskih termina.....	136
Tabela 14. Rezultati primene metoda nad medicinskim terminima.....	137
Tabela 15. Statistički značaj poboljšanja metode <i>M2</i> u odnosu na druge metode .....	138
Tabela 16. Statistički značaj razlika u označavanju pojmova .....	138
Tabela 17. Rezultati nadgledanih metoda nad kontrolnim skupom podataka.....	139
Tabela 18. Rezultati primene nekih postojećih metoda za označavanje pojmova .....	140
Tabela 19. <i>F1</i> -rezultati <i>M2</i> , <i>TreeTagger</i> i <i>RF</i> metode za različite kategorije pojmova.....	140

## **BIOGRAFIJA AUTORA**

Aldina Avdić je rođena 11.06.1987. godine u Prijepolju, Republika Srbija. Osnovnu školu i gimnaziju završila je u Prijepolju.

Prvu godinu osnovnih akademskih studija na Elektronskom fakultetu u Nišu, smer Računarska tehnika, odsek Softversko inženjerstvo upisala je 2006. godine. Diplomirala je 7. jula 2011. godine sa prosečnom ocenom 9,52 i ocenom 10,00 na završnom diplomskom radu pod nazivom „Razvoj mobilne aplikacije za udaljeni pristup *Moodle LMS*-u“, čime je stekla zvanje Diplomirani inženjer elektrotehnike i računarstva. Za ovaj diplomski rad dodeljena joj je povelja Elektronskog fakulteta „Najbolje izrađen diplomski rad u školskoj 2010/2011. godini“. Iste godine je nagrađena je stipendijom Fonda za mlade talente. Doktorske studije upisala je 2011. godine, na smeru Elektrotehnika i računarstvo na Elektronskom fakultetu u Nišu.

Od 2011. god. zaposlena je na Državnom univerzitetu u Novom Pazaru, prvo kao saradnik u naučno-istraživačkom centru, a zatim kao saradnik u nastavi. Trenutno je asistent za užu naučnu oblast Računarska tehnika na Departmanu za tehničke nauke. Njeni istraživački interesi odnose se na servise pametnog zdravstva i pametnih gradova, obradu prirodnog jezika (NLP), s posebnim naglaskom na obradu slobodnog teksta u okviru elektronskih medicinskih izveštaja. Autori/koautor je preko 50 publikacija na domaćim i međunarodnim konferencijama i časopisima u oblasti obrade prirodnog jezika i servisa pametnih gradova. Udata je, majka dvoje dece.

## IZJAVA O AUTORSTVU

Izjavljujem da je doktorska disertacija, pod naslovom

### **REALIZACIJA SERVISA PAMETNOG ZDRAVSTVA I NJIHOVA INTEGRACIJA U KONCEPT PAMETNIH GRADOVA**

koja je odbranjena na fakultetu Univerziteta u Nišu:

- rezultat sopstvenog istraživačkog rada;
- da ovu disertaciju, ni u celini, niti u delovima, nisam prijavljivala na drugim fakultetima, niti univerzitetima;
- da nisam povredila autorska prava, niti zloupotrebio/la intelektualnu svojinu drugih lica.

Dozvoljavam da se objave moji lični podaci, koji su u vezi sa autorstvom i dobijanjem akademskog zvanja doktora nauka, kao što su ime i prezime, godina i mesto rođenja i datum odbrane rada, i to u katalogu Biblioteke, Digitalnom repozitorijumu Univerziteta u Nišu, kao i u publikacijama Univerziteta u Nišu.

U Nišu, \_\_\_\_\_

Potpis autora disertacije:

*Aldina R. Avdić*

Aldina R. Avdić

**IZJAVA O ISTOVETNOSTI ELEKTRONSKOG I ŠTAMPANOG OBLIKA  
DOKTORSKE DISERTACIJE**

Naslov disertacije:

**REALIZACIJA SERVISA PAMETNOG ZDRAVSTVA I NJIHOVA INTEGRACIJA U  
KONCEPT PAMETNIH GRADOVA**

Izjavljujem da je elektronski oblik moje doktorske disertacije, koju sam predala za unošenje u **Digitalni repozitorijum Univerziteta u Nišu**, istovetan štampanom obliku.

U Nišu, \_\_\_\_\_

Potpis autora disertacije:

Aldina R. Avdić

Aldina R. Avdić



## IZJAVA O KORIŠĆENJU

Ovlašćujem Univerzitetsku biblioteku „Nikola Tesla“ da u Digitalni repozitorijum Univerziteta u Nišu unese moju doktorsku disertaciju, pod naslovom:

### **REALIZACIJA SERVISA PAMETNOG ZDRAVSTVA I NJIHOVA INTEGRACIJA U KONCEPT PAMETNIH GRADOVA**

Disertaciju sa svim prilogima predala sam u elektronskom obliku, pogodnom za trajno arhiviranje.

Moju doktorsku disertaciju, unetu u Digitalni repozitorijum Univerziteta u Nišu, mogu koristiti svi koji poštuju odredbe sadržane u odabranom tipu licence Kreativne zajednice (Creative Commons), za koju sam se odlučila.

1. Autorstvo (CC BY)

2. Autorstvo – nekomercijalno (CC BY-NC)

3. Autorstvo – nekomercijalno – bez prerade (CC BY-NC-ND)

4. Autorstvo – nekomercijalno – deliti pod istim uslovima (CC BY-NC-SA)

5. Autorstvo – bez prerade (CC BY-ND)

6. Autorstvo – deliti pod istim uslovima (CC BY-SA)

U Nišu, \_\_\_\_\_

Potpis autora disertacije:

Aldina R. Avdić

Aldina R. Avdić